Heterologous production of the antifungal polyketide antibiotic soraphen A of Sorangium cellulosum So ce26 in Streptomyces lividans

Ross Zirkle,1,2† James M. Ligon1 and István Molnár1

1Syngenta Biotechnology Inc., 3054 Cornwallis Road, Research Triangle Park, NC 27709, USA
2Department of Microbiology, North Carolina State University, 4527 South Gardner Hall, Raleigh, NC 27695, USA

The antifungal polyketide soraphen A is produced by the myxobacterium Sorangium cellulosum So ce26. The slow growth, swarming motility and general intransigence of the strain for genetic manipulations make industrial strain development, large-scale fermentation and combinatorial biosynthetic manipulation of the soraphen producer very challenging. To provide a better host for soraphen A production and molecular engineering, the biosynthetic gene cluster for this secondary metabolite was integrated into the chromosome of Streptomyces lividans ZX7. The upstream border of the gene cluster in Sor. cellulosum was defined by disrupting sorC, which is proposed to take part in the biosynthesis of methoxymalonyl-coenzyme A, to yield a Sor. cellulosum strain with abolished soraphen A production. Insertional inactivation of orf2 further upstream of sorC had no effect on soraphen A production. The genes sorR, C, D, F and E thus implicated in soraphen biosynthesis were then introduced into an engineered Str. lividans strain that carried the polyketide synthase genes sorA and sorB, and the methyltransferase gene sorM integrated into its chromosome. A benzoate-coenzyme A ligase from Rhodopseudomonas palustris was also included in some constructs. Fermentations with the engineered Str. lividans strains in the presence of benzoate and/or cinnamate yielded soraphen A. Further feeding experiments were used to delineate the biosynthesis of the benzoyl-coenzyme A starter unit of soraphen A in the heterologous host.

INTRODUCTION

Sorangium cellulosum strains are slow-growing, saprophytic and cellulytic Gram-negative myxobacteria that are also prolific producers of many different bioactive compounds, including the eukaryotic transcription inhibitor sorangicin, the cytotoxic and antitumour compound epothilone, or fungicides like jerangolid A and soraphen A (Reichenbach & Höfle, 1999). Myxobacteria also display a complex social behaviour. When starved for nutrients, myxobacteria aggregate to form a multicellular fruiting body, in which cells differentiate to form myxospores as a culmination of a complex series of developmental events. Swarming motility plays a key role in the coordinated cell movements required for the formation of fruiting bodies, and also leads to the formation of thin, spreading and merging swarms instead of colonies on agar medium in vegetative cultures.

Fermentation and industrial strain development of Sor. cellulosum strains is challenging. Sor. cellulosum strains grow slowly, with doubling times approaching 16 h (Reichenbach, 1999). Sorangium cultures do not grow readily on agar medium when plated under a certain cell density. This behaviour results from a quorum-sensing mechanism that prevents the futile growth of individual cells, reflecting the ecological need for Sorangium strains to form communities that can effectively degrade non-soluble polymeric substrates in the soil (Reichenbach, 1999). Plated above the threshold density, the primary microcolonies spread and aggregate to form macroscopic swarms. As a result of this process, distinct colonies derived from single cells do not form, making the isolation and propagation of discrete genetic events difficult.

The molecular genetic manipulation of Sor. cellulosum strains is also extremely challenging. Introduction of DNA into these strains relies on low-efficiency intergeneric conjugation from Escherichia coli. Phleomycin resistance has been the only effective selection marker described in Sor. cellulosum until recent reports on the use of a hygromycin resistance marker in strains So ce56 and So ce12 (Pradella...
et al., 2002; Kopp et al., 2004). No plasmids have been found to replicate in Sor. cellulosum strains, thus the stable maintenance of transferred DNA requires integration into the chromosome via homologous recombination between a cloned DNA fragment and a homologous chromosomal locus (Jaoua et al., 1992). Resolution of the cointegrate formed as a result of homologous recombination is extremely rare, preventing the isolation of gene replacement events.

Soraphen A is an 18-membered macrolide polyketide with an unsubstituted phenyl side ring, produced by Sor. cellulosum So ce26. Soraphen A has a unique mode of action in the inhibition of fungal acetyl-CoA carboxylase (Gerth et al., 1994; Vahlensieck et al., 1994). Due to its potent activity against plant-pathogenic fungi, soraphen A was of considerable commercial interest until it was discovered that it is a weak teratogen. The gene cluster responsible for the biosynthesis of soraphen A in Sor. cellulosum So ce26 has been cloned and sequenced (Ligon et al., 2002; Schupp et al., 1995). The cluster encodes two type I polyketide synthases (PKSs) that together contain a starter module and seven extension modules for the biosynthesis of the soraphen polyketide core, several domains that take part in the biosynthesis of the putative polyketide chain extender unit methoxymalonyl-CoA, and at least two enzymes that tailor the nascent polyketide to produce soraphen A (Ligon et al., 2002). Type I PKSs are multifunctional enzymes with domains for substrate recognition and loading (acyl transferase, AT), substrate anchoring (acyl carrier protein, ACP), condensation (ketoacyl synthase, KS) and β-keto processing (ketoreductase, KR; dehydratase, DH; enoyl reductase, ER). These domains are organized into modules and each module is responsible for one round of polyketide chain extension using acyl-CoA substrates (Hopwood, 1997). Apart from the common substrates malonyl- and methylmalonyl-CoA, the soraphen PKS also uses benzoyl-CoA for chain initiation and a ‘glycolate’ unit, probably methoxymalonyl-CoA, during chain extension.

