Synthetic and Systems Biotechnology 9 (2024) 828–833

Contents lists available at ScienceDirect

Synthetic and Systems Biotechnology

journal homepage: www.keaipublishing.com/en/journals/synthetic-and-systems-biotechnology

Original Research Article

Heterologous activation and metabolites identification of the *pks7* gene cluster from *Saccharopolyspora erythraea*

Hao Tang [a,1], Xingchi Yang [b,1], Wenzong Wang [a], Xingjun Cui [a], Wenping Wei [a], Jing Wu [a], Peng Sun [b,**], Bang-Ce Ye [a,c,*]

[a] Institute of Engineering Biology and Health, Collaborative Innovation Center of Yangtze River Delta Region Green Pharmaceuticals, College of Pharmaceutical Sciences, Zhejiang University of Technology, Hangzhou, 310014, China
[b] School of Pharmacy, Naval Medical University, Shanghai, 200433, China
[c] State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, Shanghai, 200237, China

ARTICLE INFO

Keywords:
Heterologous expression
Saccharopolyspora erythraea
6-Methylsalicyclic acid
6-Ethylsalicyclic acid
6-Ethylsalicyclic acid ester derivatives
Flavonoid

ABSTRACT

The microbial genome remains a huge treasure trove for the discovery of diverse natural products. *Saccharopolyspora erythraea* NRRL23338, the industry producer of erythromycin, has a dozen of biosynthetic gene clusters whose encoding products are unidentified. Heterologous expression of one of the polyketide clusters *pks7* in *Streptomyces albus* B4 chassis resulted in the characterization of its function responsible for synthesizing both 6-methylsalicyclic acid and 6-ethylsalicyclic acid. Meanwhile, two new 6-ethylsalicyclic acid ester derivatives were isolated as shunt metabolites. Their structures were identified by comprehensive analysis of MS and NMR experiments. Putative functions of genes within the *pks7* BGC were also discussed.

1. Introduction

Actinomycetes have long been an important source of natural products aiming at clinical drug discovery [1,2]. Among them, the classical antibiotic-producing *Streptomyces* and *Saccharopolyspora* strains have not been fully explored and most of the secondary metabolites remain unknown. Genome sequencing and bioinformatics analyses have revealed that microbial genome is a huge treasure trove for finding novel natural products [3]. The genome mining approach has greatly enhanced the discovery of natural products from Actinomycetes. Genome sequencing of *Saccharopolyspora erythraea* NRRL23338, the producer of erythromycin, revealed the presence of a dozen of biosynthetic gene clusters (BGCs) for secondary metabolites besides the erythromycin gene cluster (*ery*). The hypothetical products of these BGCs remained mostly unidentified [4].

In the current work, we focused on the *pks7* BGC, which contains a core biosynthetic gene for encoding partial reduced type-I iterative polyketide synthase (PR-PKS). Herein, we have characterized the function of the *pks7* BGC through heterologous expression using the *Streptomyces albus* B4 chassis (B4). The metabolic products including 6-methylsalicyclic acid, 6-ethylsalicyclic acid, and two new 6-ethylsalicyclic acid ester derivatives were identified.

2. Materials and methods

2.1. General experimental equipment

The NMR spectra were recorded at 300 K on a Bruker AVANCE 600 NMR spectrometer (Germany). Chemical shifts are reported in parts per million (δ), with use of the residual $CDCl_3$ signals ($\delta_H = 7.26$ ppm; $\delta_C = 77.16$ ppm) as internal standard, and coupling constants (J) are in Hz; assignments are supported by COSY, HSQC, HMBC, and NOESY experiments. HRMS data were obtained using a Bruker APEXIII 7.0 T FT-MS spectrometer in m/z, resolution 5000; an isopropyl alcohol solution of sodium iodide (2 mg/mL) was used as a reference. Semi-preparative HPLC was performed on an Agilent 1100 system (USA) equipped with

Peer review under responsibility of KeAi Communications Co., Ltd.
* Corresponding author. Institute of Engineering Biology and Health, Collaborative Innovation Center of Yangtze River Delta Region Green Pharmaceuticals, College of Pharmaceutical Sciences, Zhejiang University of Technology, Hangzhou, 310014, China.
** Corresponding author.
E-mail addresses: sunpeng78@126.com (P. Sun), yebangce@zjut.edu.cn (B.-C. Ye).
[1] The authors contribute equally.

https://doi.org/10.1016/j.synbio.2024.05.004
Received 15 December 2023; Received in revised form 28 April 2024; Accepted 7 May 2024
Available online 9 May 2024

a UV detector using a YMC-Pack ODS-A column (5 μm, 250 × 10 mm; YMC, Japan) or a Shimadzu system (Japan) with a UV detector using a Positisil ODS-P column (5 μm, 250 × 10 mm; China). Silica gel (200−300 mesh; Yantai) was used for column chromatography. Precoated Silica gel plates (HSGF254; Yantai) were used for TLC. Compounds were detected on TLC under UV light or by heating after spraying with anisaldehyde−sulfuric acid reagent.

