Gross, F., Luniak, N., Perlova, O., Gaitatzis, N., Jenke-Kodama, H., Gerth, K., Gottschalk, D., Dittmann, E., Müller, R.
Bacterial type III polyketide synthases: Phylogenetic analysis and potential for the production of novel secondary metabolites by heterologous expression in pseudomonads
(2006) Archives of Microbiology, 185 (1), pp. 28-38.

Bacterial type III polyketide synthases: Phylogenetic analysis and potential for the production of novel secondary metabolites by heterologous expression in pseudomonads

Frank Gross[1,2,3,¶#], Nora Luniak[1,#], Olena Perlova[1], Nikolaos Gaitatzis[1,‡], Holger Jenke-Kodama[4], Klaus Gerth[3], Daniela Gottschalk[2,3], Elke Dittmann[4] and Rolf Müller[1,2,3,*]

[1] Institut für Pharmazeutische Biotechnologie, Universität des Saarlandes, Postfach 151150, 66041 Saarbrücken, Germany

[2] Institut für Pharmazeutische Biologie, Technische Universität Carolo-Wilhelmina, Mendelsohnstr. 1, 38106 Braunschweig, Germany

[3] Gesellschaft für Biotechnologische Forschung mbH, Mascheroder Weg 1, 38124 Braunschweig, Germany

[4] Institut für Biologie, Humboldt Universität, Chausseestr. 117, Berlin, Germany

[#] these authors contributed equally to this work

Corresponding author:

Institut für Pharmazeutische Biotechnologie, Universität des Saarlandes, Postfach 151150, 66041 Saarbrücken, Germany

Phone: +49 681 302 5474

Fax: +49 681 302 5473

Email: rom@mx.uni-saarland.de

Present Address:

[¶] Gene Bridges GmbH, Tatzberg 47-51, 01307 Dresden, Germany

[‡] Biotica Technology Ltd., Chesterford Research Park, Little Chesterford, Nr Saffron Walden, Essex, CB10 1XL, UK

ABSTRACT

Type III polyketide synthases (PKS) were regarded as typical for plant secondary metabolism before they were found in microorganisms recently. Due to microbial genome sequencing efforts, more and more type III PKS are found, most of which of unknown function. In this manuscript, we report a comprehensive analysis of the phylogeny of bacterial type III PKS and report the expression of a type III PKS from the myxobacterium *Sorangium cellulosum* in pseudomonads. There is no precedent of a secondary metabolite that might be biosynthetically correlated to a type III PKS from any myxobacterium. Additionally, an inactivation mutant of the *S. cellulosum* gene shows no physiological difference compared to the wild type strain which is why these type III PKS are assumed to be "silent" under the laboratory conditions administered. One type III PKS (SoceCHS1) was expressed in different *Pseudomonas* sp. after the heterologous expression in *Escherichia coli* failed. Cultures of recombinant *Pseudomonas* sp. harbouring SoceCHS1 turned red upon incubation and the diffusible pigment formed was identified as 2,5,7-trihydroxy-1,4-naphthoquinone, the autooxidation product of 1,3,6,8-tetrahydroxynaphthalene. The successful heterologous production of a secondary metabolite using a gene not expressed under administered laboratory conditions provides evidence for the usefulness of our approach to activate such secondary metabolite genes for the production of novel metabolites.

Keywords

myxobacteria, type III PKS, silent gene, pseudomonads, flaviolin, heterologous expression, genome sequencing, phylogeny

INTRODUCTION

Nearly 100 000 “small molecules” from natural sources have been characterized until the end of the last century. These secondary metabolites are defined by their molecular weight not exceeding 2500 Da. They are mainly produced by plants and microorganisms, each group being responsible for the production of almost 50% of the compounds (Demain 1999). Natural products are of immense interest because they often exhibit biological activities, such as antibiotics, cytostatics, insecticidals or antifungals. On the other hand, the interest in natural products as a source for new lead structures in drug discovery has decreased in the past decades (Grabley and Thiericke 1999; Bode and Müller 2005) and as a consequence the discovery rate of novel structural classes declined (Peric-Concha and Long 2003). In parallel, the severe clinical problem of multiresistant pathogens is on the rise and numerous infectious deseases are in danger of becoming uncurable. It has been concluded that to control the resistant pathogens novel antibiotics are urgently needed in the future (Dougherty et al. 2002). In the lead structure identification process, the probability to isolate already known compounds can be lowered by choosing new resources for the screening. One possibility to do so is to take advantage of ongoing and finished microbial genome sequencing projects of secondary metabolite producers, which show that there are numerous “silent” biosynthetic gene clusters, clearly indicating an untapped genetic potential to produce novel secondary metabolites (Bode and Müller 2005). With the number of sequenced genomes rising, these silent clusters could be a valuable route to new natural products. In addition, new methods have been established to access the genomes from difficult to culture or non-culturable microorganisms via metagenomic DNA libraries (Peric-Concha and Long 2003) that can be used as sources for new biosynthetic gene clusters. Although the expression of these metagenomic libraries in *E. coli* led to the discovery of novel substances (MacNeil et al. 2001; Gillespie et al. 2002), this heterologous expression host shows some disadvantages in comparison to others (Wenzel and Müller 2005). E. g., there are several factors which favor the heterologous expression in *Streptomyces* sp. (Peric-Concha and Long 2003) and recently it was shown that *Pseudomonas* sp., especially *Pseudomonas putida* KT2440, is a suitable heterologous host for secondary

metabolite production. Using the complex polyketide/nonribosomal peptide myxobacterial myxochromide S biosynthetic pathway, production yields superior to the orginal producer could be achieved in this strain (Wenzel et al. 2005). It was also shown that there is no need in pseudomonads to introduce a broad range phosphopantetheinyl transferase to secure the essential posttranslational modification required for active carrier protein domains involved in polyketide and nonribosomal peptide biosynthesis (Gross et al. 2005).

Recently a polyketide biosynthetic pathway for the production of small aromatic molecules was discovered in bacteria. The polyketide synthase (PKS) involved was classified as type III PKS. These enzymes belong to the chalcone synthase (CHS) and stilbene synthase (STS) superfamily of PKS, which were only found to this date in plants (Bangera and Thomashow 1999; Funa et al. 1999; Moore and Hopke 2001). Today, this enzyme class includes at least 15 functionally divergent β-ketosynthases of plant and bacterial origin (Austin et al. 2004). Type III PKS are relatively modest in size (40-47 kDa) and function as homodimer (Moore and Hopke 2001). CHS is ubiquitous in higher plants and is involved in the first step in flavonoid biosynthesis by catalyzing the decarboxylative addition of three malonyl-CoA extender units to a *p*-coumaroyl-CoA starter moiety, which is derived from L-phenylalanine (Winkel-Shirley 2001). At the same catalytic sites intermolecular cyclization and aromatization of the linear phenylpropanoid-tetraketide leads to the formation of chalcone (Austin and Noel 2003). RppA, the first described bacterial type III PKS found in *Streptomyces griseus* catalyzes the formation of 1,3,6,8-tetrahydroxynaphthalene (THN) from five molecules of malonyl-CoA, which can undergo spontaneous oxidation to flaviolin (Ueda et al. 1995). THN is normally transformed into a 1,8-dihydroxynaphthalene (DHN) monomer, which polymerizes to DHN-melanin. The same pigment is found in fungi, but here THN is synthesized by an ACP-utilizing, multi-domain iterative type I PKS (Tsai et al. 1998). In contrast to the relative small RppA protein (approx. 350 – 370 amino acids in size) the fungal enzymes are large (approx. 4150 amino acids).