Heterologous expression of secondary metabolic gene clusters in surrogate hosts is emerging as a viable alternative to both classical strain and fermentation process development, and molecular biological manipulation of the native producer strain (Pfeifer & Khosla, 2001). Heterologous expression of these clusters can provide strains with better fermentation characteristics that are more amenable to further optimization, or present the only fermentation-compatible alternative as with sponge- or symbiont-derived metabolites. Surrogate hosts with well developed genetic systems and higher production levels also open up the possibility of combinatorial genetic derivatization of secondary metabolites originating from genetically intractable strains for structure–activity relationship studies. Heterologous production of type I polyketides is, however, extremely challenging because of the large size (up to 10 kDa) of the synthases, the requirement for post-translational modification of the synthases by phosphopantetheinylation, the need for the availability of the activated substrates and our limited knowledge of the regulation of the interplay of primary and secondary metabolism. Despite these difficulties, the last few years have witnessed a growing number of reports on the successful production of these metabolites in surrogate hosts. Thus, Streptomyces coelicolor and its close relative Streptomyces lividans were used to express the polyketide cores for erythromycin, oleandomycin, picromycin, megalomicin, epothilone and their genetically engineered variants (Kao et al., 1994; Shah et al., 2000; Tang et al., 1999, 2000; Volchegursky et al., 2000; Xue et al., 1999), 6-Methylsalicylate and 6-deoxyerythronolide B (6-DEB, the macroline core of erythromycin) were produced in yeast (Kealey et al., 1998), 6-DEB and yersiniabactin were manufactured in E. coli (Pfeifer et al., 2001, 2003) and epothilone was biosynthesized in a Myxococcus xanthus heterologous host (Julien & Shah, 2002). Although the initial titres with these engineered strains were usually low (0.2–50 mg l\(^{-1}\)), further improvements in the production processes by classical microbiological or metabolic engineering methodologies are becoming increasingly feasible (Desai et al., 2002; Lau et al., 2002; Murli et al., 2003).

The formidable difficulties with the fermentation, classical strain improvement and molecular genetic manipulation of soraphen production in Sor. cellulosum So ce26 made it desirable, while the characterization of the gene cluster for soraphen A production made it feasible, to investigate a heterologous expression strategy of this metabolite in a more amenable host. The downstream border of the soraphen A biosynthetic gene cluster was reported previously by our group (Schupp et al., 1995). In this study, we have established the upstream border of the soraphen biosynthetic gene cluster by insertional inactivation of sorC and orf2 in Sor. cellulosum So ce26. Next, the soraphen PKS genes sorA and sorB, together with two genes downstream of sorB, the soraphen C methyltransferase sorM and orf4 of unknown function (Ligon et al., 2002), were cloned into Streptomyces expression vectors and integrated into the Str. lividans ZX7 chromosome. The genes upstream of sorA that were implicated in soraphen A production (sorR, sorC, sorD, sorF and sorE), were also introduced into the engineered Str. lividans strain together with a gene from Rhodopseudomonas palustris that encodes a benzoate-coenzyme A ligase (Eglund et al., 1995). The resulting strains of Str. lividans produced soraphen A when sodium benzoate or \(\textit{trans}\)-cinnamic acid was supplied to the fermentations. Further feeding experiments were used to shed light on the biosynthesis of the soraphen starter unit benzoyl-CoA in the engineered Str. lividans strain.

**METHODS**

**Strains and plasmids.** The streptomycin-resistant mutant SJ3 (Jaoua et al., 1992) of Sor. cellulosum So ce26 (Gerth et al., 1994) was used for these studies. E. coli DH10B (Invitrogen) and SURE (Stratagene) were used for routine cloning, and E. coli ET 12567 with plasmid pUZ8002 (MacNeil et al., 1992) was utilized for

---

**R. Zirkle, J. M. Ligon and I. Molnár**
conjugative plasmid transfer. *Str. lividans* ZX7 (John Innes Centre, Norwich, UK) is a derivative of *Str. lividans* 66. *Candida albicans* (type strain NRRL Y-12983; Agricultural Research Service Culture Collection, Peoria, IL, USA) is sensitive to soraphen A and was used as an indicator strain for soraphen production assays. *E. coli* cloning vectors pBluescript II SK (Stratagene), pCR-Blunt II-Topo (Invitrogen) and pNEB193 (New England Biolabs) were used for routine cloning. The mobilizable plasmid pCIB312 was used for conjugative plasmid transfer (Schupp et al., 1995). The expression vector pTBK (I. Molnár, unpublished) contains the thiopeptin-inducible promoter PtpA (Murakami et al., 1989) upstream of the cloning sites *PacI* and *PmlI*, the kanamycin resistance marker *aphII* (Kieser et al., 2000) and the 631 bp intact *attP* site (Kuhstoss & Rao, 1991) for site-specific recombination into the chromosome of streptomycetes. The expression vector pTBBH (I. Molnár, unpublished) contains the PtpA promoter upstream of the cloning sites *PacI* and *PmlI*, a hygromycin resistance marker (Kieser et al., 2000) and the IS117 transposase and attachment site *attM* (Henderson et al., 1990) for site-specific integration into the chromosomes of streptomycetes. Plasmid pTUE (I. Molnár, unpublished) contains the *Streptomyces* origin of replication and replication protein gene from pIJ101 (Kieser et al., 2000), the PtpA promoter upstream of the cloning sites *PacI* and *PmlI* and an erythromycin resistance marker (Kieser et al., 2000). All three expression vectors also contain the pUC18 origin of replication and an ampicillin resistance marker for propagation in *E. coli*.

**Media and growth conditions.** *E. coli* strains were grown at 37°C in Luria broth or on Luria agar with the appropriate antibiotics. *Sor. cellulosum* So ce26 S13 was grown at 30°C on S42 agar or in GS1 broth (Jaoua et al., 1992). *Str. lividans* ZX7 and its derivatives were grown on ISP-2 agar for sporulation, R5 agar for protoplast regeneration and YEME broth for soraphen production (Kieser et al., 2000), supplemented with antibiotics if necessary. *C. albicans* was grown at 30°C in Bacto potato glucose broth (PDB) or agar (PDA; Becton Dickinson).

**DNA manipulations.** Routine cloning and transformation procedures for *E. coli* were as described by Sambrook & Russell (2001). PCR for cloning was performed using *Taq* polymerase and, for the analysis of strains, with Herculase polymerase (Stratagene) with 10% (final conc.) dimethyl sulfoxide supplementing each reaction. Oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA, USA). Genomic DNA was isolated from *Sor. cellulosum* So ce26 S13 with the Puregene Kit (Genta) and from *Str. lividans* ZX7 with DNAzol (Invitrogen). The isolation of plasmid DNA from *Str. lividans* ZX7 was done with Tip-100 columns (Qiagen) as described by Kieser et al. (2000).