2.2. Bacterial strains, plasmids and fermentation conditions

Escherichia coli Trans5α (TransGen Biotech, Beijing, China) was used as cloning host and *E. coli* ET12567 (pUZ8002) as donor strain for *E. coli-S. albus* conjugation [5]. The collection number of *S. erythraea* NRRL23338 is DSM 40517 [4]. *S. albus* B4 was used as the chassis for heterologous expression of *pks7* [6]. The pSET152 plasmid was used to integrate the *pks7* sequence into B4 genome [7]. Apramycin (50 mg/mL) was used for selection in *E. coli*. For *E. coli* ET12567 (pUZ8002), additional kanamycin (50 mg/mL) and chloramphenicol (25 mg/mL) were utilized. Apramycin (50 mg/mL) and nalidixic acid (25 mg/mL) were utilized for exconjugants selection. Antibiotics were added to the medium at a ratio of 1:1000. *E. coli* strains were grown in LB liquid media (5 g/L yeast extract, 10 g/L tryptone, and 10 g/L NaCl) or maintained on LB agar plate containing corresponding antibiotics. *S. erythraea* NRRL23338 was grown in TSB (30 g/L tryptone soy broth) liquid media for the preparation of genomic DNA as PCR template. B4 and its recombinant strains were grown in TSB liquid media, while mannitol soy flour (MS) solid media (20 g/L soybean powder, 20 g/L mannitol, 20 g/L agar and pH adjusted to 7.5) was used to prepare spores. MS solid media and SFM liquid media (20 g/L glucose, 0.5 g/L $K_2HPO_4{\cdot}3H_2O$, 0.5 g/L $MgSO_4{\cdot}7H_2O$, 4 g/L yeast powder, 4 g/L tryptone, 0.01 g/L $FeSO_4{\cdot}7H_2O$, 1 g/L KNO_3 and pH adjusted to 7.2) were used for fermentation [8]. For liquid fermentation, the spores of B4 and its derivatives were activated in seed media (45 g/L tryptone soy broth, 10 g/L glucose, 9 g/L yeast extract and 2.2 g/L $MgSO_4{\cdot}7H_2O$) at 30 °C for 36 h [8], and then transferred into SFM liquid media. All liquid fermentations were carried out at 220 rpm and 30 °C for 5 days. For solid fermentation, B4 and its derivatives were cultivated on MS agar plate at 30 °C for 5 days.

2.3. DNA isolation, manipulation, and E. Coli-S. albus conjugation

Purification of genomic DNA from *S. erythraea* employed a TIANamp Bacteria DNA kit (Tiangen Biotech, China). Plasmids were isolated from *E. coli* using Easy Pure Plasmid Miniprep Kit (TransGen Biotech, China). Using Phanta Max Super-Fidelity DNA polymerase (Vazyme Biotech, China), polymerase chain reaction (PCR) was carried out in a Mastercycler Thermocycler (Eppendorf, Germany). Primers in this study were all synthesized by Sangon Biotech (China). DNA fragments were purified using a gel extraction kit (TransGen Biotech, China). Restriction endonucleases were acquired from Thermo Fisher Scientific (USA). Hieff Clone Multi One Step Cloning Kit (Yeasen, China) was used to insert PCR amplicons into linear plasmid. Sequencing was performed by GENEWIZ (China).

E. coli-S. albus conjugation was performed in accordance with the procedure described previously [9]. B4 spores were incubated at 37 °C in a shaker at 220 rpm for 3 h after undergoing heatshock treatment at 50 °C for 10 min. The mixture of spores and *E. coli* ET12567 donors at various ratios was then cultured on MS plates with 20 mM $MgCl_2$ added as an additive. The plates were covered with 1 mL of H_2O holding apramycin and nalidixic acid after being incubated for 14–16 h at 30 °C. Following incubation lasting 3–5 days at 30 °C, the resultant exconjugants were detected and picked into TSB liquid medium with apramycin.

2.4. Construction of recombinant strains

pSET152 was digested with XbaI and BamHI for linearization. The *pks7* coding sequence was amplified from *S. erythraea* genome as disassembled into three parts flanked by homologous sequences with adjacent DNA fragments (Fig. S1). Part 1 was obtained by two rounds of PCR by amplification successively using the primer pairs (forward primer 5′-CAGCTAGCTCAGTCCTAGGTATAATGCTAGCGTGCGTGCTTCCGAAACGTTC-3′ and reverse primer 5′-TTGATGGTCGCCATCGTGCCCTTGCCTTCAGCACGCCACA-3′) and (forward primer 5′-TTGGGCTGCAGGTCGACTCTAGATTGACAGCTAGCTCAGTCCTAGGTATAATGCT-3′ and reverse primer 5′-TTGATGGTCGCCATCGTGCCCTTGCCTTCAGCACGCCACA-3′) to make *pks7* under the control of Pj23119. Part 2 was amplified using the primer pair (forward primer 5′-CGTGCTGAAGGCAAGGGCACGATGGCGACCATCAA-3′ and reverse primer 5′- TCGATCTCGTCTACGCGCACGTCAATTCCTCCTCATCTCAGAACCACGGGTACCT-3′), while Part 3 was amplified using the primer pair (forward primer 5′-CGTGGTTCTGAGATGAGGAGGAATTGACGTGCGCGTAGACGAGATCGA-3′ and reverse primer 5′- ATATCGCGCGCGGCCGCGGATCCCTATTTCGCTTCGCTCGACAGG-3′). The fragments were then inserted into linear plasmid. Correct colony with assembly of Pj23119 promoter and *pks7* was confirmed by PCR using M13 primer pairs (forward primer 5′-GTAAAACGACGGCCAGT-3′ and reverse primer 5′-CAGGAAACAGCTATGAC-3′) and sequencing (Fig. S2). Considering the rearrangement of genes within *pks7*, the ribosome-binding site sequence was inserted between *SACE_5309* and *SACE_5306* using homology arm primers (5′-TCGATCTCGTCTACGCGCACGTCAATTCCTCCTCATCTCAGAACCACGGGTACCT-3′ and 5′-CGTGGTTCTGAGATGAGGAGGAATTGACGTGCGCGTAGACGAGATCGA-3′). The resulting assembly was then transformed into B4 by intergeneric conjugation and the B4-*pks7* strain was thus obtained.

B4-*pks7*-5306KO, the *SACE_5306* deletion mutant, was constructed in a similar way as B4-*pks7*. Part 4 was amplified using the primer pair (5′-CGTGCTGAAGGCAAGGGCACGATGGCGACCATCAA -3′ and 5′-ATATCGCGCGCGGCCGCGGATCCTCATCTCAGAACCACGGGTACCT -3′). Parts 1 and 4 were then inserted into linear plasmid. PCR using M13 primer pairs and sequencing confirmed successful construction of the assembly of Pj23119 promoter and *pks7*-5306KO (Fig. S2). The resulting assembly was then transformed into B4 by intergeneric conjugation and the B4-*pks7*-5306KO strain was thus obtained.

