

## Isolation, Characterization, and Heterologous Expression of the Biosynthesis Gene Cluster for the Antitumor Anthracycline Steffimycin†

Sonia Gullón,<sup>1,‡</sup> Carlos Olano,<sup>1,‡</sup> Mohamed S. Abdelfattah,<sup>2</sup> Alfredo F. Braña,<sup>1</sup> Jürgen Rohr,<sup>2</sup> Carmen Méndez,<sup>1</sup> and José A. Salas<sup>1\*</sup>

Departamento de Biología Funcional e Instituto Universitario de Oncología del Principado de Asturias (I.U.O.P.A.), Universidad de Oviedo, 33006 Oviedo, Spain,<sup>1</sup> and Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, Lexington, Kentucky 40536-0082<sup>2</sup>

Received 30 March 2006/Accepted 31 March 2006

**The biosynthetic gene cluster for the aromatic polyketide steffimycin of the anthracycline family has been cloned and characterized from “*Streptomyces steffisburgensis*” NRRL 3193. Sequence analysis of a 42.8-kbp DNA region revealed the presence of 36 open reading frames (ORFs) (one of them incomplete), 24 of which, spanning 26.5 kb, are probably involved in steffimycin biosynthesis. They code for all the activities required for polyketide biosynthesis, tailoring, regulation, and resistance but show no evidence of genes involved in L-rhamnose biosynthesis. The involvement of the cluster in steffimycin biosynthesis was confirmed by expression of a region of about 15 kb containing 15 ORFs, 11 of them forming part of the cluster, in the heterologous host *Streptomyces albus*, allowing the isolation of a biosynthetic intermediate. In addition, the expression in *S. albus* of the entire cluster, contained in a region of 34.8 kb, combined with the expression of plasmid pRHAM, directing the biosynthesis of L-rhamnose, led to the production of steffimycin. Inactivation of the *stfX* gene, coding for a putative cyclase, revealed that this enzymatic activity participates in the cyclization of the fourth ring, making the final steps in the biosynthesis of the steffimycin aglycon similar to those in the biosynthesis of jadomycin or rabelomycin. Inactivation of the *stfG* gene, coding for a putative glycosyltransferase involved in the attachment of L-rhamnose, allowed the production of a new compound corresponding to the steffimycin aglycon compound also observed in *S. albus* upon expression of the entire cluster.**

Anthracycline antibiotics are aromatic polyketide compounds widely used as therapeutics for the treatment of a variety of cancers (37). Daunorubicin and doxorubicin were the first identified anthracyclines, isolated from *Streptomyces* spp. in the early 1960s, and remain clinically widely used today as single agents or in adjuvant therapy in the treatment of hematological and solid tumors, such as acute leukemia, high-grade lymphoma, breast cancer, and bladder cancer. This has prompted the study of this family of natural products to the level of their biosynthesis in order to develop new bioactive compounds by recombinant DNA technology (33). In particular, three anthracycline compounds have been studied so far: daunorubicin-doxorubicin (18), nogalamycin (59, 56), and aclacinomycin (42, 43). These studies have revealed the assembly of this type of compound by type II polyketide synthases (PKS) that consist of several discrete proteins carrying a set of iteratively used enzyme activities.

Steffimycin (Fig. 1A) belongs to the anthracycline family of compounds and contains in its structure several distinctive characteristics not present in other anthracyclines: (i) the C-10 keto group, to which there is usually a carboxymethyl group

(nogalamycin and aclacinomycin) attached or neither substitution nor a functional group (daunorubicin); (ii) two methoxy groups at C-2 and C-8 not present in nogalamycin, aclacinomycin, or daunorubicin; and (iii) a neutral deoxysugar instead of the aminosugar normally attached to C-7 in other anthracyclines. Steffimycin was isolated in 1967 as antibiotic U-20,661 (2) from “*Streptomyces steffisburgensis*,” which also produces 8-demethoxy steffimycins (23). Later, other members of this subfamily of compounds were isolated: aranciamycin from *Streptomyces echinatus* (19, 49); steffimycin B from *Streptomyces elgreteus* (6), which produces in addition steffimycin and steffimycin C (7); and finally steffimycin D from a *Streptomyces* sp. (52).

Steffimycin was shown to be nontoxic in mice but highly cytotoxic in mammalian cell cultures (44). However, it showed a low inhibitory effect on the growth of mouse leukemia L1210 (32) and P388 cells (57). Nevertheless, steffimycin was recently shown to induce a high apoptotic response in HCT116 colon carcinoma cells expressing p53 by inducing DNA damage (12). In addition, steffimycin D derivatives inhibit the growth of *ras* oncogene-expressed cells (52), steffimycin B analogues containing chemical modification in C-3 are highly active against mouse leukemia P388 cells (57), and aranciamycin derivatives were found to inhibit DNA synthesis of Yoshida sarcoma tumor cells (4).