It was shown that RppA is involved in the biosynthesis of melanin in *S. griseus*, because the substrate mycelium of an inactivation mutant contained no melanin-like pigments (Funa et al. 1999). In the meantime more homologues of RppA have been identified from

various actinomycetes, indicating the involvement of this enzyme in one of the general melanin biosynthetic pathways in this genus (Funa et al. 1999). At the same time it was reported that a type III PKS (PhlD) from *Pseudomonas fluorescens* Q2-87 synthesizes 2,4-diacetylphloroglucinol, one of the components responsible for the biocontrol activity of this strain (Bangera and Thomashow 1999). PhlD uses acetoacetyl-CoA as starter unit, adds one malonyl-CoA moiety and performs ring closure and aromatization to form monoacetylphloroglucinol. Although the number of bacterial type III PKS is increasing with the number of sequenced bacterial genomes (Saxena et al. 2003), the physiological function of these enzymes is unknown in most cases. This may even hold true if a reaction product can be obtained *in vitro*, but not isolated *in vivo* (Sankaranarayanan et al. 2004).

In ongoing genome sequencing projects of the Gram-negative myxobacteria *Myxococcus xanthus* DK1622 and *Sorangium cellulosum* So ce56 genes encoding putative type III PKS were identified (one type III PKS in *M. xanthus* and two in *S. cellulosum*). There are no plausible products corresponding to these enzymes known from both microorganisms, although they were intensely screened for secondary metabolite production, especially *S. cellulosum* So ce56 [(Gerth et al. 2003) and unpublished results]. This leads to the conclusion that these genes are “silent” under the cultivation conditions so far exhibited in the laboratory. Here, we present the heterologous expression and identification of the biosynthetic product of a *rppA*-like type III PKS from *S. cellulosum* in *Pseudomonas* sp.. This demonstrates the activation of a “silent” gene due to heterologous expression. In addition, a comprehensive phylogenetic analysis of bacterial type III PKS is provided.

MATERIALS AND METHODS

The bacterial strains and plasmids used in this study are listed in Table 1. *E. coli* and pseudomonads were routinely grown in Luria-Bertani (LB) (Miller 1972) medium at 37ºC and 30ºC, respectively. *S. cellulosum* So ce56 was grown at 30°C in liquid M-medium and plated on P-agar plates as described previously (Pradella et al. 2002). To analyze secondary metabolite production, strain So ce56 was grown in liquid P-Medium (Pradella et al. 2002). The following antibiotic concentrations were used: carbenicillin, 50 µg ml^{-1}; kanamycin, 50 µg ml^{-1}; hygromycin, 100 µg ml^{-1}.

Recombinant DNA techniques

Restriction enzymes and T4 DNA ligase were purchased from New England Biolabs and used according to the manufacturer's instructions. Taq-polymerase was obtained from Gibco BRL and cloned *Pfu*-polymerase was purchased from Stratagene. Both enzymes were employed in accordance with the supplied instruction manual. Plasmid DNA was purified from pseudomonads using standard methods (Sambrook et al. 1989), whereas for *E. coli* a simplified variant was used. Cell suspension-, cell lysis- and sodium acetate buffer were added in equal amounts (250 µl), after adding lysis- and acetate buffer, respectively, suspensions were immediately mixed and centrifuged, no phenol/chloroform extraction was performed. Plasmid DNA purification for critical down stream processes was performed using the NucleoSpin Plasmid Kit (Macherey-Nagel, Düren, Germany). DNA from *S. cellulosum* So ce56 was isolated using the Puregene® Genomic DNA Purification Kit (Gentra, Minneapolis, USA) according to manufacturer`s protocol.

Construction of expression plasmid. The DNA fragment encoding the first type III PKS s*oceCHS1* from *S. cellulosum* So ce56 was amplified using primer FG15 (5'-GTAG*GCGGCCGC*ATGGCCACACTGTGCAGGCCG-3') and FG16 (5'-CCCG*GAATTC*ACCGCCCGCCGGAGGACG-3'). The introduced restriction sites are depicted as italics. The PCR mixture (25 µl) contained 1xPCR-buffer, 2.5 mM $MgCl_2$, 2.5 % DMSO, 100 µM dNTPs, 600 nM of each primer, 100 ng of genomic DNA and 1 unit Taq-polymerase. PCR cycling included 30 cylces at 95ºC (30 s), 61ºC (30 s) and

72ºC (90 s), followed by 10 min at 72ºC to ensure an A-overhang. The amplicon was cloned using the TOPO TA cloning kit (Invitrogen, Karlsruhe, Germany) according to the manufacturer`s instructions to generate plasmid pTOPO/CHS(FG15/16). The fidelity of the amplicon was verified by nucleotide sequencing. This construct was used as template in a PCR reaction with primers FG19a (5'-GGGA*CATATG*GCCACACTGTGCAGGCCG-3') and FG20a (5'-CGC*CTCGAG*CCGCCCGCCGGAGGACGAAC-3') to introduce restriction sites for cloning into pCYB1. The PCR mixture (25 µl) contained 1x*Pfu*-buffer, 100 µM dNTPs, 500 nM of each primer, 10 ng of plasmid template and 1.25 units *Pfu*-polymerase. PCR cycling included 30 cycles at 95ºC (30 s), 61ºC (30 s) and 72ºC (2 min). The amplicon was restricted with *Nde*I and *Xho*I and ligated into accordingly prepared pCYB1 to generate plasmid pFG141, creating a C-terminal intein fusion protein. Again the fidelity was verified by nucleotide sequencing. The DNA-fragment that contains the fusion protein was excised with *Eco*RV and *Swa*I and cloned into *Eco*RV-digested and phosphatase treated pJB861 (Blatny et al. 1997) generating pFG154-1 and pFG154-2, respectively, representing both possible integration orientations.

The broad host plasmid harbouring the type III PKS intein fusion was transferred into *P. putida*, *P. syringae* and *P. stutzeri* by triparental mating as described previously (Hill et al. 1994). The successful transfer into each strain was confirmed by plasmid isolation and restriction digests.

Construction of inactivation plasmid. Plasmid pOPB20 was constructed after cloning of part of the *rppA*-like gene, obtained by PCR amplification into a pSUP102 derivative as described previously (Pradella et al. 2002). The restriction sites *Bam*HI and *Hin*dIII required for the further cloning steps were introduced into the oligonucleotide primers (marked in italics). A PCR product of 777 bp for the inactivation of *soceCHS1* obtained with the primer pair OP26 (5' – TGTCCGCGAACAG*GGATCC*TAG – 3') and OP27 (5' – CTCACC*AAGCTT*TCGAGATCC -3') was cloned first into the pCR2.1TOPO vector, digested with *Xma*I, filled-in with Klenow fragment of DNA polymerase I and religated. Next the fragment was cut out using enzymes *Bam*HI and *Hin*dIII and cloned into pSUPHyg digested with appropriate restriction enzymes to form the plasmid pOPB20.

Conjugation into *Sorangium cellulosum* So ce56

The plasmid pOPB20 was introduced into the methylation-deficient *E. coli* strain ET12567 containing pUB307 and used for mobilization into *S. cellulosum* So ce56 as described previously (Pradella et al. 2002; Kopp et al. 2004). Insert-specific homologous recombination leads to the inactivation of the target gene. Hygromycin resistant colonies were observed after 10–12 days of growth. Exconjugants were verified using PCR and Southern hybridization (data not shown).

Fermentation conditions

Fermentation in *E. coli*: Initially, purification of the fusion protein was attempted from *E. coli*. 200 ml LB containing carbenicillin were inoculated 1:100 with an overnight culture of the respective *E. coli* DH10B clone, incubated at 37ºC on a rotary shaker to an optical density at 600 nm (OD_{600}) of 0.6. Isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.4 mM and the cells were incubated additional 20 h at 16ºC before harvesting the cells by centrifugation. The cell pellet was resuspended in 5 ml column buffer (20 mM Tris/HCl, 500 mM NaCl, 1 mM EDTA and 0.1% Triton X-100; pH 8.0) and lysed by ultrasonication. The suspension was dispensed into microcentrifuge tubes and spun for 10 min at 4ºC and 14,000 rpm. The supernatant and pellet were analyzed after SDS-PAGE with Western-Blot.