**Gene disruptions.** Genomic DNA of *Sor. cellulosum* So ce26 S13 was used as template for PCR reactions to amplify internal fragments of the *sorC* and *orf2* coding regions. The SCI primers (Table 1) that contain NolI sites were used to amplify a 1506 bp fragment internal to the coding region of *sorC* and cloned into pCR-Blunt II-Topo to create pTorC. Plasmid pTorC was digested with NolI and the *sorC* fragment was cloned into NolI-digested pCIB132 to generate pCIB132sorC. Similarly, the O2 primers (Table 1) were used to generate a 1995 bp fragment internal to *orf2*, and the ampiclon was cloned into pCR-Blunt-Topo to generate pTorF2ko. The *orf2* fragment was then cloned into pCIB132 to create pCIB132sorCko as described above for pCIB132sorC. Introduction of plasmids pCIB132sorCko and pCIB132orF2ko into *Sor. cellulosum* So ce26 S13 by conjugation and selection of transconjugants were carried out as described by Jaoua et al. (1992).

**Construction of soraphen expression vectors.** The primers used in the construction of the following vectors are listed in Table 1. To construct an integrative expression plasmid with the *sorA* gene, a PCR fragment containing the start codon to a *Smal* site within the *sorA* gene was amplified from cosmid pM15-5 (Ligon et al., 2002). The 370 bp PCR product was cloned as a *KpnI–SpeI* fragment by sites incorporated from the primers (PS1) into *KpnI*/ *SpeI*-digested pBluescript II SK(+) to yield pSorAs. A second PCR reaction was used to amplify the 3' end of *sorA* from cosmid p98/I (Ligon et al., 2002) from a HindIII site to the stop codon. The PCR product was cloned as an *SphI–XbaI* fragment by sites incorporated from the amplification primers (PS2) into *SphI*/ *XbaI*-digested pBluescript II SK(+) to yield pSorArev that was further digested with *SmaI* and HindIII and ligated to a 6.3 kb *SmaI–HindIII* fragment isolated from cosmids pM15-5. The resulting plasmid, pSorA3m15, was then digested with *HindIII* and ligated with a 10.6 kb HindIII fragment from cosmid p98/I. The resulting plasmid, pSorA3, was digested with *PsciI/Pmel* and the 19 kb fragment was ligated into the *PsciI/Pmel*-digested expression vector, pTBK (I. Molnár, unpublished), yielding pSorA.

An integrative expression plasmid that contains the genes *sorB*, *orf4* and *sorM* was constructed as follows. A 546 bp fragment containing the *sorB* start codon to a *BstI* site within *sorB* was PCR-amplified from cosmid p98/1 and cloned into *PsciI/Pmel*-digested pNEB193 as a *PsciI–Pmel* fragment using restriction sites from the amplification primers (PS3) to yield pSorBs. This plasmid was digested with *BstI* and *SfdI* (present in PS3) and ligated to a 29.8 kb *BstI–SfdI* fragment from cosmid p98/1 containing the rest of *sorB*, the flanking genes *orf4* and *sorM*, and the C-terminal part of *orf5*. The resulting plasmid, pNEBsorB, was digested with *PsciI/Pmel* and the 30.3 kb fragment was ligated into the *PsciI/Pmel*-digested expression vector pTBBH (I. Molnár, unpublished), yielding pSorB4M.

Two expression plasmids containing the genes upstream of *sorA* were constructed. *sorR* was PCR-amplified from cosmid pM15-5 and cloned into *XbaI/NdeI*-digested pNEB193 as an *NdeI–XbaI* fragment using sites introduced by the amplification primers (PS4) to yield pSorR. A DNA fragment containing the genes *sorC*, *sorD*, *sorF* and *sorE* was amplified by PCR from cosmid pM15-5 with primers (PS5) that introduced a ribosome-binding site (AGGAGG) upstream of *sorC*, and was cloned into the *PcrI–BluntII*-Topo vector, creating pTorC*sorD*E. This plasmid was digested with *EcoRI* and *SpeI* (sites introduced by PS5) and the 6 kb fragment was ligated into *EcoRI*–*SpeI*-digested pSor to yield pNEBsorRCDFE. The *badA* gene was PCR-amplified (PS6) from plasmid pPE202 (Egdall et al., 1995) with *SpeI* and *XbaI* restriction sites as well as the ribosome-binding site (AGGAGG) incorporated into the PCR product, and was cloned into *PcrI–BluntII*-Topo vector, creating pTbadA. This plasmid was digested with *SpeI/XbaI* and the fragment was cloned into *PcrI/XbaI*-digested pNEBsorRCDFE to create pNEBsorRCDFE+BL. The *Paci–Pmel* fragments of plasmids pNEBsorRCDFE and pNEBsorRCDFE+BL were cloned into the *Streptomyces* expression vector pTUE (I. Molnár, unpublished), yielding pSorRCDFE and pSorRCDFE+BL, respectively.

**Construction of *Str. lividans* ZX7 strains producing soraphen.** Plasmids pSorA and pSorB4M were co-transformed into protoplasts of *Str. lividans* ZX7 by polyethylene-glycol-mediated transformation (Kieser et al., 2000). The transformants were selected on R5 medium with hygromycin (100 μg ml⁻¹) and kanamycin (25 μg ml⁻¹) for the two integration events. Resistant strains were analysed by Southern hybridizations (data not shown) and strains with the expected integration events were labelled SorAB. Plasmids pSorRCDFE or pSorRCDFE+BL were transformed into protoplasts of SorAB and selected with hygromycin and kanamycin for the integration events and erythromycin (200 μg ml⁻¹) for the presence of the plasmid. These strains were analysed for the structural integrity of the two integrated expression cassettes by a series of PCR...
reactions covering the sorA (primer pairs SA1–4; Table 1) and the sorB, orf4 and sorM genes (primer pairs SB1–6, Table 1), and the ends of these overlapping 3–6 kb amplicons were sequenced and compared to the published sequence of the soraphen cluster (accession no. U24241). The plasmid pTUE was also transformed into Str. lividans ZX7 to create strain TUE. The presence of the plasmids pTUE, pSorRCDFE or pSorRCDFE +BL was ascertained by plasmid isolation and restriction analysis. Three isolates were selected from each transformed strain and named TUE, SorABRCDFE and SorABRCDFE +BL.