2.5. HPLC analysis of the metabolites of recombinant strains

After cultivation for 5 days in SFM liquid media or on MS agar plate, the spores and freeze-dried fermentation broth were extracted with methanol under ultrasonic processing for 10–15 min. The mixture was centrifuged at 10,700×*g*, 10 min at room temperature and the supernatant was then processed with a 0.22 μm syringe filter. For HPLC analysis, detections were carried out on Agilent 1260 system, using a Diamonsil Plus C18-A analytical column (5 μm, 180 × 4.6 mm; DiKMA, China) with a flow rate of 1.0 mL/min. The crude extract was injected into HPLC using a 20.0 μL volume and detected at wavelengths of 210, 260, and 300 nm. Mobile phase A was water with 0.1 % trifluoroacetic acid and mobile phase B was acetonitrile with 0.1 % trifluoroacetic acid. Method 1: 0–20 min, 5–30 % B; 20–30 min, 30–100 % B; 30–34 min, 100 % B; 34–35 min, 100-5% B and 35–40 min, 5 % B. Method 2: 0–15 min, 5–50 % B; 15–20 min, 50–100 % B; 20–24 min, 100 % B; 24–25 min, 100-5% B and 25–30 min, 5 % B.

2.6. Isolation of compounds **1**–**4**

The culture broths of strain B4-*pks7* with SFM liquid media (10 L) were extracted with EtOAc (3 × 5 L), and the organic extract was concentrated under reduced pressure to afford a residue (1.5 g). The residue was subjected to semi-preparative HPLC purification to give compounds **1** (1.5 mg, t_R = 20.5 min) and **2** (1.0 mg, t_R = 22.5 min). Semi-preparative HPLC was carried out on a Shimadzu system equipped with a UV detector using a Positisil ODS-P column (5 μm, 250 × 10 mm). Mobile phases A (water with 0.1 % trifluoroacetic acid) and B (acetonitrile with 0.1 % trifluoroacetic acid) were used at a flow rate of 3.0

mL/min with the following gradient: 0–20 min, 5–30 % B; 20–30 min, 30–100 % B; 30–34 min, 100 % B; 34–35 min, 100-5% B and 35–40 min, 5 % B.

The spores harvested from 300 MS plates were extracted with EtOAc (3 × 5 L), and the organic extract was concentrated under reduced pressure to afford a residue (0.8 g). The residue was subjected to HPLC purification (75 % MeOH, 2.5 mL/min) to give compounds **3** (1.6 mg, t_R = 29 min) and **4** (1.0 mg, t_R = 44 min). Semi-preparative HPLC was carried out on an Agilent 1100 system equipped with a UV detector using a YMC ODS-A column (5 μm, 250 × 10 mm).

2.7. Physical and spectroscopic data of compounds 1–4

Compound **1**: white amorphous powder; HRESIMS *m/z*: 153.0607 ($[M + H]^+$, calcd 153.0552 for $C_8H_9O_3$, error 0.55 ppm); R_f 0.61 (CH_2Cl_2/MeOH 15:1).

Compound **2**: white amorphous powder; HRESIMS *m/z*: 167.0697 ($[M + H]^+$, calcd 167.0708 for $C_9H_{11}O_3$, error −0.11 ppm); R_f 0.63 (CH_2Cl_2/MeOH 15:1).

Compound **3**: white amorphous powder; ^1H and ^{13}C NMR data see Table 1; HRESIMS *m/z*: 403.1158 $[M + H]^+$ (calcd 403.1182 for $C_{24}H_{19}O_6$, error −0.24 ppm); R_f 0.45 (CH_2Cl_2/MeOH 15:1).

Compound **4**: white amorphous powder; ^1H and ^{13}C NMR data see Table 1; HRESIMS *m/z*: 419.1144 $[M + H]^+$ (calcd 419.1131 for $C_{24}H_{19}O_7$, error 0.11 ppm); R_f 0.40 (CH_2Cl_2/MeOH 15:1).

3. Results and discussion

3.1. Bioinformatic analysis of pks7 BGC from S. erythraea

The *S. erythraea* genome possesses at least 25 BGCs for secondary metabolites biosynthesis including eleven polyketide genes [4]. Besides the identified BGCs of *ery*, *rpp* [10], and *pks8* [11], the products of eight polyketide BGCs (*pks1-pks7*, and *pke* [12]) remains unknown [4]. The genome of *S. erythraea* (NCBI accession number AM420293) was re-analyzed bioinformatically using antiSMASH [13]. The *pks7* gene cluster was found within a large BGC that contained 36 genes spanning ~46 kb. Referencing to gene annotations from the KEGG database [14], we analyzed each gene in this cluster (Table S1). *SACE_5308* is the core biosynthetic gene, and is annotated as 6-methylsalicyclic acid synthase.

Table 1
The ^1H NMR and ^{13}C NMR data for compounds **3** and **4**[a].

No.	3		4	
	δ_C, type	δ_H (*J* in Hz)	δ_C, type	δ_H (*J* in Hz)
2	153.8, CH	8.41, s	154.9, CH	8.43, s
3	122.7, C		121.4, C	
4	174.3, C		179.5, C	
5	127.1, CH	7.94, d (8.7)	159.1, C	
6	116.1, CH	6.92, d (8.7)	99.7, CH	6.19, s
7	164.6, C		161.9, C	
8	102.2, CH	6.84, s	94.2, CH	6.36, s
9	157.8, C		157.8, C	
10	115.7, C		103.7, C	
1′	130.1, C		128.9, C	
2′, 6′	130.3, CH	7.67, d (8.6)	130.4, CH	7.68, d (8.5)
3′, 5′	121.4, CH	7.27, d (8.6)	121.6,CH	7.28, d (8.5)
4′	150.1, C		150.2, C	
1″	120.4, C		120.4,C	
2″	155.0, C		154.8, C	
3″	113.5, CH	6.85, d (8.0)	113.4, CH	6.82, d (8.0)
4″	131.1, CH	7.23, t (8.0)	131.1, CH	7.25, t (8.0)
5″	118.9, CH	6.77, d (8.0)	119.0, CH	6.79, d (8.0)
6″	142.2, C		142.3, C	
7″	166.8, C		166.7, C	
8″	26.2, CH_2	2.66, q (7.6)	26.2, CH_2	2.66, q (7.6)
9″	15.7, CH_3	1.21, t (7.6)	15.7, CH_3	1.20, t (7.6)

[a] In DMSO-d_6; 600 MHz for ^1H and 150 MHz for ^{13}C.