Fermentation in *Pseudomonas* sp.: 500 ml LB containing kanamycin were inoculated with 5 ml overnight culture of *P. putida*/pFG154-1. After incubation at 30°C on a rotary shaker until the OD_{600} reached 0.6 the temperature was lowered to 16°C and incubated was continued for 48 h in total. The cells were harvested by centrifugation, resuspended in 12.5 ml column buffer and lysed by ultrasonication or French press. The cell debris was removed by centrifugation (Sorvall R-25, SS34 rotor, 13,000 rpm, 20 min at 4°C). An aliquot of the cell free extract was stored for analysis with SDS-PAGE and Western-Blot. As alternative media for fermentations terrific broth (TB) (12 g/l tryptone, 24g/l yeast extract, 4 ml/l glycerol, 2.31 g/l KH_2PO_4 and 12.54 g/l K_2HPO_4) and maize-molasses medium (MMM) (25 g/l molasses, 2.1 g/l corn flour, 10 g/l malt extract, 30 g/l sucrose, 10 g/l yeast extract and 2 g/l K_2HPO_4) were used. These fermentations were carried out at 16°C and 22°C after OD_{600} reached 0.6.

The culture volume for flaviolin production was 200 ml LB supplemented with kanamycin. The incubation was carried out at 30°C for approximately 60 h on a rotary shaker.

Flaviolin isolation

The cells were separated from the culture broth after fermentation by centrifugation. The supernatant was acidified with a fifth of the culture volume of 6 M hydrochloric acid, and then extracted twice with one volume ethylacetate. Ethyl acetate was removed *in vacuo* and the residue was dissolved in 1 ml methanol. The extracts were analyzed by HPLC (Dionex p680 solvent delivery system, PDA 100 Photodiode Array Detector and Macherey & Nagel Nucleodur C18 RP-column 125x2 mm with 3µm particle size; Buffer A: water, Buffer B: acetonitrile, solvent gradient from 5 % B at 0 min to 20 % B in 5 min, from 20 % B to 30 % B in 10 min and to 100 % B in 15 min employing a flow rate of 0.5 ml min^{-1}) and HPLC-MS (Bruker Daltonics HCT plus with Macherey & Nagel Nucleodur C18 RP-column 125x2 mm with 3 µm particle size). Flaviolin was identified by UV- and mass spectroscopy based on reference substance as well as NMR techniques.

Labeling with ^{14}C-acetate:

2 ml LB supplemented with kanamycin were inoculated to an OD_{600} of 0.1 with a saturated culture of *P. putida*/pFG154-1 and incubated at 30°C on a rotary shaker. Feeding of 10 µCi ^{14}C-acetate in total started 12 h after inoculation over a 7 h period. The last addition of the labeled acetate was followed by additional incubation at 30°C for one hour. The cells were separated from the culture broth by centrifugation. The cells and the supernatant were acidified and extracted with ethyl acetate; the solvent was reduced to 10 µl by nitrogen flow. Both extracts and the aqueous phase of the supernatant after extraction were analyzed with thin layer chromatography on silica plates (Merck). All samples were developed with chloroform/methanol 1:9 as mobile phase. Radioactivity was detected with a phosphoimager (Fujifilm FLA-3000).

Affinity chromatography

The RppA-homolog was heterologously expressed as a C-terminal intein fusion protein (CN-Impact, New England Biolabs) and purified according to the instructions of the

manufacturer with the following variations. The flow through the column was achieved by gravity. All steps were carried out at 4°C. The eluate was concentrated with Centriplus YM-10 and Centricons YM-10 (Millipore). The proteins were identified by Western Blot as described elsewhere(Gross et al. 2005).

Analysis of secondary metabolites formed by *S. cellulosum* So ce56

Culture extracts of *S. cellulosum* So ce56 wild-type as well as mutants were prepared after growth of 100 ml of bacterial culture in P-medium with addition of adsorber resin XAD™16 in 250 ml Erlenmeyer flask for 2 weeks. The XAD resin was harvested and eluted with methanol. The supernatant was extracted with ethyl acetate and after evaporation eluted with methanol. Extracts were analyzed using HPLC as described previously (Pradella et al. 2002). Moreover, different synthetic media were used for So ce56 cultivation (Gerth and Müller 2005), as well as different culture volumes (up to 2 l) and incubation times before the secondary metabolite production was analysed.

Phylogenetic analysis

PKS III amino acid sequences of bacteria, fungi, and plants were acquired from public databases via NCBI. BlastP searches (http://www.ncbi.nlm.nih.gov/BLAST) were conducted using the RppA sequence of *Streptomyces griseus* (Ueda et al. 1995) and the CHS2 sequence of *Medicago sativa* as query sequences. In addition, BlastP searches were performed at the NCBI microbial genome platform (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi). Amino acid sequences were aligned using ClustalX (Thompson et al. 1997) and adjusted manually using the MacClade software version 4.03 (Maddison and Maddison 2000). Phylogenetic relations were inferred by Bayesian estimation, maximum parsimony (MP) and the neighbor-joining (NJ) method. For Bayesian estimation the MrBayes program version 3 (Huelsenbeck and Ronquist 2001) was used employing the JTT amino acid replacement model (Jones et al. 1992) and a gamma distribution approximated by four categories to represent among site rate heterogeneity. The Markov chain Monte Carlo analysis (MCMC) ran 1.5 million generations with four independent chains. The consensus tree and posterior clade probabilities were calculated from the trees of the convergence state

of the MCMC. MP analysis was performed using the heuristic search option of PAUP*4.0b10 (Swofford 2002) with gaps being treated as missing data and branch swapping by tree-bisection-reconnection (TBR). Constant and uninformative characters were excluded from the analysis. The NJ method was conducted using the modules Seqboot, Protdist, Neighbor, and Consens of the PHYLIP software version 3.62 (Felsenstein 2002). Bootstrap analysis was done with 1000 pseudo-replicate sequences.

Nucleotide sequence accesion numbers

The nucleotide sequences of *soceCHS1* and *soceCHS2* have been deposited in GenBank under the accession numbers DQ156078 and DQ156079, respectively. The nucleotide sequence of *mxCHS* is annotated as MXAN6639 at TIGR-CMR (comprehensive microbial resources).

Results

The genomes of *Sorangium cellulosum* So ce56 and *Myxococcus xanthus* DK1622 harbour type III PKS genes

S. cellulosum So ce56 and *M. xanthus* DK1622 are the two myxobacterial strains whose genome is currently being sequenced. During our screening of the preliminary sequences of the ongoing *S. cellulosum* genome project (http://www.genetik.uni-bielefeld.de/GenoMik/cluster6.html) several interesting and unusual genes were detected. The available partial sequence data was also searched for the presence of type III PKS genes. Two corresponding ORFs were identified in *S. cellulosum* So ce56 and one in *M. xanthus* DK1622. The first ORF in *S. cellulosum* So ce56, named *soceCHS1*, is 1092 nucleotides long and the deduced polypeptide (363 aa) has a molecular weight of 39,568.5 Da. The protein exhibits a significant homology to RppA (identities: 244/349 (69%), similarities: 294/349 (83%), no gaps). The second ORF, named *soceCHS2*, is 1116 nucleotides long, the deduced protein (371 aa) has a molecular weight of 39,149.3 Da, 38% identical amino acids (143/371) and 52% similar amino acids (195/371) with 25 gaps over 371 residues (6%) in comparison to the putative chalcone synthase 2 of *Deinococcus radiodurans* R1. The type III PKS gene from *M. xanthus* DK1622, mxCHS, is 1083 nucleotides long, the corresponding protein (360 aa) has a molecular weight of 38,984.8 Da, with 40.5% identity (177/437) and 48.5% similarity (212/437) with 87 gaps over 437 amino acids (19.5%) to a putative chalcone synthase from *Magnetospirillum magnetotacticum*. All three type III PKS from myxobacteria contain the conserved type III PKS catalytical triad, consisting of cysteine, histidine and asparagines (Austin and Noel 2003). For SoceCHS1 these correspond to Cys138, His270 and Asn303, for SoceCHS2 Cys167, His296 and Asn329 and for MxCHS Cys148, His284 and Asn317, respectively. SoceCHS1 has also in common with other THNS in the database an extension of the C-terminus, varying in length up to 25 amino acid residues in comparison to all other type III PKS characterized to date (Izumikawa et al. 2003).