**Fermentation of Str. lividans strains.** To test for the production of soraphen A, spores and vegetative mycelia from about 2 cm² patches of Str. lividans strains TUE, SorAB, SorABRCDFE and SorABRCDFE +BL were inoculated into 25 ml YEME, supplemented with 20 μg erythromycin ml⁻¹ for the plasmid-containing strains, and incubated at 30 °C with shaking at 225 r.p.m. for 2 days. Six milliliters of these starter cultures was used to inoculate 600 ml YEME supplemented with 5 μg thiostrepton ml⁻¹, or 5 μg thiostrepton ml⁻¹ and 20 μg erythromycin ml⁻¹ as applicable. The cultures were grown at 30 °C with shaking at 225 r.p.m. Some of the cultures were also supplemented with one of the following (final concn): 2–5 mM trans-cinnamic acid, 5 mM L-phenylalanine, 5 mM sodium benzoate, 5 mM phenylpyruvic acid, 2–5 mM phenylacetate or 2–5 mM benzaldehyde (all from Sigma), where each stock solution had been adjusted to pH 7·5 with sodium hydroxide. After 3 days of growth, the cultures were freeze-dried and extracted with 500 ml methanol with shaking at 100 r.p.m. overnight at room temperature. The mixtures were centrifuged at 9000 g for 1 h, the methanol phases were concentrated in a SpeedVac and the extracts were redissolved in a final volume of 10 ml methanol. Fermentations were carried out with three independent isolates of each strain and repeated twice.

**Sor. cellulosum fermentations.** To analyse strains of Sor. cellulosum for the production of soraphen, 250 ml G51t media were
inoculated with 25 ml 5-day-old starter cultures and incubated at 30°C with shaking at 225 r.p.m. After 12 days, the cultures were freeze-dried and extracted with 100 ml methanol by slow mixing overnight. The mixtures were centrifuged at 9000 g for 30 min, then the methanol phases were concentrated in a SpeedVac and redissolved in 5 ml methanol. Fermentations were carried out with four

**Bioassay for soraphen A production.** Ten to 50 μl from the methanol extracts was spotted onto PDA plates and allowed to dry. An overnight culture of *C. albicans* was used to seed soft PDA (0.7% agar) and used to overlay the PDA plates containing the dried extracts. Plates were incubated at 30°C for a maximum of 3 days and observed every day for the inhibition of growth of *C. albicans*.

**HPLC and LC-MS analysis.** Extracts from strains were analysed by reverse phase HPLC and LC-MS. HPLC was performed on an HP1100 series (Agilent Technologies) fitted with a Synergy 4 micron MAX-PR 80 Å 150 × 4.6 mm column (Phenomenex). Mobile phase A was 0-1% trifluoroacetic acid (TFA) in water and mobile phase B was 0.0075% TFA in acetonitrile. Elution was performed with a gradient of 25-85% B over 15 min at 40°C with a flow rate of 0.8 ml min⁻¹ and monitored by a diode array detector. Soraphen A was detected by its absorption at 210 nm at a retention time of 14 min.

LC analysis of extracts was carried out on a HP1100 (Agilent Technologies) with the column described above. The gradient was 25–85% methanol in water in the presence of 0-1% sodium formate and 0-1% formic acid with a flow rate of 0-4 ml min⁻¹. Separation was monitored by a Finnigan LCQ Classic Mass Spectrometer (ThermoFinnigan) using positive mode ESI (capillary temp, 275°C; capillary voltage, 3-22 V; spray needle voltage, 4-54 kV).

**RESULTS**

**The upstream border of the soraphen cluster**

The soraphen biosynthetic gene cluster has been cloned, sequenced and characterized (Ligon et al., 2002). Genes downstream of the soraphen C methyltransferase gene, *sorM* (Fig. 1), showed homologies to primary metabolic genes in sugar metabolism (*orf6* and *orf7*) and to hypothetical proteins of unknown function (*orf5*). Gene disruptions in this region had no effect on the production of any soraphen congener in *Str. lividans* (Schupp et al., 1995), and consequently these genes were considered not to be part of the cluster. Genes immediately upstream of the *sorA* PKS (Fig. 1) have been proposed to take part in soraphen B oxidation (*sorR*) and in the formation of the putative polyketide extender unit methoxymalonyl-CoA (*sorC,D,E*). The putative products of genes further downstream of *sorE* had tentatively been proposed to form the border of the soraphen cluster (Ligon et al., 2002), based on their lack of homology to other proteins in the databases (*orf1* and *orf2*) or their similarity to only conserved hypothetical proteins (*orf3*). However, there was no direct evidence that suggested the involvement of the *sorC,D,E-* gene products in soraphen biosynthesis or, conversely, the absence of such involvement of the *orf1–3* gene products.

Gene disruption in *Sorangium* cells is a technically challenging task that can be accomplished by insertional inactivation only, necessitated by the prohibitively low frequency by which genetic replacements are recovered. Insertional inactivation is in itself an inefficient process and has been described to require at least 1 kb of homologous DNA to direct the recombination process (Pradella et al., 2002). Since the genes *sorD, E* and *orf3* are relatively short and thus do not offer practical targets for gene disruption, the genes *orf2* and *sorC* were targeted to establish the upstream border of the soraphen biosynthetic gene cluster. Internal fragments of *sorC* and *orf2* were cloned into the mobilizable plasmid pCIB132 and conjugated into *Str. cellosupum* (Jaoua et al., 1992) as described in Methods. Phleomycin-resistant transconjugants were isolated and the integration events at *sorC* or *orf2*, respectively, were confirmed by Southern hybridization (results not shown). Bioassays with *C. albicans* as an indicator strain (see Methods) showed the presence of large zones of growth inhibition for the wild-type SJ3 control strain and for the strains where *orf2* was insertionally inactivated, while the strains with the *sorC* knockout displayed no zones of inhibition (Fig. 2). Further analysis of the extracts by HPLC and LC-MS confirmed that the production of soraphen A remained undisturbed in the *orf2* knockout strain, but was

**Fig. 1.** Heterologous expression strategy for the soraphen cluster. The arrows are not drawn to scale and their shading is coordinated between the cluster and the expression cassettes. Arrows with stripes represent ORFs that are not proposed to take part in soraphen biosynthesis. (a) Genes in the soraphen biosynthetic locus. (b) Expression cassettes in plasmids used to reconstruct the soraphen biosynthetic cluster in *Str. lividans*. The triangles represent the PtipA promoter. Only the 5’ half of *orf5* is included in pSorB4M. BL, *badA* gene from *Rhodopseudomonas palustris* (Egland et al., 1995).
completely abolished in the sorC mutant (Fig. 2). Thus, the upstream border of the cloned soraphen locus was localized between orf2 and sorC.