The rest genes can be classified as additional biosynthesis, transporters, regulators and others. Therefore, it is unlikely that this cluster can produce an antibiotic backbone that uses 6-methylsalicyclic acid as a moiety like chlorothricin [15], and we tentatively decided to focus on the four genes (*SACE_5306-5309*) in *pks7*, based on the analysis by Oliynyk et al. [4]. *SACE_5306-5308* are conserved in some *Streptomyces* strains, although the gene arrangements differ (Fig. S3) [15–17]. SACE_5308 showed high sequence homology to ChlB1 and MdpB (Fig. 1A), which are responsible for production of 6-methylsalicyclic acid in chlorothricin and maduropeptin biosynthesis, respectively (Fig. 1B). The protein SACE_5306 was homologous to ChlB2, a putative discrete ACP protein while SACE_5307 was homologous to ChlB3, which was named 3-oxoacyl-ACP synthase.

3.2. Heterologous expression of the pks7 BGC in B4

To discover the function of *pks7* BGC from *S. erythraea*, heterologous expression strategy was conducted with B4 by inserting of a strong constitutive promoter Pj23119 before the rearranged *pks7* cluster (Fig. 1C) [18]. The recombinant strains B4-*pks7*, B4-*pks7*-5306KO and the host strain B4 were cultivated in SFM liquid medium and MS agar plate, respectively. The metabolite profiles of the three strains were compared (Fig. 2). Two additional metabolites (**1** and **2**) appeared in B4-*pks7* and B4-*pks7*-5306KO fermented with SFM liquid media. The fermentation of strains B4-*pks7* and B4-*pks7*-5306KO with MS agar plate media resulted in the production of another two compounds (**3**–**4**) (Fig. 3). The four compounds (**1**–**4**) were successfully isolated and purified by chromatographic separations. The above results indicated that *pks7* was activated both in B4-*pks7* and B4-*pks7*-5306KO strains.

3.3. Structure elucidation of 1–4

Compound **1**, with a molecular formula $C_8H_8O_3$, was determined using HRESIMS with a positive ion at *m/z* 153.0607 ($[M + H]^+$, calcd 153.0552 for $C_8H_9O_3$) (Fig. S4). The structure of **1** was identified to be 6-methylsalicyclic acid by comparing its NMR spectroscopic data with those reported in literatures (Fig. S4) [19]. Compound **1** has previously been reported as an alarm or trail pheromone in ants [20] and as a bacterial metabolite [16]. Considering the sequence homology between SACE_5308 and the bacterial PKS responsible for synthesizing 6-methylsalicyclic acid, **1** was undoubtedly the direct product of SACE_5308 (Fig. 1A and B).

Compound **2** has a molecular formula of $C_9H_{10}O_3$, which was determined by HRESIMS with a positive ion at *m/z* 167.0697 ($[M + H]^+$, calcd 167.0708 for $C_9H_{11}O_3$) (Fig. S5). Compound **2** was identified to be 6-ethylsalicyclic acid by comparing its NMR spectroscopic data with those reported in early literature [21]. Previously, compound **2** was reported as a catalytic product of the 6-methylsalicyclic acid synthetase from *Penicillium patulum* [22] and as a bacterial metabolite from *Mycobacterium phlei* [21]. On the basis of the proposed function of SACE_5308, **2** was also the direct product of SACE_5308 (Fig. 1A and B).

Compound **3** was obtained an amorphous powder. The molecular formula was assigned to be $C_{24}H_{18}O_6$ on the basis of HRESIMS data (*m/z* obsd 403.1158 $[M + H]^+$, calcd 403.1182 for $C_{24}H_{19}O_6$) (Fig. S6). The ^1H NMR spectrum displayed the presence of a series of aromatic protons and an ethyl group (δ_H 1.21, t, J = 7.6 Hz, H-9″; 2.26, q, J = 7.6 Hz, H-8″) (Table 1). The ^{13}C NMR and distortionless enhancement by polarization transfer (DEPT) spectra showed 24 carbon signals that corresponded to 22 sp^2 (11 CH and 11 C) and 2 sp^3 carbon atoms (1 CH_3 and 1 CH_2). These data were typical of a flavonoid skeleton. Comprehensive analysis of 1D and 2D NMR experiments allowed the structural assignment of **3** (Fig. S6). A set of AMX type aromatic proton resonances at δ_H 7.94 (1H, d, J = 8.7 Hz, H-5), 6.92 (1H, d, J = 8.7 Hz, H-6) and 6.84 (1H, s, H-8) were assigned to A ring, which was supported by the HMBC cross-peaks from H-5 to C-4 and C-7, and from H-6 to C-8 and C-10. The aromatic singlet at δ_H 8.41 (1H, s, H-2) combined with the HMBC cross-