Figure 1

Phylogenetic analysis of bacterial type III PKS

In order to assess the evolutionary relations between the myxobacterial type III PKS proteins and other members of the chalcone synthase family a comprehensive phylogenetic analysis was performed including all bacterial and fungal sequences available in the public databases and a selection of sequences from plants. Figure 1 shows the phylogenetic tree inferred by Bayesian estimation. Using FabH sequences of archaebacteria, eubacteria and plant plastids as the outgroup, type III PKS sequences are distributed to three bacterial, one fungal and one plant clade. The first bacterial clade comprises two subclades, one that is built from RppA-like sequences of streptomycetes and a second one that mainly contains type III PKS from bacilli and mycobacteria. The bacterial clade II consists exclusively of actinobacterial sequences that include DpgA proteins involved in vancomycin antibiotic biosynthesis (Pfeifer et al. 2001). The third clade of bacterial type III PKS is the most heterogeneous group. Amongst others, it comprises sequences from actinobacteria, sphingobacteria, cyanobacteria as well as γ- and δ-proteobacteria. None of these proteins has been characterized biochemically up to now. The eukaryotic PKS III are located at the top of the tree forming two well-defined neighbor branches. The first branch contains all fungal sequences, whereas the second one is built up from the plant type III PKS, which form a very homogeneous group, compared to the other clades. This group includes representative sequences of different chalcone synthase-like proteins. The NJ and MP methods of tree reconstruction gave the same tree topology, except that in the parsimony tree the positions of bacterial groups I and II were interchanged (data not shown). The phylogenetic tree of type III PKS proteins is in agreement with earlier studies that reported a clear segregation of sequences from bacteria and plants (Moore and Hopke 2001) and a separate group of fungal sequences (Seshime et al. 2005), respectively. SoceCHS1 is located in the RppA branch of the first bacterial clade being closely related to RppA of *Streptomyces antibioticus*. MxCHS falls into the other branch of this clade together with sequences from *Magnetospirillum*

magnetotacticum and diverse actinobacteria, whereas SoceCHS2 is a member of the bacterial clade III.

Effect of the inactivation of the SoceCHS1 gene in *S. cellulosum* So ce56

In order to investigate the physiological role of the gene product of SoceCHS1 in *S. cellulosum* So ce56, the gene was inactivated by insertion of the plasmids pOPB20 into the coding region of the *S. cellulosum* chromosome and the verified mutants were compared to the wild-type. No change in secondary metabolite pattern and fatty acid composition of cell wall lipids produced under the laboratory conditions described in materials and methods was observed (data not shown).

Cloning of the *S. cellulosum* So ce56 CHS1 type III PKS.

The ORF identified in the genome of *S. cellulosum* So ce56 was used to deduct the corresponding oligonucleotide primers, which introduced artificial restriction sites in the PCR reaction. The putative type III PKS was cloned into pCYB1 (NEB) as described above to generate a C-terminal intein-chitin binding domain (CBD) fusion. After the expression of the gene in *E. coli* failed, the DNA fragment representing the fused protein together with a functional promotor was subcloned into the broad host range vector pJB861 to achieve expression in *Pseudomonas* sp. as described earlier (Gross et al. 2005).

Expression of SoceCHS1 in *E. coli* and in *Pseudomonas* sp.

As first case study, the protein SoceCHS1 should be purified and tested in enzyme assays to shed light on its possible role in the metabolism of *S. cellulosum* So ce56. To facilitate purification after expression, the protein was affinity tagged with the intein-CBD. The *E. coli* cell free extract was analyzed by Western Blot (Gross et al. 2005), but no signal could be observed, not even for the intein-CBD protein, which would indicate *in vivo* cleavage of the protein (data not shown). As this attempt to express SoceCHS1 in *E. coli* failed, we have chosen pseudomonads strains as alternative heterologous hosts for the expression of the protein. After subcloning into a broad host range plasmid, the fusion protein was expressed in *P. putida* KT2440 (Bagdasarian et al. 1981) and purified as described in Material and Methods. The cell free extract and the purified protein was

analyzed using Western Blot (Figure 2) showing that the expression of the fusion protein was successful, but that the main portion of it is cleaved *in vivo,* as indicated by the presence of a major signal at approx. 55 kDa, representing the intein-chitin-binding domain. The variation of culture media (see Figure 2) and fermentation temperature (data not shown) has little influence on the *in vivo* cleavage. Despite the fact that there was a certain amount of fused protein left in the cell free extract (Figure 2), we could not purify enough enzyme by means of affinity chromatography to conduct *in vitro* assays.

Figure 2

On the other hand the *in vivo* cleavage indicates that the native enzyme is present in the cytosol of the cell. In fact, cultures of *P. putida*/pFG154 expressing SoceCHS1 turned red when the incubation time was prolonged to approximately 60 h and carried out at 30°C. This indicated the SoceCHS1 dependent production of a red pigment in the heterologous host *P. putida*/pFG154-1.

Identification of the product synthesized by *P. putida*/pFG154

The fermentation broth of cultures from *P. putida*/pFG154 were extracted as described in Material and Methods. The extraction of the acidifed culture broth with ethyl acetate was superior to the method using XAD for flaviolin (data not shown). The HPLC-chromatogram showed an additional peak compared to a negative control (Figure 3). The UV-spectrum of this peak is consistent with authentic flaviolin. The nature of the produced compound could be verified by HPLC-MS/MS using authentic flaviolin as a reference compound. The production of flaviolin by *P. putida*/pFG154 could also be confirmed by feeding ^{14}C-acetate to the cultures. The radioactive spot co-migrated with flaviolin (data not shown). In addition, flaviolin was isolated from 1 l culture broth as described above and further purified using silica gel chromatography resulting in ~6 mg of pure compound which was characterized by mass spectrometry and ^{1}H-NMR (LC/MS: ESI (+) m/z 207 $[M+H]^+$; MS/MS (precursor ion at m/z 207) m/z 179; ^{1}H-NMR (500 MHz, d_6-DMSO) δ [ppm]: 5,51 (1H, br s, H-3), 6,42 (1H, d, *J*= 2,14 Hz, H-7), 6,8 (1H, d, *J*= 2,14 Hz, H-9)).

Figure 3

The plasmid harboring the gene for the *soceCHS1*-intein fusion was also successfully introduced by conjugation into *P. syringae* pv. *tomato* DC3000 (Cuppels 1986) and *P. stutzeri* DSM10701 (Carlson et al. 1983). The production of flaviolin could be verified by HPLC analysis for recombinant strains of *P. stutzeri*/pFG154 (data not shown). The yield was not as good as for *P. putida*/pFG154 using the same fermentation conditions. The production in *P. syringae*/pFG154 could only be verified by HPLC-MS-analysis, because the yield was very low even after prolonged fermentation times (data not shown).