Here, we report the cloning and characterization of the gene cluster for steffimycin biosynthesis from “*Streptomyces steffisburgensis*” NRRL 3193. Heterologous expression in *Streptomyces albus* and disruption experiments showed the involvement of

\* Corresponding author. Mailing address: Departamento de Biología Funcional e Instituto Universitario de Oncología del Principado de Asturias (I.U.O.P.A.), Universidad de Oviedo, 33006 Oviedo, Spain. Phone and fax: 34-985-103652. E-mail: jasalas@uniovi.es.

† Supplemental material for this article may be found at <http://asm.org/>.

‡ S. Gullón and C. Olano contributed equally to this work.

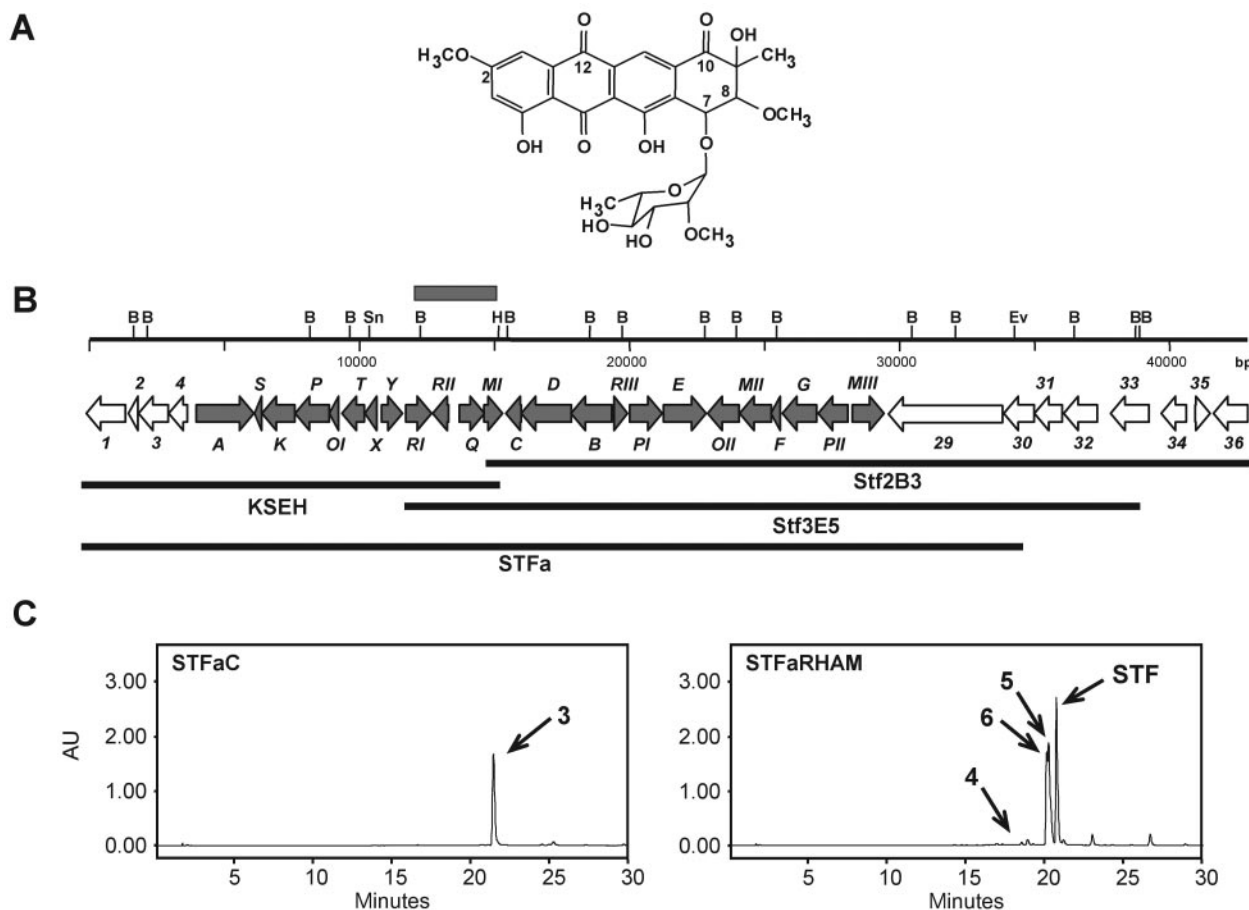


FIG. 1. (A) Structure of steffimycin. (B) Genetic organization of the steffimycin biosynthesis gene cluster. White arrows, genes not involved in steffimycin biosynthesis; gray arrows, genes involved in steffimycin biosynthesis. The gray box represents the 2.9-kb HindIII-BamHI fragment used as a probe for cosmid library screening. KSEH, STFa, Stf2B3, and Stf3E5 represent DNA regions present in plasmids pEM4KSEH and pEM4STFa and in cosmids Stf2B3 and Stf3E5, respectively. (C) HPLC analysis of *S. albus* strains STFaC (containing pEM4STFa and pEM4) and STFaRHAM (containing pEM4STFa and pRHAM). Peaks 3, 4, 5, 6, and STF correspond to compounds in Fig. 4. AU, arbitrary units; B, BamHI; H, HindIII; Sn, SnaBI; Ev, EcoRV; STF, steffimycin.

the gene cluster in the biosynthesis of steffimycin, leading to the isolation and characterization of three biosynthetic intermediates. One of them allowed the identification of a mode for the cyclization of the fourth ring different than that previously described for the biosynthesis of other anthracyclines, such as nogalamycin or daunorubicin, where an intermediate containing a carboxymethyl group at C-10 is required, and similar to that described for the biosynthesis of the angucyclines jadomycin and rabelomycin. In addition, the steffimycin gene cluster provides new enzymatic tools for the modification of other anthracyclines for the generation of novel anticancer compounds.