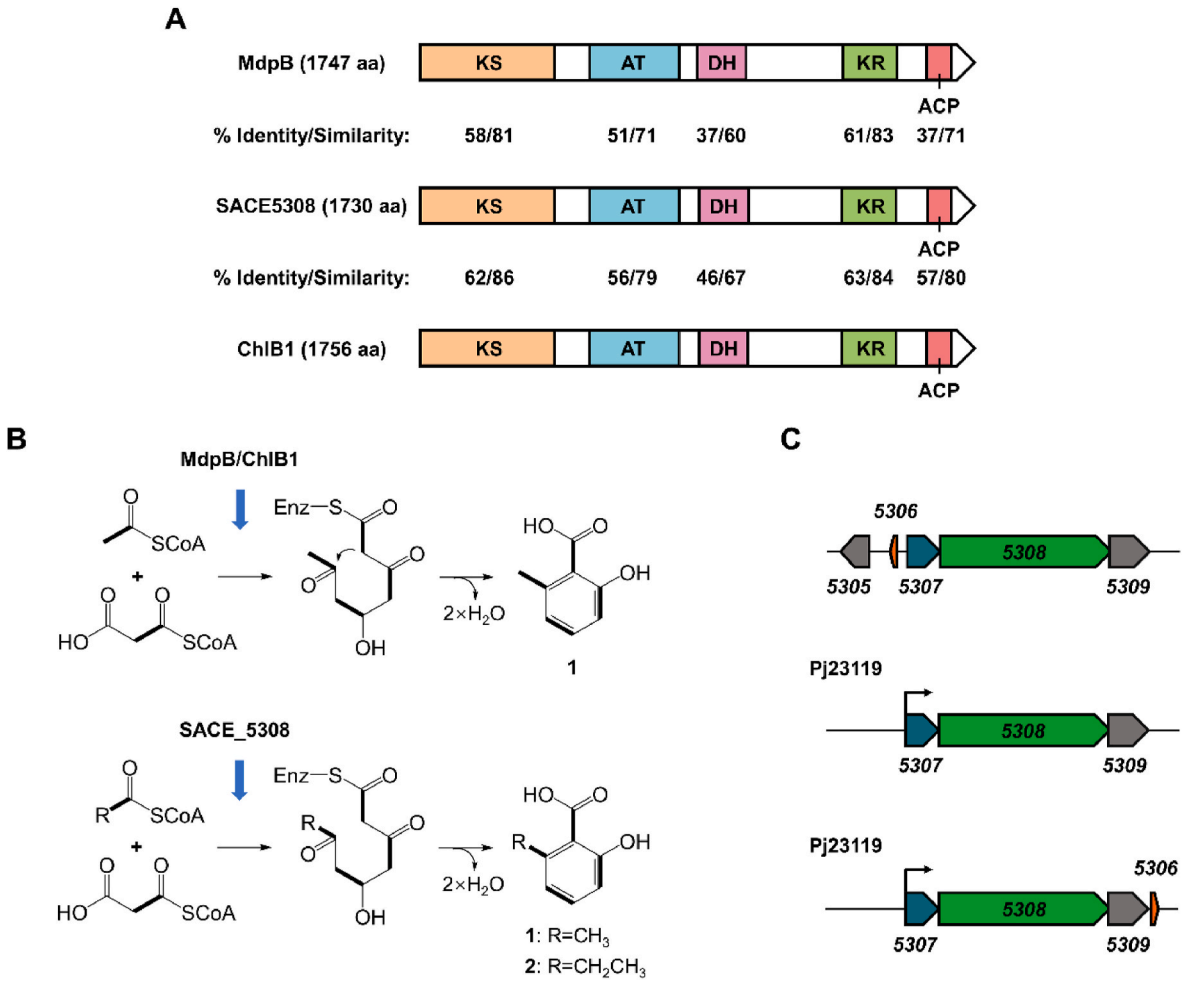

Fig. 1. Comparisons of SACE_5308 with selected bacterial PR-PKSs and organization of *pks7*. **A**) Domain organization of SACE_5308 and comparisons of domains between SACE_5308 and homologous proteins. **B**) Proposed biosynthetic pathways catalyzed by MdpB/ChlB1 and SACE_5308. **C**) Genetic organization of native *pks7* in *S. erythraea*, artificial *pks7* (*SACE_5307-5309*) in B4-*pks7*-5306KO, and artificial *pks7* (*SACE_5306-5309*) in B4-*pks7*.

peaks from H-2 to C4, C-9, and C-1′ suggested an isoflavonoid type ring B structure. The ^{1}H spectrum also displayed signals for AA'BB' type aromatic proton resonances at δ_H 7.67 (2H, d, J = 8.6 Hz, H-2′,6′) and 7.27 (2H, d, J = 8.6 Hz, H-3′,5′), indicating characteristics of *p*-substituted ring C structure. Additionally, a 1,2,3-trisubstituted benzene as ring D was determined by the ^{1}H NMR spectrum δ_H 6.85 (1H, d, J = 8.0 Hz, H-3″), 7.23 (1H, t, J = 8.0 Hz, H-4″), and 6.77 (1H, d, J = 8.0 Hz, H-5″), and by ^{1}H–^{1}H COSY spectrum. The HMBC cross-peaks from H-3″ to C-1″ and C-7″, from H-4″ to C-2″ and C-6″, from H_2-8″ to C-1″ and C-5″, and from H_3-9″ to C-6″ established an 6″-ethylsalicyclic acid unit (Fig. 4). The structure of **3** shows a high similarity to daidzein, an isoflavonoids widely distributed in plant [23,24]. Careful comparison of NMR data of **3** with those of daidzein revealed that the ^{13}C resonance of C-4′ was shielded (δ_C 150.1 in **3** and 157.2 in daidzein), while those of C-3'/5′ was deshielded (δ_C 121.4 in **3** and 115.0 in daidzein). These results indicated the 6″-ethylsalicyclic acid unit was connected at the hydroxy group of C-4'. Therefore, **3** was determined to be daidzein-4'-(6″-ethylsalicylate).

Compound **4**, as an amorphous powder, has a molecular formula of $C_{24}H_{18}O_7$ as determined by HRESIMS data (*m/z* obsd 419.1144 $[M+H]^+$, calcd 419.1131 for $C_{24}H_{19}O_7$) (Fig. S7). The molecular weight of **4** is 16 mass unit more than that of **3**, suggesting the presence of an additional oxygen atom. The 1D NMR spectra of **4** were highly close to that of **3**, indicating a similar conjugate of isoflavonoid and 6-ethylasliclic acid. The difference in ^{1}H spectrum was observed for the two singlets for aromatic proton resonances at δ_H 6.19 (1H, s, H-6) and 6.36 (1H, s, H-8). The ^{13}C and DEPT spectra revealed the presence of an additional oxygenated quaternary carbon (δ_C 159.1, C-5) and two shielded aromatic methines (δ_C 99.7, C-5; 94.2, C-8). These phenomena were consistent with a 5-OH in ring A, similar to the NMR data of genistein [24,25]. Full analysis of the 1D and 2D NMR experiments revealed that the remaining structure was the same as that of **3** (Fig. S7). The comparison of NMR data of **4** with those of genistein [24,25] confirmed the connection of 6″-ethylsalicyclic acid unit with the hydroxy group of C-4'. Compound **4** was then assigned to be genistein-4'-(6″-ethylsalicylate).

In this work, *pks7* from *S. erythraea* was successfully identified by heterologous expression, which is responsible for synthesizing **1** or **2**. The core gene *SACE_5308* acts as both 6-methylsalicyclic acid and 6-ethylsalicyclic acid synthase. The proteins encoded by a pair of genes, *SACE_5306* and *SACE_5307*, were annotated as putative discrete ACP protein and 3-oxoacyl-ACP synthase, respectively. It has been proposed that **1** underwent activation by SACE_5307 and then loaded onto SACE_5306 for further decoration. Removal of *SACE_5306* did not influence the production of **1**–**4** in the B4-*pks7*-5306KO strain (Fig. 2). It was suggested that the P450 gene *SACE_*5309 might be involved in the online modification of **1** or **2**. However, no related products were detected. The hidden function of SACE_5309 may require further experimental proof.