DISCUSSION

It is still controversely discussed in the literature if PKS III enzymes that are ubiquitous in higher plants and only rarely present in modern eubacteria have originally evolved in plants or if ancestors of these enzymes already existed in ancient eubacteria and were mostly lost during prokaryotic evolution. PKS enzymes generally share high structural and mechanistic similarities with fatty acid biosynthesis enzymes (Austin and Noel 2003; Jenke-Kodama et al. 2005). Type III PKS enzymes are most similar to the FabH type of β-ketoacyl synthase involved in fatty acid biosynthesis in eubacteria and plants. The analysis of the phylogenetic distribution of FabH and type III PKS sequences from different types of organisms suggests that first bacterial type III PKS evolved from the FabH proteins in eubacteria and that these proteins are the ancestors of the diversity of type III PKS enzymes that is common in higher plants. The phylogenetic tree could well reflect a continuous evolution of type III PKS. Alternatively, bacterial and fungal type III PKS must be explained as a result of a series of horizontal gene transfer events (HGTs) from plants. With the increasing number of type III PKS that are becoming known due to bacterial genome sequencing the latter hypothesis seems unlikely. First, type III PKS enzymes in bacteria are more similar to the FabH ancestors of both bacteria and plants, whereas plant FabH and type III PKS sequences are very distant from each other in the phylogenetic trees (see figure 1). Second, bacterial type III PKS sequences show a much higher diversity than plant type III PKS and only one subgroup of bacterial sequences clusters in the neighborhood of the plant sequences. Interestingly, this bacterial subgroup also contains so far uncharacterized sequences encoded in cyanobacterial genomes. Thus, there is the possibility that plant type III PKS sequences originate from the cyanobacterial endosymbiont that was the ancestor of modern chloroplasts. A third fact that contradicts the theory of a horizontal gene transfer from plants is that type III PKS sequences are found in single strains of all major subgroups of bacteria and that only a minority of these strains are plant associated. However, the assessment of the evolutionary history of type III PKS must still be regarded as preliminary, as the majority of bacterial type III PKS is still uncharacterized and in particular those sequences that fall in the subgroup that is most similar to the plant sequences. The myxobacterial sequences investigated in this study cluster in two of the three different bacterial subgroups of type III PKS. The

SoceCHS1 sequence falls into the subbranch of actinobacterial RppA like sequences that also includes the PhlD sequence from the γ-Proteobacterium *P. fluorescens*. This position in the phylogenetic tree presumably indicates an HGT event from actinobacteria for both proteobacterial sequences. HGT events are also discussed for type I PKS in proteobacteria (Jenke-Kodama et al., 2005). It is difficult to evaluate whether the other two type III PKS sequences from myxobacteria are also results of HGT events or the result of a speciation process as they fall into subbranches of the phylogenetic tree that contain sequences representing all major subclasses of bacteria.

In this study, we show that the biosynthetic potential behind these type III PKS of unknown function ("silent" biosynthetic genes) can be used by heterologous expression for the production of novel secondary metabolites.

The genome sequencing projects of the myxobacteria *M. xanthus* DK1622 and *S. cellulosum* So ce56 revealed the presence of one type III PKS in *M. xanthus* and two in *S. cellulosum*. The *in silico* analysis of the three sequences showed no peculiarity with respect to the molecular mass of the deduced protein, peptide length or important conserved amino acid residues compared to other type III PKS in the database. The extensive screening of *S. cellulosum* under different culture conditions (Gerth et al. 2003; Gerth and Müller 2005) did not identify flaviolin or related compounds with a naphthoquinone ring as a product of the metabolism. In fact, no product correlated to a type III PKS is known to date from any myxobacterium. In addition, the knock-out of *soceCHS1* in *S. cellulosum* had no effect on the morphology, pigmentation and secondary metabolite spectrum of the mutant compared to the wild type. These findings lead to the assumption that the RppA-homologue SoceCHS1 is a gene not expressed under laboratory conditions. This is probably also the fact for MxCHS and SoceCHS2. We performed knock-out mutageneses for both genes in the respective strains. Both mutants were undistinguishable from their respective wild type strain (O. Perlova, G. Gerth and R. Müller, unplished). Attempts to express both genes heterologously are currently under way. The latter two enzymes cluster in the phylogenetic tree (see figure 1) with PKS10/PKS11 and PKS18 from *Mycobacterium tuberculosis*, respectively. It was shown *in vitro*, that PKS11 and PKS18 can utilize long-chain aliphatic acyl coenzyme-A substrates as starter units and extend these with two or three acetate units derived from

malonyl-CoA to generate α-pyrones (Saxena et al. 2003). Neither in *M. tuberculosis* (Sankaranarayanan et al. 2004) nor in *M. xanthus* and *S. cellulosum* such long chain α-pyrones could be detected. For *M. tuberculosis* it is known that the organism harbors a striking amount of long-chain and very long-chain fatty acids (Russell et al. 2002), which are part of cell wall lipids. It seems possible that PKS11 and PKS18 are involved in the biosynthesis of these fatty acids. In analogy, SoceCHS2 and MxCHS could also play a role in the biosynthesis of such long chain fatty acids, but corresponding knock-out mutants show no differences to the respective wild types in the fatty acids composition of cell wall lipids (data not shown; compare (Bode et al. 2005; Dickschat et al. 2005)). Further studies to elucidate possible roles of *soceCHS2* and *mxCHS* in the metabolism of *S. cellulosum* So ce56 and *M. xanthus* are under way in our laboratory.

The attempts to express SoceCHS1 as C-terminal fusion protein in *E. coli* failed. One reason for this could be the high GC-content of myxobacterial DNA (67.5 % for this ORF) so that *E. coli* is not equipped suitably with the required tRNAs. For example, the codons TTT and TTC for Phe are used nearly to the same extend (57.42 % and 42.58 %, respectively). In comparison, *M. xanthus* tends heavily towards the usage of TTC with 94.7 % and only 5.3 % for TTT. Another example is the codon GAC for Asp, which has a triplet frequency of 32.22 % in *E. coli* and 83.9 % in *M. xanthus* (data from TIGR-CMR). Since it was recently shown that *Pseudomonas* sp. are suitable hosts for the expression of myxobacterial genes (Gross et al. 2005; Wenzel et al. 2005), the fusion gene was subcloned into a broad host range vector and conjugated into *P. putida*, *P. syringae* and *P. stutzeri*. The expression of SoceCHS1 could be achieved in *P. putida* (Fig. 2), *P. syringae* and *P. stutzeri*. However, *in vivo* cleavage appears as a major problem using the intein-chitin binding domain (CBD) fusion as C-terminal tag. This phenomenon has been described earlier (Gross et al. 2005). It is evident from the strong signal corresponding to the molecular weight of the intein-CBD that the expression was successful, but the majority of the fused protein is already cleaved in the cell, which makes successful purification employing this system impractical.