**Construction of Str. lividans ZX7 strains carrying the soraphen A gene cluster**

To transfer the soraphen A gene cluster to *Str. lividans* ZX7, the soraphen A PKS genes, sorA and sorB, the methyltransferase gene, sorM, and orf4 were introduced into the chromosome of *Str. lividans* ZX7 using two compatible integrative expression plasmids. Both sorA and the sorB-orf4-sorM putative operon were placed under the control of the thiostrepton-inducible PtipA promoter (Murakami *et al*., 1989) that is widely used in *Streptomyces* expression vectors. The expression plasmid containing the sorA gene (Fig. 1) was integrated into the chromosome of *Str. lividans* ZX7 by site-specific integration mediated by the integrase and attP site from the *Streptomyces* phage φC31 (Kuhstoss & Rao, 1991). The expression plasmid (Fig. 1) containing the sorB-orf4-sorM putative operon was integrated into the chromosome of *Str. lividans* ZX7 by site-specific integration mediated by the transposase and attM site from IS117, a transposable element from *Str. coelicolor* A3(2) (Henderson *et al*., 1990). The chromosomal integration of both expression plasmids containing the sorA and the sorB-orf4-sorM cassettes were confirmed by Southern hybridization analysis (data not shown) and the corresponding strain was named *Str. lividans* SorAB.

The genes upstream of sorA were supplied to *Str. lividans* SorAB on a replicative expression vector that is compatible with the pSorA and the pSorB4M integrative vectors (Fig. 1). During the course of this work, it was brought to our attention that a small ORF encoding an ACP, not reported in our previous analysis, is located between the sorD and the sorE genes (Tin-Wein Yu, personal communication). This ORF, that we named sorF, extends from...
bp 8930 to 9181 (on the complementary strand) in the published sequence of the soraphen cluster (U24241) and would link the transcriptionally coupled sorCD genes to the sorE gene in an apparently open operon.

Thus, the sorR gene and the sorCDFE operon were placed under the control of the thiostrepton-inducible PtpA promoter in an artificial operon setting. Since benzoate is the starter unit for soraphen biosynthesis (Hill et al., 2003), but no genes for benzoate biosynthesis or activation were found clustered with the soraphen biosynthetic genes, the benzoate-coenzyme A ligase gene badA from *Rhodopseudomonas palustris* (Egliand et al., 1995) was also fused to this artificial operon in construct pSorRCDFE-badA. This gene has previously been shown to be necessary to supply benzoyl-CoA for the biosynthesis of phenyl-substituted lactones in a *Saccharopolyspora erythraea* expression system utilizing the starter module of *Sac. erythraea* (Wilkinson et al., 2001). We have inserted copies of a *Streptomyces* ribosome-binding site (AGGAGG) into sorRCDFE, the sorCDFE-badA artificial operons 12 bases upstream of the start codon of sorC and also upstream of the badA gene, as described in Methods. This ribosome-binding site occurs in the PtpA promoter and is widely distributed in genes that are highly expressed in *Streptomyces* (Strohl, 1992). The expression plasmids pSorRCDFE and pSorRCDFE-badA were transformed into *Str. lividans* SorAB to produce *Str. lividans* SorAB-badA and *Str. lividans* SorABRCDFE-badA. To ensure that no spurious recombination events disrupted the integrated PKS genes on the *Str. lividans* chromosome, the sorA and sorB-orf4-sorM expression cassettes were analysed in a series of overlapping PCR reactions as described in Methods. Due to the high level of homology amongst the repeated PKS domains, Southern hybridizations using probes derived from the soraphen PKS were of limited use for this purpose.

### Soraphen production with the engineered *Str. lividans* strains fed with benzoate

Strains of *Str. lividans* ZX7 carrying the expression vector pTUE, and *Str. lividans* SorAB, SorABRCDFE and SorABRCDFE-badA were fermented in the presence of thioestrepton to induce the expression of the sorA, and the sorB4M, sorCDFE or sorCDFE-badA operons. The fermentations were analysed by Candida bioassays, HPLC and LC-MS, but no trace of soraphen A or soraphen-related compounds were detected in any of the cultures (Table 2). L-Phenylalanine was shown previously to be the precursor of benzoyl-CoA used for soraphen biosynthesis by *Sor. cellulosum* So ce26 (Gherth et al., 2003; Hill & Thompson, 2003). The *Str. lividans* fermentations were repeated with L-phenylalanine at the same concentration, but again no production of soraphen could be detected. Since benzoate feeding and the benzoate-coenzyme A ligase encoded by badA were both found to be necessary for the production of phenyl-substituted lactones in engineered *Sac. erythraea* systems (Wilkinson et al., 2001), we have also tried feeding benzoate to the cultures. *Str. lividans* SorAB-badA and *Str. lividans* SorABRCDFE-badA, but none of the other strains, produced a compound that showed antifungal activity against *C. albicans* (Fig. 3). HPLC analysis revealed a small peak co-migrating with authentic soraphen A in extracts of *Str. lividans* SorAB-badA and *Str. lividans* SorABRCDFE, but not in *Sor. cellulosum* So ce26 (Gerth et al., 2003). Feeding trans-cinnamate led to the production of soraphen in *Str. lividans* SorAB and *Str. lividans* SorABRCDFE (but not in TUE or SorAB), in an amount similar to that obtained with benzoate supplementation (Table 2). Phenylpyruvate and phenylacetate are intermediates in the anaerobic catabolism of phenylalanine to benzoyl-CoA in *Thauera aromatica* (Schneider et al., 1997), and benzaldehyde has been suggested as a possible intermediate in benzoate biosynthesis in plants (Moore & Hertweck, 2002, and ibid.). Feeding these compounds to the engineered *Str. lividans* strains, however, did not lead to a detectable production of soraphen A (Table 2).