The formation of **1** and **2** might be attributed to the robustness of AT domain of SACE_5308 for accepting either acetyl-CoA or propionyl-CoA as substrates. Further experiments, including in vitro biochemical assays, are required to elucidate the mechanism of the formation of **1** and

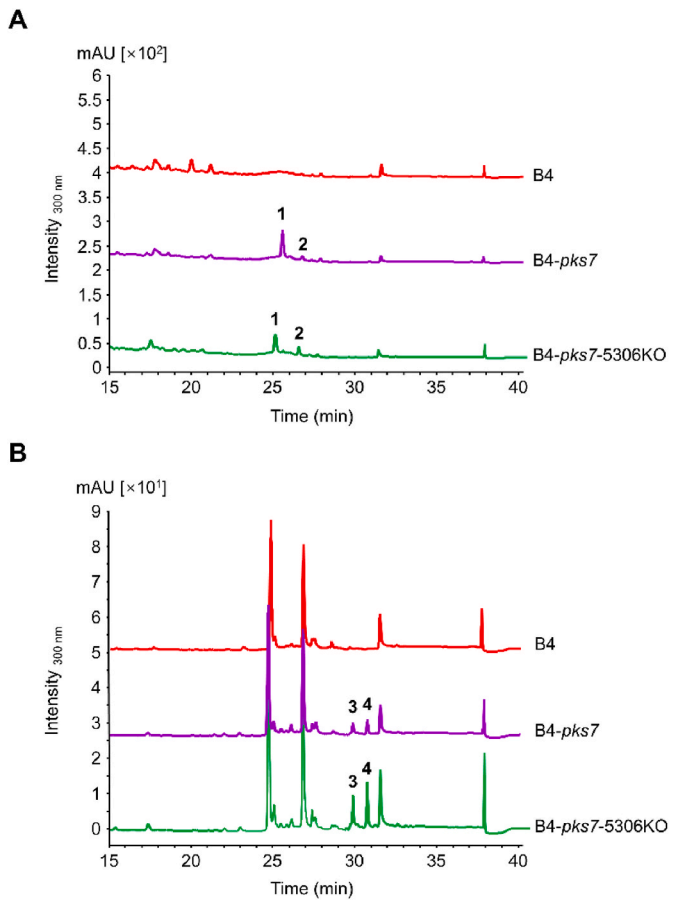

Fig. 2. HPLC analysis of secondary metabolites in B4 and its recombinant strains. **A** HPLC chromatograms, monitored at 300 nm, of B4, B4-*pks7* and B4-*pks7*-5306KO cultivated in SFM liquid medium. **B** HPLC chromatograms, monitored at 300 nm, of B4, B4-*pks7* and B4-*pks7*-5306KO cultivated on MS agar plate.

2. Component analysis showed that methyl esters of both **1** and **2** were present in the headspace extracts of *S. erythraea* [26]. These results designated *pks7* as a common biosynthesis BGC for **1** and **2**. The homologous enzyme which is experimentally validated for synthesizing both 6-methylsalicyclic acid and 6-ethylsalicyclic acid is a 6-methylsalicyclic acid synthetase from *Penicillium patulum* [22]. Our study provided a bacterial example. As shown in Fig. 2A, the yield of **2** was lower than that of **1**, which almost led to a failure to detect its presence. The addition of 50 mM propionate to SFM liquid medium resulted in an obvious increase for **2**. However, the yield of **2** was still lower than **1** (Fig. S8). Previous studies have discovered that the relative abundance of intracellular acetyl-CoA (381.2 ± 21.1) was dramatically higher than that of propionyl-CoA (15.6 ± 0.6) when the recombinant *S. albus* strain was cultivated in liquid fermentation media [27]. Therefore, when cultured in SFM liquid media, the mere addition of exogenous propionate could hardly offset the significant disparity between the intracellular concentrations of acetyl- CoA and propionyl-CoA.

Both **3** and **4** are esterification products of **2** with flavonoids. Flavonoids are widely spread natural products mainly from plants and exhibit potentially pharmacological characteristics, including anticancer, anti-oxidant, anti-inflammatory, and anti-infection [28,29]. Numerous kinds of flavonoids have also been extracted from actinomycetes and fungi [30]. However, the origin of flavonoids is unclear in microbes. The discovery of flavonoid derivatives including daidzein and genistein from actinomycetes and fungi is often accompanied by using malt extract or soybean flour as culture media [31]. Additionally, actinomycetes and fungi can utilize the flavonoids as substrates via whole-cell biotransformation [30]. Recently, the biosynthetic origin of flavonoids has been characterized both in bacteria and fungi [32–35]. However, no homologous genes of reported flavonoid BGCs, whether bacterial or fungal origin, could be found in *S. albus* genome. Considering that **3** and **4** are only present in the fermentation extract with MS media but not with SFM media, they are supposed to be esterification products of **2** with daidzein and genistein in MS media. The esterification may be realized through unknown biotransformation.

4. Conclusion

In summary, we have successfully characterized the function of *pks7* from *S. erythraea* for synthesizing both 6-methylsalicyclic acid and 6-ethylsalicyclic acid. Four compounds, including two new 6-ethylsalicyclic acid ester derivatives, were isolated and elucidated. Heterologous expression has been proven to be an efficient approach to mine the

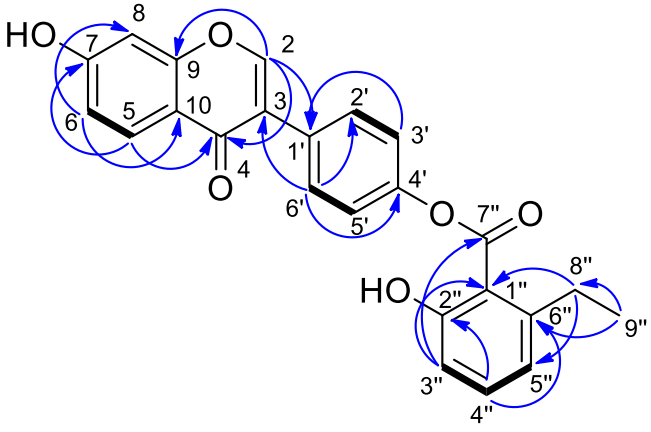

Fig. 4. Key COSY (bold bond) and HMBC (blue arrows) correlations of **3**.