Figure 4

In case of SoceCHS1 the failure of the fusion system proved lucky, because the produced and fusion freed enzyme was active *in vivo* carrying out the reaction sequence depicted in figure 4. The red pigment was produced in good yield (~6 mg/l) without optimizing production conditions by the recombinant *P. putida* strain. The chemical nature of the molecule was verified by NMR and HPLC-MS as 2,5,7-trihydroxy-1,4-naphthoquinone or flaviolin, the product of spontaneous oxidation of 1,3,6,8-tetrahydroxynaphthalene (THN). This finding was further strengthened with the feeding of ^{14}C-labeled acetate which was incorporated into flaviolin after activation within the cell to malonyl-CoA. The heterologous expression of SoceCHS1 was also successful in *P. syringae* and *P. stutzeri*, although the yield was not as good as in *P. putida*. Interestingly, one of the first examples of the heterologous expression of a so called "silent" gene also involved a RppA-homologue from the actinomycete *Streptomyces coelicolor* A3(2) (Izumikawa et al. 2003). In the genome of this bacterium three putative type III PKS were identified (Bentley et al. 2002) and one of these encodes a RppA-like protein, which was heterologously expressed in active form in *E. coli* although the codon usage of the gene is similar to SoceCHS1. Since a knock-out of this type III PKS in an *act*$^-$, *red*$^-$, *whiE*$^-$ triple-mutant (which is only slightly pigmented in comparison to the wild type strain) showed no phenotypic difference, it was concluded that this gene is silent (Izumikawa et al. 2003) and not involved in the biosynthesis of a diffusible pigment. This is in contradiction to the conclusion drawn from the inactivation of RppA in *Sacharopolyspora erythraea* (Cortes et al. 2002); these authors state that the widely spread RppA genes in streptomycetes are implicated in the biosynthesis of diffusible pigments. The organization of the operon in which RppA is located in *S. erythraea*, *S. avermitilis*, *S. griseus* and *S. coelicolor* is very similar. RppA is always found adjacent to a cytochrome P_{450}-dependent monooxidase, which is believed to carry out a tailoring enzymatic step giving rise to the final product. The sequences of these P_{450} enzymes exhibit weak, but significant homologies to oxidases in the biosynthesis gene clusters of chloroeremomycin and balhimycin (Cortes et al. 2002) which are believed to establish biphenyl and biphenyl ether cross-links in the respective molecules (Pelzer et al. 1999; Bischoff et al. 2001). In the neighborhood of *soceCHS1* a P_{450} monooxidase could not be detected.

The production of flaviolin using a "silent" gene from *S. cellulosum* described in this manuscript shows again that pseudomonads offer good qualities as heterologous hosts for the expression of myxobacterial genes and the production of the respective compounds in good yields (Wenzel and Müller 2005). The similar GC-content compared to several important secondary metabolite producers (e.g. streptomycetes, myxobacteria) increases the chances that the expression is successful. A further advantage of *Pseudomonas* sp. as heterologous host, especially for biosynthetic gene clusters of polyketides, nonribosomal pepetides and hybrids thereof, is that pseudomonads themselves are producers of secondary metabolites and therefore harbour some of the required biosynthetic machinery, as opposed to *E. coli*. Nevertheless, further engineering of pseudomonads hosts to accomplish more sophisticated biosyntheses is needed and under way in our laboratory.

ACKNOWLEDGMENTS

Authentic flaviolin reference substance was a kind gift from Prof. Dr. H. Anke, Kaiserlautern. The authors highly appreciate the help of D. Krug and P. Meiser in the chemical analysis of flaviolin. This work was supported by the German Ministry for Education and Research (BMB+F, BioFuture program) and the Deutsche Forschungsgemeinschaft (DFG) within the Schwerpunktprogramm “evolution of metabolic diversity”.

Legends to the figures:

Figure 1

Phylogenetic tree of type III PKS protein sequences from bacteria, fungi, and plants as inferred by Bayesian estimation. Numbers above branches indicate posterior clade probability values. The tips of the tree give the species names, protein names, if annotated, and the accession numbers. Plant sequences are assigned to the different type III PKS families. CHS: chalcone synthase; STS: stilbene synthase; BPS: benzophenone synthase; BAS: benzalacetone synthase; ACS: acridone synthase; RS: resveratrol synthase. Bacterial and plant FabH proteins were chosen as outgroup.

Figure 2

Western Blot of SoceCHS1 expressed in *P. putida* in different media at 16°C. 1-3: cell free extracts; 4-6: insoluble fractions; 1+4: LB; 2+5: TB; 3+5: MM medium, repectively; A: SoceCHS1 intein fusion; B: intein chitin binding domain

Figure 3

HPLC-chromatogram of culture extracts from *P. putida* wild type and *P. putida*/pFG154-1 expressing SoceCHS1; the peak at approx. 8 min. corresponds to flaviolin.
Inlet: Overlay of the UV-spectrum of the main peak in the HPLC chromatogram with authentic flaviolin

Figure 4

Reaction scheme of 1,3,6,8-tetrahydroxynaphthalene synthase (THNS)

REFERENCES

Austin MB et al. (2004) Crystal structure of a bacterial type III polyketide synthase and enzymatic control of reactive polyketide intermediates. J Biol Chem 279:45162-45174

Austin MB, Noel JP (2003) The chalcone synthase superfamily of type III polyketide synthases. Nat Prod Rep 20:79-110

Bagdasarian M et al. (1981) Specific-purpose plasmid cloning vectors. II. Broad host range, high copy number, RSF1010-derived vectors, and a host-vector system for gene cloning in Pseudomonas. Gene 16:237-247

Bangera MG, Thomashow LS (1999) Identification and Characterization of a Gene Cluster for Synthesis of the Polyketide Antibiotic 2,4-Diacetylphloroglucinol from *Pseudomonas fluorescens* Q2-87. J. Bacteriol. 181:3155-3163

Bentley SD et al. (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). Nature 417:141-147

Bischoff D et al. (2001) The Biosynthesis of Vancomycin-Type Glycopeptide Antibiotics-The Order of the Cyclization Steps. Angew Chem Int Ed Engl 40:4688-4691

Blatny JM, Brautaset T, Winther-Larsen HC, Karunakaran P, Valla S (1997) Improved Broad-Host-Range RK2 Vectors Useful for High and Low Regulated Gene Expression Levels in Gram-Negative Bacteria. Plasmid 38:35-51

Bode HB, Dickschat JS, Kroppenstedt RM, Schulz S, Müller R (2005) Biosynthesis of iso-fatty acids in myxobacteria: iso-even fatty acids are derived by alpha-oxidation from iso-odd fatty acids. J Am Chem Soc 127:532-533

Bode HB, Müller R (2005) The impact of bacterial genomics on natural product research. Angew Chem Int Ed:in press

Carlson CA, Pierson LS, Rosen JJ, Ingraham JL (1983) *Pseudomonas stutzeri* and related species undergo natural transformation. J Bacteriol 153:93-99.

Cortes J, Velasco J, Foster G, Blackaby AP, Rudd BAM, Wilkinson B (2002) Identification and cloning of a type III polyketide synthase required for diffusible pigment biosynthesis in *Saccharopolyspora erythraea*. Molecular Microbiology 44:1213-1224

Cuppels DA (1986) Generation and characterization of Tn5 insertion mutations in *Pseudomonas syringae* pv. *tomato*. Appl Environ Microbiol 51:323-327

Demain AL (1999) Pharmaceutically active secondary metabolites of microorganisms. Appl Microbiol Biotechnol 52:455-463

Dickschat JS, Bode HB, Kroppenstedt RM, Müller R, Schulz S (2005) Biosynthesis of iso-fatty acids in myxobacteria. Org Biomol Chem 3:2824-2831

Dougherty TJ, Barrett JF, Pucci MJ (2002) Microbial genomics and novel antibiotic discovery: new technology to search for new drugs. Curr Pharm Des 8:1119-1135

Felsenstein J (2002) PHYLIP (phylogeny inference package). Version 3.6. In. Distributed by the author

Funa N, Ohnishi Y, Fujii I, Shibuya M, Ebizuka Y, Horinouchi S (1999) A new pathway for polyketide synthesis in microorganisms. Nature 400:897-899

Gerth K, Müller R (2005) Studies on the regulation of secondary metabolism of *Sorangium cellulosum* So ce56. Development of simple media which allow

investigations into the synthesis of secondary metabolites. J. Biotechnology submitted

Gerth K, Pradella S, Perlova O, Beyer S, Müller R (2003) Myxobacteria: Proficient producers of novel natural products with various biological activities - past and future biotechnological aspects with the focus on the genus *Sorangium*. J Biotechnol 106:233-253

Gillespie DE et al. (2002) Isolation of Antibiotics Turbomycin A and B from a Metagenomic Library of Soil Microbial DNA. Appl. Environ. Microbiol. 68:4301-4306