### Table 2. Soraphen A production by *Str. lividans* strains

<table>
<thead>
<tr>
<th>Str. <em>lividans</em> strain</th>
<th>Soraphen production upon feeding</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NF</td>
</tr>
<tr>
<td>TUE</td>
<td>–</td>
</tr>
<tr>
<td>SorAB</td>
<td>–</td>
</tr>
<tr>
<td>SorABRCDFE</td>
<td>–</td>
</tr>
<tr>
<td>SorABRCDFE-badA</td>
<td>–</td>
</tr>
</tbody>
</table>

NF, No feeding; Phe, phenylalanine; B, benzoate; Cin, cinnamate; PP, phenylpyruvate; PA, phenylacetate; BA, benzaldehyde; ND, not done.
DISCUSSION

The biosynthetic gene cluster for soraphen A production in *Sor. cellulosum* So ce26 (Schupp et al., 1995) contains genes for a Type I PKS involved in the biosynthesis of the soraphen core, substrate supply genes encoding enzymes proposed to be involved in methoxymalonyl-CoA biosynthesis and genes encoding polyketide tailoring activities that process the nascent product of the PKS (Ligon et al., 2002). Gene disruption has been used to demonstrate the involvement of the PKS genes sorA and sorB in soraphen biosynthesis and localized the downstream border of the gene cluster 5' to the sorM soraphen C O-methyltransferase (Schupp et al., 1995). In this study, we have used insertional inactivation of the sorC and the orf2 genes (Fig. 1) to establish the upstream border of the cluster. Disruption of orf2 had no effect, while insertional inactivation of sorC completely abolished soraphen A production in *Sor. cellulosum* So ce26 SJ3 (Fig. 2), indicating that the upstream border of the cluster lies somewhere between these two genes. While orf2 shows no homology to known genes, sorC forms an operon with the sorD and sorE oxidases, and the newly described sorF ACP (Tin-Wein Yu, University of Washington, personal communication), which together have been proposed to biosynthesize the unusual polyketide extender unit methoxymalonyl-CoA (Ligon et al., 2002). PKSs most commonly utilize malonyl- or methylmalonyl-CoA substrates, but unorthodox positioning of hydroxyl- or methoxy groups in some polyketides in combination with biosynthetic labelling experiments has suggested the use of extender units derived from glucose or glycerol. Such 'glycolate' units are seen in ansamitocin, geldanamyacin, leucomycin, FK520 and FK506, as well as in soraphen A (Haber et al., 1977; Hill et al., 1998; Omura et al., 1983). Methoxymalonate biosynthetic subclusters have been characterized in the ansamitocin and FK520 biosynthetic gene clusters (*asm13-17* and *fkbGHIJK*, respectively; Wu et al., 2000; Yu et al., 2002). Disruption of *asm13-17* abrogated ansamitocin production in *Actinosynnema pretiosum* (Yu et al., 2002). Furthermore, co-expression of *asm13-17* in *Str. lividans* with a cassette of the erythromycin PKS in which one of the AT domains had been replaced with an AT that apparently specifies methoxymalonate (Kato et al., 2002) led to the production of a 6-DEB derivative with a methoxymalonate incorporated in the expected position. The products of sorD and sorE, which are homologous to *asm13* and 15, and to *fkbK* and I, respectively, are proposed to be involved in the oxidation of glyceryl-ACP to hydroxymalonyl-ACP. sorF, *asm14* and

![Fig. 3. Analysis of soraphen A production in *Str. lividans* strains. (a) Bioassay of soraphen A production in fermentations with *Str. lividans* SorAB fed with benzoate (1) and three independent isolates of *Str. lividans* SorABRCDFE + BL fed with benzoate (2–4). (b) HPLC trace of an extract from a fermentation with *Str. lividans* SorABRCDFE + BL fed with benzoate (soraphen A retention time 14.0 min). (c) LC-MS selected ion chromatogram (m/z 541–545) of an extract from a benzoate-supplemented fermentation of *Str. lividans* SorABRCDFE + BL (soraphen A retention time 34.8 min). (d) LC-MS of the peak corresponding to soraphen A in the extract from a benzoate-supplemented fermentation of *Str. lividans* SorABRCDFE + BL (soraphen A M+Na ion, m/z 543).](image-url)
fkbJ all encode type II ACP that might tether the substrate during these oxidations. The loading of the unknown glycerate metabolite onto the ACP and O-methylation of hydroxymalonate to methoxymalonate requires the asm16 and 17, or the fkbH and G products, respectively. The soraphen cluster contains no homologues of asm16 or fkbH, but SorC that features AT, ACP and O-methyltransferase domains residing on a single enzyme represents a unique alternative to the functions of these gene products. The presence of two ACPs (sorF and the ACP domain within sorC) within the proposed soraphen methoxymalonate synthase is unprecedented and would require further investigation. The failure of the sorC knockout strain to produce soraphen A implies that either a functional SorC, or the products of the downstream genes sorD, F and E that might have been affected by the polar effects of the insertion, are necessary for the biosynthesis of soraphen A. In either case, the putative operon formed by the sorC, D, F and E genes is indispensable for soraphen A production and thus was included in our design of a heterologous production strategy (Fig. 1) that would also allow a further functional dissection of this operon.

Heterologous production of polyketide natural products or their derivatives may circumvent the need for strain optimization and fermentation development for every native production strain, some of which cannot (sponges, symbionts, unculturable strains) or have not (environmental DNA libraries) been grown in standard laboratory conditions. Further, horizontal transfer of natural product biosynthetic gene clusters into host strains with well-developed genetics should facilitate the creation of ‘unnatural’ natural products by combinatorial genetics (Pfeifer & Khosla, 2001), a task that has proven especially elusive with the soraphen A producer strain Sor. cellulosum So ce26.