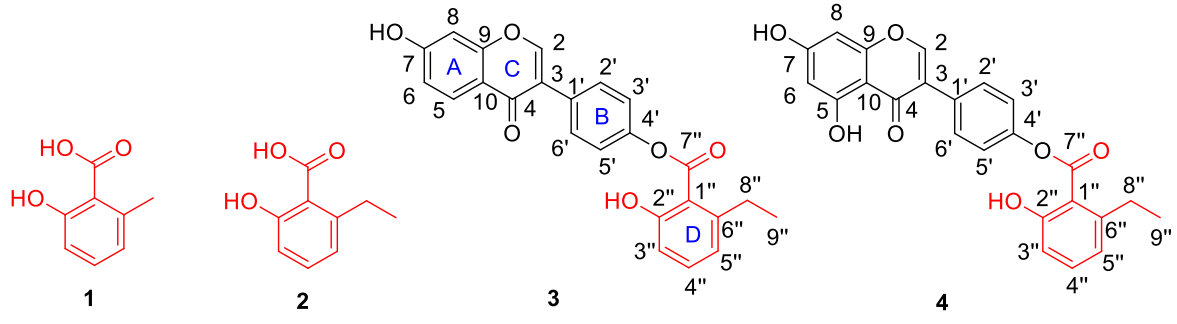

Fig. 3. Structures of compounds **1**–**4**.

hidden BGCs in *S. erythraea*. Besides *pks7*, there are still numerous BGCs in *S. erythraea* that have not been identified. Also, considerable numbers of hidden BGCs exist in industrial *Streptomyces*, which await for further genome mining, metabolite identification, and enzymatic function characterization.

Funding

This work received financial support from the National Key Research and Development Program of China (Nos. 2018YFA0900404 and 2022YFC2804105), and the National Natural Science Foundation of China (No. 82273849).

CRediT authorship contribution statement

Hao Tang: Data curation, Formal analysis, Investigation, Visualization, Writing – original draft. **Xingchi Yang:** Data curation, Formal analysis, Investigation, Visualization. **Wenzong Wang:** Investigation. **Xingjun Cui:** Investigation. **Wenping Wei:** Writing – review & editing. **Jing Wu:** Investigation. **Peng Sun:** Conceptualization, Investigation, Writing – review & editing, Supervision, Funding acquisition. **Bang-Ce Ye:** Conceptualization, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are grateful to Prof. Andriy Luzhetskyy (Universität des Saarlandes) for providing the *Streptomyces albus* B4 strain.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.synbio.2024.05.004.