#### MATERIALS AND METHODS

**Strains, culture conditions, and plasmids.** “*Streptomyces steffisburgensis*” NRRL 3193, the steffimycin producer, and mutants were routinely grown on tryptone soy broth. For protoplast formation, they were grown on R5 liquid medium (21). Regeneration of protoplasts after transformation was carried out on solid R5 agar plates using standard procedures (21); after regeneration, clones were grown on agar plates containing A medium for sporulation (14). *Streptomyces albus* J1074 (8) was used as a host for expression of steffimycin genes. *Escherichia coli* DH10B (Invitrogen) and ET12567 (28) were used for subcloning. *Escherichia coli* ED8767 (38) was used to propagate a cosmid library.

Plasmids pOJ260 (3) and pEM4 (41) were used for gene replacement and gene expression, respectively. pBLUNT (Invitrogen) was used for cloning of PCR products. pFBA and pLHyg were used as donors of the apramycin resistance gene *aac(3)IV* and the hygromycin resistance gene *hyg*, respectively (26, 39). pKC505 (45) was used to construct a cosmid library. pRHAM was used for expression in *S. albus* as a source of genes coding for enzymes involved in L-rhamnose biosynthesis (46). When antibiotic selection of transformants was needed, 100  $\mu$ g/ml ampicillin, 20  $\mu$ g/ml tobramycin, 25  $\mu$ g/ml apramycin, 50  $\mu$ g/ml thiostrepton, or 50  $\mu$ g/ml hygromycin was used.

**DNA manipulation.** DNA manipulations were performed according to standard procedures for *E. coli* (47) and for *Streptomyces* (21). A cosmid library of “*S. steffisburgensis*” genomic DNA was constructed in pKC505 according to procedures published in the literature (45). Ligations were in vitro packaged using the Gigapack III Gold packaging extract kit according to the manufacturer’s handbook (Stratagene). A number of the resulting *E. coli* transductants were picked and transferred to 96-well microtiter plates containing Luria broth medium and 100  $\mu$ g/ml ampicillin. Clones were replica plated onto Luria agar plates containing ampicillin. After overnight growth at 37°C, colonies were transferred to nylon membrane filters for in situ colony hybridization analysis according to published methods (47) and screened using a labeled probe that was generated using the DIG DNA labeling and detection kit (Roche).

**Isolation of the steffimycin gene cluster.** In order to generate a mutant defective in steffimycin production, we used a sequence coding for a ketoacyl synthase (KS) from “*S. steffisburgensis*” present in the databases under accession no. ABO24979. An internal fragment encompassing nucleotides 352 to 995 was