Biosynthetic gene clusters with type I modular PKS have been expressed in several host strains. Thus, the synthase for 6-dEB has been expressed in E. coli together with the Sfp phosphopantheteinyl transferase to yield 10 mg of the macroline product 1. E. coli does not produce methylmalonyl-CoA, so the biosynthesis of this substrate of the PKS had to be engineered into the host, requiring a substantial number of genetic modifications (Dayem et al., 2002; Pfeifer et al., 2001). The Sorangium-derived natural product epothilone (Molnár et al., 2000) has been expressed in the myxobacterium host Myxococcus xanthus with an initial yield of 0.2 mg l−1 (Julien & Shah, 2002). Industrial antibiotic-producer Actinomyces strains, optimized for high production levels during fermentation, seem to be especially well suited for heterologous production of polyketides (Martin et al., 2003; Rodriguez et al., 2003), but these strains are often proprietary and their molecular genetic manipulation is far from trivial. Str. coelicolor, and its close relative Str. lividans, have been used much more extensively for the production of complex polyketides, including the macrocyclic core of erythromycin (Kao et al., 1994; Ziemann & Betlach, 1999), megalomicin (Volchegursky et al., 2000), picromycin (Tang et al., 1999), oleandomycin (Shah et al., 2000) or epothilone (Tang et al., 2000). Engineered variants of these metabolites (Tang et al., 2000; Xue et al., 1999), or glycosylated derivatives (Tang & McDaniel, 2001), were also produced in Str. lividans, with the initial titre of the heterologously produced metabolites ranging from 0.2 to 50 mg l−1 in the two host strains (Pfeifer & Khosla, 2001). Both Str. lividans and Str. coelicolor are able to synthesize malonyl- and methylmalonyl-CoA that serve as substrates for these PKS. Both strains also activate type I PKS by phosphopantheteinylolation, with the Str. coelicolor PPTase reported to be promiscuous in recognizing diverse synthases as its substrate (Cox et al., 2002). Str. coelicolor also has a wide range of transporters (Bentley et al., 2002) that allowed the secretion of heterologously produced aromatic polyketides (Pfeifer & Khosla, 2001). Based on its close taxonomic position, Str. lividans would also be likely to offer a broad-spectrum PPTase and diverse antibiotic transport systems, and has been described to support similar levels of production of type I polyketides as an expression host to those of Str. coelicolor (Ziemann & Betlach, 1999). The absence of a requirement for methylation-free DNA for genetic transformation of Str. lividans, however, made this strain our preferred choice for the heterologous production of soraphen. Although single-plasmid expression systems have proved useful for the production of complex polyketides, the discovery that PKS complexes readily assemble from subunits expressed from multiple, chromosomally integrated or plasmid-based constructs (Kuhstoss et al., 1996; Ziemann & Betlach, 2000) has simplified the assembly, transfer and later modification of these large gene clusters that often cover 30–100 kb.

In the present work, we have used a three-construct approach, integrating the PKS genes (together with a post-PKS methyltransferase and an ORF of unknown function) into two loci in the Str. lividans chromosome for added stability, and supplying the sorR oxidase and the methoxymalonate biosynthetic operon sorCDFE on an autonomously replicating plasmid for added convenience. All these genes were placed under the control of the thiostrepton-inducible promoter PtipA (Murakami et al., 1989) to provide for regulated expression of the cluster during fermentation (Ali et al., 2002; aspects under investigation in our laboratory) that is also decoupled from the developmental regulation of secondary metabolic promoters (Kyung et al., 2001).

The yield of heterologously produced soraphen A was less than 0.3 mg l−1 in these experiments, six to ten times lower than the initial reported titres with the wild-type Sor. cellulosum So ce26 strain (Gerth et al., 1994), but within the range of other type I polyketides produced in Str. lividans or Str. coelicolor (Pfeifer & Khosla, 2001). Although several years of industrial fermentation optimization and strain development had increased the titre of So ce26 fermentations to 150 mg l−1 (Gerth et al., 1994) and later to 1 g l−1 (Gerth et al., 2003), the strain is still largely
refractory to molecular genetic manipulations, and challenging for classical genetic manipulations, while its slow growth (generation time 16 h; Reichenbach, 1999) makes industrial-scale fermentations contamination-prone and the volumetric productivity is low. Despite the low initial titres, the Str. lividans soraphen production system might offer a more attractive starting point for fermentation optimization, and microbiological and genetic strain development. Indeed, both classical fermentation optimization and metabolic engineering has been used to increase low initial production levels in heterologous polyketide production systems. Introduction of an additional methylmalonyl-CoA supply pathway increased 6-dEB production in Str. coelicolor from 40 to 200 mg l\(^{-1}\) (Lombo et al., 2001), while media and fermentation regimen optimization boosted the production levels to 1·5 g l\(^{-1}\) (Desai et al., 2002). Fermentation optimization and utilization of an absorber resin to sequester the produced polyketide allowed epothilone titres to be increased from 0·2 to 23 mg l\(^{-1}\) in a Mucoraceae xanthus heterologous production system (Julien & Shah, 2002; Lau et al., 2002). A concerted approach of metabolic engineering, expression system optimization and fermentation development allowed the production of g l\(^{-1}\) quantities of 6-dEB even in E. coli (Murli et al., 2003; Pfeifer et al., 2002). The availability of the Str. coelicolor genome sequence (Bentley et al., 2002), together with our increasing understanding of its transcriptome and metabolome (Aignone-Rossa et al., 2002; Huang et al., 2001) also raises the possibility of a rational and rapid metabolic engineering of this host and, by extension, its close relative Str. lividans. Random mutagenesis coupled with whole-genome shuffling could provide an evolutionary approach aimed at rapid optimization of strains for polyketide metabolite production (Zhang et al., 2002).

The biosynthesis of soraphen A is initiated by AT1a in the soraphen starter module loading a benzoate unit from benzoyl-CoA onto ACP1a (Ligon et al., 2002; Wilkinson et al., 2001). The biosynthetic origin of the benzoyl-CoA substrate in Sor. cellulosum is not clear, however. Precursor feeding studies showed the incorporation of label into soraphen from cinnamate and phenylalanine, the latter often quoted as the ultimate source for the phenyl moiety of soraphen A (Gerth et al., 2003; Hill & Thompson, 2003). Phenylalanine might be converted to cinnamate via a phenylalanine ammonia lyase (PAL); although these enzymes are rare in bacteria in general, a PAL is present in the myxalamin biosynthetic gene cluster of the myxobacterium Stigmatella (Silakowski et al., 2001) and a PAL was also detected in the Sor. cellulosum So ce56 genome project (Gerth et al., 2003), indicating that this pathway might be more prevalent in myxobacteria. Interestingly, benzoate fed to So ce26 during fermentation does not seem to be utilized for soraphen biosynthesis (Gerth et al., 2003; Hill et al., 2003), probably due to the lack of a suitable coenzyme A ligase.