Grabley S, Thiericke R (1999) The Impact of natural Products on Drug Discovery. In: Grabley S, Thiericke R (eds) Drug Discovery from Nature. Springer - Verlag Berlin Heidelberg New York

Gross F, Gottschalk D, Müller R (2005) Posttranslational modification of myxobacterial carrier protein domains in *Pseudomonas* sp. by an intrinsic phosphopantetheinyl transferase. Appl Microbiol Biotechnol 68:66-74

Hill D et al. (1994) Cloning of Genes Involved in the Synthesis of Pyrrolnitrin from *Pseudomonas fluorescens* and Role of Pyrrolnitrin Synthesis in Biological Control of Plant Disease. Appl. Environ. Microbiol. 60:78-85

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17:754-755

Izumikawa M et al. (2003) Expression and characterization of the type III polyketide synthase 1,3,6,8-tetrahydroxynaphthalene synthase from *Streptomyces coelicolor* A3(2). J Ind Microbiol Biotechnol 30:510-515

Jenke-Kodama H, Sandmann A, Müller R, Dittmann E (2005) Evolutionary implications of bacterial polyketide synthases. Mol Biol Evol 22:2037-2039

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8:275-282

Kopp M et al. (2004) Critical variations of conjugational DNA transfer into secondary metabolite multiproducing *Sorangium cellulosum* strains So ce12 and So ce56: development of a mariner-based transposon mutagenesis system. Journal of Biotechnology 107:29-40

MacNeil IA et al. (2001) Expression and isolation of antimicrobial small molecules from soil DNA libraries. J Mol Microbiol Biotechnol 3:301-308

Maddison WR, Maddison WP (2000) MacClade version 4.0. In: Associates S (ed), Sunderland, MA

Miller J (1972) Experiments in molecular genetics. Cold Spring Harbor Laboratories, Cold Spring Harbor, NY

Moore BS, Hopke JN (2001) Discovery of a new bacterial polyketide biosynthetic pathway. Chembiochem 2:35-38

Pelzer S et al. (1999) Identification and analysis of the balhimycin biosynthetic gene cluster and its use for manipulating glycopeptide biosynthesis in *Amycolatopsis mediterranei* DSM5908. Antimicrobial Agents and Chemotherapy 43:1565-1573

Peric-Concha N, Long PF (2003) Mining the microbial metabolome: a new frontier for natural product lead discovery. Drug Discov Today 8:1078-1084

Pfeifer V et al. (2001) A polyketide synthase in glycopeptide biosynthesis - The biosynthesis of the non-proteinogenic amino acid (S)-3,5-dihydroxyphenylglycine. J Biol Chem 276:38370-38377

Pradella S, Hans A, Sproer C, Reichenbach H, Gerth K, Beyer S (2002) Characterisation, genome size and genetic manipulation of the myxobacterium *Sorangium cellulosum* So ce56. Archives of Microbiology 178:484-492

Russell DG, Mwandumba HC, Rhoades EE (2002) Mycobacterium and the coat of many lipids. J. Cell Biol. 158:421-426

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: A laboratory manual, 2 edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA

Sankaranarayanan R, Saxena P, Marathe UB, Gokhale RS, Shanmugam VM, Rukmini R (2004) A novel tunnel in mycobacterial type III polyketide synthase reveals the structural basis for generating diverse metabolites. Nat Struct Mol Biol 11:894-900

Saxena P, Yadav G, Mohanty D, Gokhale RS (2003) A new family of type III polyketide synthases in *Mycobacterium tuberculosis*. Journal of Biological Chemistry 278:44780-44790

Seshime Y, Juvvadi PR, Fujii I, Kitamoto K (2005) Discovery of a novel superfamily of type III polyketide synthases in *Aspergillus oryzae*. Biochem Biophys Res Commun 331:253-260

Swofford D (2002) PAUP* 4.0: Phylogenetic analysis using parsimony. In. Sinauer Associates, Sunderland, MA

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876-4882

Tsai H-F, Chang YC, Washburn RG, Wheeler MH, Kwon-Chung KJ (1998) The Developmentally Regulated alb1 Gene of Aspergillus fumigatus: Its Role in Modulation of Conidial Morphology and Virulence. J. Bacteriol. 180:3031-3038

Ueda K, Kim KM, Beppu T, Horinouchi S (1995) Overexpression of a gene cluster encoding a chalcone synthase-like protein confers redbrown pigment production in Streptomyces griseus. J Antibiot (Tokyo) 48:638-646.

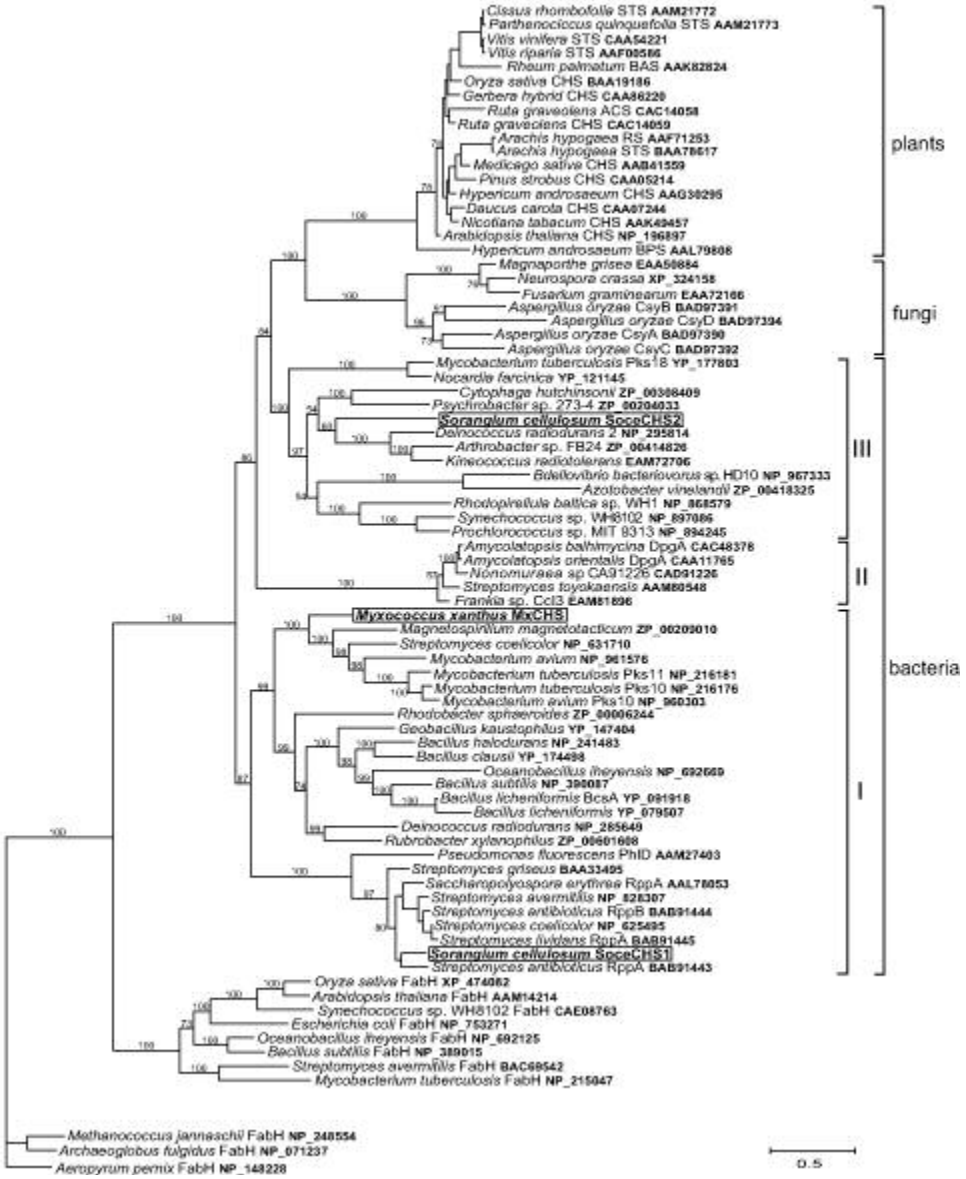
Wenzel S, Müller R (2005) Recent developments towards the heterologous expression of complex bacterial natural product biosynthetic pathways. Curr Op Biotechnol:in press