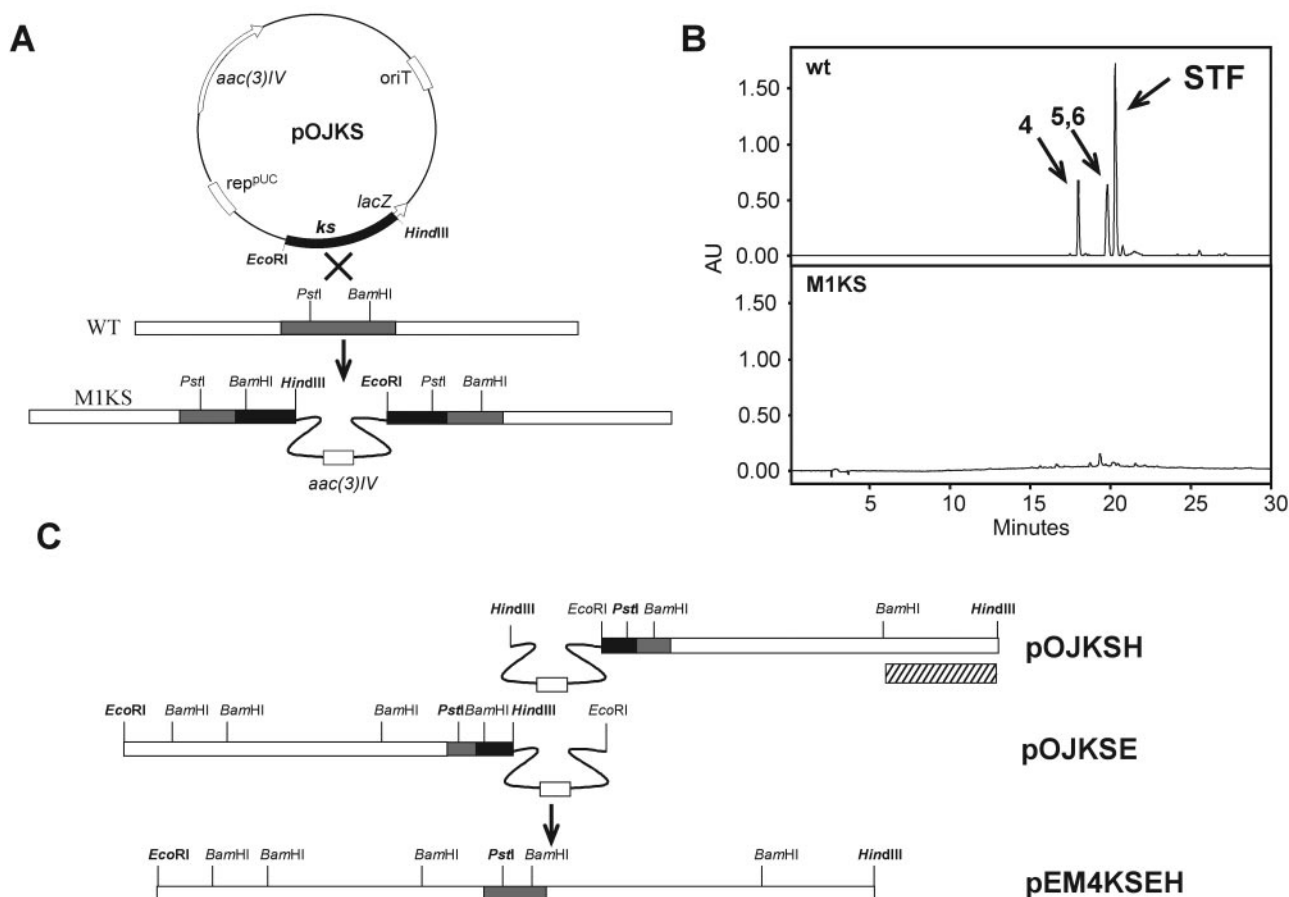


FIG. 2. (A) Strategy to generate mutant MIKS in “*S. steffisburgensis*.” The black and gray boxes represent the PCR-amplified fragments used for disruption and the same region in the “*S. steffisburgensis*” chromosome, respectively. (B) HPLC analysis of “*S. steffisburgensis*” wild type (wt) and strain MIKS. Peaks 4, 5, 6, and STF correspond to compounds in Fig. 4. (C) Strategy for recovery of the KS DNA flanking region. The dashed box represents the 2.9-kb *HindIII*-*BamHI* fragment used as a probe for cosmid library screening. *aac(3)IV*, apramycin resistance gene; AU, arbitrary units; STF, steffimycin.

obtained by PCR from “*S. steffisburgensis*” total DNA using the oligoprimers KSSTEFI (5'-TTTAAGCTTACCTGGCAGGTCTCCGAGG-3'; the *HindIII* site is underlined) and KSSTEFII (5'-AGAATTCATGGAGCTGACCGGTGTCCG-3'; the *EcoRI* site is underlined). The PCR conditions used were 97°C for 5 min; 30 cycles of 95°C for 30 s, 65°C for 45 s, and 72°C for 1 min; and a final extension cycle at 72°C for 10 min. *Pfx* DNA polymerase (Invitrogen) and 2.5% dimethylsulfoxide (DMSO) were used for all amplifications. A PCR product of 643 bp was cloned into *HindIII*-*EcoRI*-digested pOJ260 to obtain plasmid pOJKS, which was introduced into “*S. steffisburgensis*” by protoplast transformation, and transformants were selected with apramycin. Southern analysis of the mutant strain was performed using pOJKS as a probe.