The substrate specificity of AT1a, at least in its original module context, seems to be rather strict for benzoyl-CoA, as was shown with an engineered polyketide synthetase in Sac. erythraea. This strain produced exclusively the expected phenyl-substituted lactone upon benzoate supplementation and co-expression of an additional benzoyl CoA ligase, but failed to produce any lactones without benzoate feeding and the additional ligase (Wilkinson et al., 2001). As with Sac. erythraea, there is no evidence that Str. lividans would biosynthesize benzoyl-CoA as part of its normal metabolism. Str. lividans was shown not to degrade benzoate (although it catabolizes p-hydroxybenzoate; An et al., 2000; Grund & Kutzner, 1998) and neither a benzoyl-CoA ligase, nor a PAL can be detected in the Str. coelicolor genome sequence (Bentley et al., 2002). It was thus not surprising that Str. lividans SorABRCDFE did not produce soraphen, even in fermentations where benzoate was supplied to the strain (Table 2). Co-expressing a benzoyl-CoA ligase, the badA gene from Rhodospseudomonas palustris with the soraphen gene cluster, however, allowed the Str. lividans strain SorABRCDFE+BL to produce soraphen A when benzoate was supplied to the medium. The strict requirement for both an external source of benzoate and for an activating enzyme highlights the importance of engineering an appropriate metabolic route for benzoyl-CoA into Str. lividans if this strain is to be utilized for soraphen production.

Benzoyl-CoA was shown to be biosynthesized via several routes in different organisms (Fig. 4). None of these pathways could be shown to supply benzoyl-CoA from phenylalanine for soraphen biosynthesis in Str. lividans, as exemplified by the absence of soraphen production with all the engineered strains in fermentations without feeding or with phenylalanine supplementation (Table 2). Still, we were interested to map out whether parts of these pathways exist in our host and could be utilized to provide the soraphen starter unit by feeding a suitable intermediate. Anaerobic β-oxidation converts phenylalanine to benzoyl-CoA via phenylpyruvate and phenylacetate in Thauera aromatica (Schneider & Fuchs, 1998; Schneider et al., 1997; Route 1 on Fig. 4). This catabolic pathway seems to be lacking in Str. lividans under the tested conditions, as shown by the absence of soraphen production in all the engineered strains fed with phenylpyruvate or phenylacetate. In plants and the marine actinomycete Str. maritimus, PAL converts phenylalanine to cinnameate that is activated to cinnamoyl-CoA by a CoA ligase, and this intermediate is converted to benzoyl-CoA via β-oxidation (Route 2; Moore et al., 2002). While the Str. lividans host does not harbour a PAL, cinnameate could be fed to both SorABRCDFE and SorABRCDFE+BL to effect soraphen biosynthesis. A cinnameate/4-coumarate CoA ligase has recently been characterized in Str. coelicolor (Kaneko et al., 2003), so its putative Str. lividans homologue should be able to activate cinnameate and commit it to β-oxidation. Generic β-oxidation enzymes were previously shown to complement for the conversion of cinnamoyl-CoA to benzoyl-CoA in ‘Str. maritimus’ mutants inactivated in
dedicated benzoyl-CoA biosynthetic genes located within the enterocin cluster (Xiang & Moore, 2003). Since β-oxidation proceeds via CoA-activated intermediates, and its last step, the thiolase-catalysed conversion of β-ketophenylpropionyl-CoA yields benzoyl-CoA directly without the involvement of a benzoate-CoA ligase (Moore et al., 2002), the badA gene incorporated into SorABRCDFE + BL was superfluous for soraphen biosynthesis upon cinnamate feeding as shown by the production of soraphen by SorABRCDFE. A non-oxidative, retro-aldol path from cinnamoyl-CoA through benzaldehyde was also reported in plants (Route 3; Moore & Hertweck, 2002; and ibid.). Feeding benzaldehyde to the engineered Str. lividans strains did not provide for soraphen production, so the conversion of cinnamate to benzoyl-CoA in Str. lividans appears to proceed predominantly through the β-oxidative route. As the main focus of these experiments was to establish an effective and convenient supply route for the soraphen starter unit, the stability and bioavailability of the fed precursors during fermentation were not addressed separately from their accessibility to soraphen biosynthesis.

In conclusion, two biosynthetic routes could be utilized for the provision of the soraphen starter unit in Str. lividans. The first route required benzoate feeding and the expression of a benzoyl-CoA ligase together with the soraphen biosynthetic genes (Fig. 4, Route 4). The second route required cinnamate feeding and seems to rely on an intrinsic cinnamoyl-CoA ligase and β-oxidation enzymes of the host strain (Route 2). While both benzoate and cinnamate are toxic at higher concentrations, and even at the concentrations used in these experiments (2–5 mM for cinnamate and 5 mM for benzoate) would inhibit the growth of the Str. lividans strains to some extent, continuous feeding during a controlled fermentation, or feeding less toxic precursors like phenyl-substituted fatty acids with odd-numbered chain length (a suggestion that was also made by Leadlay and coworkers; Wilkinson et al., 2001) could lead to an increased production of soraphen. Introduction of a PAL into the engineered Str. lividans strains might be able to couple the existing cinnamate degradation pathway to the endogenous phenylalanine pool; significantly, the PAL from 'Str. maritimus' has been shown to provide cinnamate when expressed in Str. coelicolor (Moore et al., 2002).

The heterologous production of soraphen thus required the augmentation of the Str. lividans host with both a methoxymalonate and a benzoyl-CoA biosynthetic pathway in addition to the expression of the soraphen polyketide synthetase. The success of this approach serves as a further example for the biosynthetic utility of this genetically well characterized strain and confirms the feasibility of engineering for the lateral transfer of secondary metabolic pathways from industrially and genetically intransigent strains into fermentation- and genetic-manipulation-friendly hosts. Increasing the titre of soraphen production by the engineered Str. lividans strains via fermentation optimization, strain improvement and metabolic engineering will allow
us to utilize the flexibility offered by this expression system to decipher the exact role of each gene product encoded in the soraphein gene cluster and to shed more light on the more arcane processes of soraphein A biosynthesis like the introduction of the C9-C10 double bond. Such approaches would also provide strains with better prospects for industrial-scale fermentations and open up new possibilities for the production of ‘unnatural’ natural products by combinatorial biosynthesis.

ACKNOWLEDGEMENTS

We would like to thank Caroline S. Harwood (University of Iowa) for the kind gift of plasmid pPE202; Tin-Wein Yu (University of Washington) for insightful discussions and for bringing the existence of sorF to our attention; Eric Miller, Robert Upchurch and Hosni Hassan (North Carolina State University) for helpful discussions and guidance to R.Z.; Makoto Ono, An Hu and Racella McNair (Syngenta Biotechnology) for their assistance with sequence analysis; James Pomes and Belle Abrera (Syngenta Biotechnology) for soraphen analysis; and Alyssa Gulledge for her assistance with fermentations.

REFERENCES