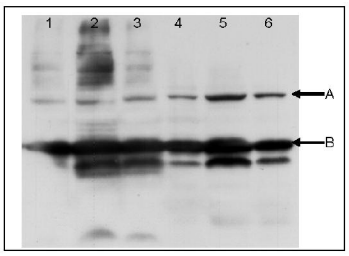
Wenzel SC, Gross F, Zhang Y, Fu J, Stewart AF, Müller R (2005) Heterologous expression of a myxobacterial natural products assembly line in pseudomonads via red/ET recombineering. Chem Biol 12:349-356

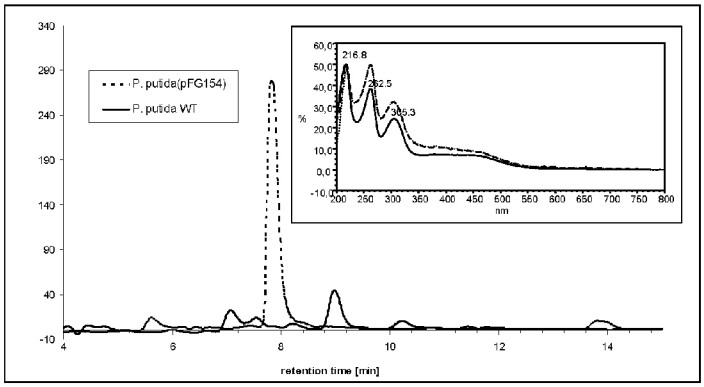
Winkel-Shirley B (2001) Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiology 126:485-493

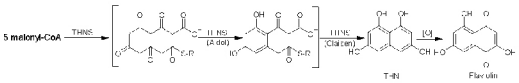
Table 1: Strains and Plasmids used in this study

	Relevant characteristics	Reference
Strains		
E. coli DH10B		Stratagene
E. coli TOP10		Invitrogen
Pseudomonas putida KT2440	wildype	(Bagdasarian et al. 1981)
Pseudomonas syringae pv. *tomato* DC3000	wildtype	(Cuppels 1986)
Pseudomonas stutzeri DSM10701	wildtype	(Carlson et al. 1983)
Sorangium cellulosum So ce56	wildtype,	GBF strain collection
Sorangium cellulosum B20-7 and B20-8	*Sorangium cellulosum* So ce56 *rppA*$^-$	this study
Plasmids		
pCR2.1 TOPO	TA cloning vector for PCR fragments, amp^R, kan^R	Invitrogen
pCYB1	amp^R, pMB1ori, VMA Intein, CBD	New England Biolabs
pJB861	kan^R, RK2*ori*, *ori*T, P_m-promoter	(Blatny et al. 1997)
pSUPHyg	pSUP102 derivative	(Pradella et al. 2002)
pTOPO/CHS(FG15/16)	*rppA* cloned in pCR2.1 TOPO	this study
pFG141	*rppA* cloned as *NdeI-XhoI* fragment in pCYB1	this study
pFG154-1 and -2	*rppA* subcloned as EcoRV-SwaI fragment in pJB861	this study
pOPB20	inactivation fragment for *SoceCHS1* cloned into pSUPhyg	this study


Cissus rhombifolia STS AAM21772
Parthenocissus quinquefolia STS AAM21773
Vitis vinifera STS CAA54221
Vitis riparia STS AAF08586
Rheum palmatum BAS AAK82824
Oryza sativa CHS BAA19186
Gerbera hybrid CHS CAA86220
Ruta graveolens ACS CAC14058
Ruta graveolens CHS CAC14059
Arachis hypogaea RS AAF71253
Arachis hypogaea STS BAA78617
Medicago sativa CHS AAB41559
Pinus strobus CHS CAA05214
Hypericum androsaemum CHS AAG30295
Daucus carota CHS CAA07244
Nicotiana tabacum CHS AAK49457
Arabidopsis thaliana CHS NP_196897
Hypericum androsaemum BPS AAL79808
Magnaporthe grisea EAA56884
Neurospora crassa XP_324158
Fusarium graminearum EAA72186
Aspergillus oryzae CsyB BAD97391
Aspergillus oryzae CsyD BAD97394
Aspergillus oryzae CsyA BAD97390
Aspergillus oryzae CsyC BAD97392
Mycobacterium tuberculosis Pks18 YP_177803
Nocardia farcinica YP_121145
Cytophaga hutchinsonii ZP_00388409
Psychrobacter sp. 273-4 ZP_00204033
Sorangium cellulosum SoceCHS2
Deinococcus radiodurans 2 NP_295814
Arthrobacter sp. FB24 ZP_00414826
Kineococcus radiotolerans EAM72706
Bdellovibrio bacteriovorus sp. HD10 NP_967333
Azotobacter vinelandii ZP_00418325
Rhodopirellula baltica sp. WH1 NP_868579
Synechococcus sp. WH8102 NP_897086
Prochlorococcus sp. MIT 9313 NP_894245
Amycolatopsis balhimycina DpgA CAC48378
Amycolatopsis orientalis DpgA CAA11765
Nonomuraea sp CA91226 CAD91226
Streptomyces toyokaensis AAM80548
Frankia sp. Ccl3 EAM81696
Myxococcus xanthus MxCHS
Magnetospirillum magnetotacticum ZP_00209010
Streptomyces coelicolor NP_631710
Mycobacterium avium NP_961576
Mycobacterium tuberculosis Pks11 NP_216181
Mycobacterium tuberculosis Pks10 NP_216176
Mycobacterium avium Pks10 NP_960303
Rhodobacter sphaeroides ZP_00006244
Geobacillus kaustophilus YP_147404
Bacillus halodurans NP_241483
Bacillus clausii YP_174488
Oceanobacillus iheyensis NP_692669
Bacillus subtilis NP_390087
Bacillus licheniformis BcsA YP_091918
Bacillus licheniformis YP_079507
Deinococcus radiodurans NP_285649
Rubrobacter xylanophilus ZP_00601608
Pseudomonas fluorescens PhlD AAM27403
Streptomyces griseus BAA33495
Saccharopolyspora erythrea RppA AAL78053
Streptomyces avermitilis NP_826307
Streptomyces antibioticus RppB BAB91444
Streptomyces coelicolor NP_625695
Streptomyces lividans RppA BAB91445
Sorangium cellulosum SoceCHS1
Streptomyces antibioticus RppA BAB91443
Oryza sativa FabH XP_474062
Arabidopsis thaliana FabH AAM14214
Synechococcus sp. WH8102 FabH CAE08763
Escherichia coli FabH NP_753271
Oceanobacillus iheyensis FabH NP_692125
Bacillus subtilis FabH NP_389015
Streptomyces avermitilis FabH BAC69542
Mycobacterium tuberculosis FabH NP_215047
Methanococcus jannaschii FabH NP_248554
Archaeoglobus fulgidus FabH NP_071237
Aeropyrum pernix FabH NP_148228
0.5
plants
fungi
bacteria
III
II
I




P. putida(pFG154)
P. putida WT
216.8
262.5
305.3
%
nm
retention time [min]


5 melonyl-CoA
THNS
II NS
(Aldol)
IIIHS
(Claisen)
IOMT
T-HN
Flavusin