References

[1] Newman DJ, Cragg GM. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. J Nat Prod 2020;83:770–803. https://doi.org/10.1021/acs.jnatprod.9b01285.
[2] Newman DJ, Cragg GM. Natural products as sources of new drugs from 1981 to 2014. J Nat Prod 2016;79:629–61. https://doi.org/10.1021/acs.jnatprod.5b01055.
[3] Rutledge PJ, Challis GL. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. Nat Rev Microbiol 2015;13:509–23. https://doi.org/10.1038/nrmicro3496.
[4] Oliynyk M, Samborskyy M, Lester JB, Mironenko T, Scott N, Dickens S, et al. Complete genome sequence of the erythromycin-producing bacterium *Saccharopolyspora erythraea* NRRL23338. Nat Biotechnol 2007;25:447–53. https://doi.org/10.1038/nbt1297.
[5] Mazodier P, Petter R, Thompson C. Intergeneric conjugation between *Escherichia coli* and *Streptomyces* species. J Bacteriol 1989;171:3583–5. https://doi.org/10.1128/jb.171.6.3583-3585.1989.
[6] Myronovskyi M, Rosenkranzer B, Nadmid S, Pujic P, Normand P, Luzhetskyy A. Generation of a cluster-free *Streptomyces albus* chassis strains for improved heterologous expression of secondary metabolite clusters. Metab Eng 2018;49:316–24. https://doi.org/10.1016/j.ymben.2018.09.004.
[7] Bierman M, Logan R, Obrien K, Seno ET, Rao RN, Schoner BE. Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. Gene 1992;116:43–9. https://doi.org/10.1016/0378-1119(92)90627-2.
[8] Rang J, He H, Yuan S, Tang J, Liu Z, Xia Z, et al. Deciphering the metabolic pathway difference between *Saccharopolyspora pogona* and *Saccharopolyspora spinosa* by comparative proteomics and metabonomics. Front Microbiol 2020;11:396. https://doi.org/10.3389/fmicb.2020.00396.
[9] Liu Q, Xiao LP, Zhou YJ, Deng KH, Tan GY, Han YC, et al. Development of *Streptomyces* sp. FR-008 as an emerging chassis. Synth Syst Biotechnol 2016;1:207–14. https://doi.org/10.1016/j.synbio.2016.07.002.
[10] Cortes J, Velasco J, Foster G, Blackaby AP, Rudd BAM, Wilkinson B. Identification and cloning of a type III polyketide synthase required for diffusible pigment biosynthesis in *Saccharopolyspora erythraea*. Mol Microbiol 2002;44:1213–24. https://doi.org/10.1046/j.1365-2958.2002.02975.x.
[11] Sun H, Ho CL, Ding F, Soehano I, Liu XW, Liang ZX. Synthesis of (*R*)-mellein by a partially reducing iterative polyketide synthase. J Am Chem Soc 2012;134:11924–7. https://doi.org/10.1021/ja304905e.
[12] Boakes S, Oliynyk M, Cortes J, Bohm I, Rudd BA, Revill WP, et al. A new modular polyketide synthase in the erythromycin producer *Saccharopolyspora erythraea*. J Mol Microbiol Biotechnol 2004;8:73–80. https://doi.org/10.1159/000084562.
[13] Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, Medema MH, et al. AntiSMASH 6.0: Improving cluster detection and comparison capabilities. Nucleic Acids Res 2021;49:W29–35. https://doi.org/10.1093/nar/gkab335.
[14] Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 1999;27:29–34. https://doi.org/10.1093/nar/27.1.29.
[15] Jia XY, Tian ZH, Shao L, Qu XD, Zhao QF, Tang J, et al. Genetic characterization of the chlorothricin gene cluster as a model for spirotetronate antibiotic biosynthesis. Chem Biol 2006;13:575–85. https://doi.org/10.1016/j.chembiol.2006.03.008.
[16] Shao L, Qu XD, Jia XY, Zhao QF, Tian ZH, Wang M, et al. Cloning and characterization of a bacterial iterative type I polyketide synthase gene encoding the 6-methylsalicyclic acid synthase. Biochem Biophys Res Commun 2006;345:133–9. https://doi.org/10.1016/j.bbrc.2006.04.069.
[17] Van Lanen SG, Oh TJ, Liu W, Wendt-Pienkowski E, Shen B. Characterization of the maduropeptin biosynthetic gene cluster from *Actinomadura madurae* ATCC 39144 supporting a unifying paradigm for enediyne biosynthesis. J Am Chem Soc 2007;129:13082–94. https://doi.org/10.1021/ja073275o.
[18] Huang H, Zheng GS, Jiang WH, Hu HF, Lu YH. One-step high-efficiency CRISPR/Cas9-mediated genome editing in *Streptomyces*. Acta Biochim Biophys Sin 2015;47:231–43. https://doi.org/10.1093/abbs/gmv007.
[19] Baur B, Storch K, Martz KE, Goettert MI, Richters A, Rauh D, et al. Metabolically stable dibenzo[*b,e*]oxepin-11(6*H*)-ones as highly selective p38 MAP kinase inhibitors: optimizing anti-cytokine activity in human whole blood. J Med Chem 2013;56:8561–78. https://doi.org/10.1021/jm401276h.
[20] Blum MS, Morel L, Fales HM. Chemistry of the mandibular gland secretion of the ant *Camponotus vagus*. Comp Biochem Physiol 1987;86B:251–2. https://doi.org/10.1016/0305-0491(87)90286-0.
[21] Dain JG, Ernst LA, Campbell IM, Bentley R. The formation of 6-ethylsalicylic acid by *Mycobacterium phlei*. Biomed Mass Spectrom 1974;1:57–61. https://doi.org/10.1002/bms.1200010112.
[22] Dimroth P, Ringelmann E, Lynen F. 6-Methylsalicylic acid synthetase from *Penicillium patulum*: some catalytic properties of the enzyme and its relation to fatty acid synthetase. Eur J Biochem 1976;68:591–6. https://doi.org/10.1111/j.1432-1033.1976.tb10847.x.
[23] Ogata T, Yahara S, Hisatsune R, Konishi R, Nohara T. Isoflavan and related compounds from *Dalbergia odorifera*. II. Chem Pharm Bull 1990;38:2750–5. https://doi.org/10.1248/cpb.38.2750.
[24] Dong CX, Wu KS, Shi SP, Tu PF. Flavanoids from clematis hexapetala. J Chin Pharm Sci 2006;15:15–20.
[25] Madan S, Singh GN, Kohli K, Ali M, Kumar Y, Singh RM, et al. Isoflavonoids from *Flemingia strobilifera* (L) R. Br. roots. Acta Pol Pharm 2009;66:297–303.
[26] Citron CA, Rabe P, Dickschat JS. The scent of bacteria: headspace analysis for the discovery of natural products. J Nat Prod 2012;75:1765–76. https://doi.org/10.1021/np300468h.
[27] Tan GY, Deng KH, Liu XH, Tao H, Chang YY, Chen J, et al. Heterologous biosynthesis of spinosad: an omics-guided large, polyketide synthase gene cluster reconstitution in *Streptomyces*. ACS Synth Biol 2017;6:995–1005. https://doi.org/10.1021/acssynbio.6b00330.
[28] Romano B, Pagano E, Montanaro V, Fortunato AL, Milic N, Borrelli F. Novel insights into the pharmacology of flavonoids. Phytother Res 2013;27:1588–96. https://doi.org/10.1002/ptr.5023.
[29] Al-Maharik N. Isolation of naturally occurring novel isoflavonoids: an update. Nat Prod Rep 2019;36:1156–95. https://doi.org/10.1039/c8np00069g.
[30] Wang JF, Liu SS, Song ZQ, Xu TC, Liu CS, Hou YG, et al. Naturally occurring flavonoids and isoflavonoids and their microbial transformation: a review. Molecules 2020;25:5112. https://doi.org/10.3390/molecules25215112.
[31] Maskey RP, Asolkar RN, Speitling M, Hoffmann V, Grun-Wollny I, Felck WF, et al. Flavones and new isoflavone derivatives from microorganisms: isolation and structure elucidation. Z Naturforsch B 2003;58b:686–91.
[32] Meinert H, Yi D, Zirpel B, Schuiten E, Geißler T, Gross E, et al. Discovery of novel bacterial chalcone isomerases by a Sequence-Structure-Function-Evolution strategy for enzymatic synthesis of (*S*)-flavanones. Angew Chem Int Ed Engl 2021;60:16874–9. https://doi.org/10.1002/anie.202107182.
[33] Zhang HJ, Li ZX, Zhou S, Li SM, Ran HM, Song ZL, et al. A fungal NRPS-PKS enzyme catalyses the formation of the flavonoid naringenin. Nat Commun 2022;13:6361. https://doi.org/10.1038/s41467-022-34150-7.
[34] Zhang W, Zhang X, Feng DD, Liang YJ, Wu ZY, Du SY, et al. Discovery of a unique flavonoid biosynthesis mechanism in fungi by genome mining. Angew Chem Int Ed Engl 2023;62:e202215529. https://doi.org/10.1002/anie.202215529.
[35] Moore BS, Hertweck C, Hopke JN, Izumikawa M, Kalaitzis JA, Nilsen G, et al. Plant-like biosynthetic pathways in bacteria: from benzoic acid to chalcone. J Nat Prod 2002;65:1956–62. https://doi.org/10.1021/np020230m.