Isolation of flanking DNA regions surrounding the KS was achieved by *HindIII* and *EcoRI* digestions (in two separate reactions) of total DNA recovered from the mutant strain. Religation of each digestion and transformation into *E. coli* DH10B rendered two pOJ260 derivatives containing approximately 7.5 kb each, pOJKSH (from *HindIII* digestion and religation) and pOJKSE (from *EcoRI* digestion and religation). To recover the inserts, pOJKSH was digested with *PstI*-*HindIII* and pOJKSE was digested with *PstI*-*EcoRI*, and both fragments were cloned together into pEM4 digested with *EcoRI*-*HindIII*. The resultant plasmid, pEM4KSEH, was used for expression in *S. albus* and for sequencing. A 2.9-kb *HindIII*-*BamHI* fragment isolated from pOJKSH was used as a probe for cosmid library screening (Fig. 1B and 2C).

**DNA sequencing and analysis.** DNA sequencing was performed on double-stranded DNA templates by the dideoxynucleotide chain termination method (48) with the Cy5 Autocycle Sequencing Kit (Pharmacia Biotech); an Alf-express automatic DNA sequencer (Pharmacia) was used. Computer-aided database

searching and sequence analysis were carried out with the University of Wisconsin Genetics Computer Group software (9) and the BLAST program (1).

**Cloning of the steffimycin gene cluster.** In order to clone the entire steffimycin gene cluster in a single construct, pEM4KSEH was digested with *HindIII* and *EcoRV* that cut within *stfMIII* and within the thiostrepton resistance gene in pEM4, respectively. Then, a 19.7-kb *HindIII*-*EcoRV* fragment from cosmid Stf2B3, containing from the *StfMIII* C-terminal coding region to *orf30*, was subcloned into pEM4KSEH digested with *HindIII*-*EcoRV*, generating pEM4STF. In this construct, all genes from *orf1* to *orf30* were present, with *orf30* incomplete. pEM4STF lacks a portion of the thiostrepton resistance gene, making selection in *Streptomyces* impossible. To solve this, a 1.6-kb *EcoRV*-*SmaI* fragment from pFBA (26) containing the apramycin resistance gene *aac(3)IV* was subcloned into *EcoRV*-digested pEM4STF, leading to pEM4STFa. This final construct, pEM4STFa, was used for heterologous expression in *S. albus*, together with pEM4 (41) or pRHAM (46).

**Inactivation of *stfX* and *stfG*.** The *stfX* gene was inactivated by *SnaBI* digestion of pEM4KSEH and cloning of the *hyg* gene (in the same direction as *stfX*) from *EcoRV*-digested pLHyg. The resultant construct, pEM4KSEHX-Hyg, was used for expression in *S. albus* to verify metabolite production.

To inactivate the *stfG* gene, an internal fragment of 442 bp was amplified by PCR from “*S. steffisburgensis*” total DNA using oligoprimers G3 (5'-CTGCTGTGGGGGCCGGATCTCTTC-3') and G4 (5'-CGCCGAGGGAGATGCGATCCGGAT-3'). The PCR conditions used were 97°C for 5 min; 15 cycles of 95°C for 30 s, 70°C to 55°C for 45 s, and 72°C for 1 min; 15 cycles of 95°C for 30 s, 63°C for 45 s, and 72°C for 1 min; and a final extension cycle at 72°C for 10 min. The PCR product was cloned into pBLUNT. The resultant plasmid was digested with *EcoRI*, and the fragment was cloned into pOJ260 to obtain plasmid pOJG, which

was introduced into "*S. steffisburgensis*" by protoplast transformation, and transformants were selected with apramycin. Southern analysis of the mutant strain (strain OJG) was performed using pOJG as a probe.

**Analysis of steffimycin production by HPLC and liquid chromatography (LC)-mass spectrometry.** Steffimycin production was assessed qualitatively by growing the "*S. steffisburgensis*" wild-type strain, mutants, or *S. albus* expressing pEM4 derivatives on solid R5A medium. After 5 days at 30°C, agar plugs containing 1.5 ml of agar media were extracted with 1 ml ethyl acetate, and the presence of steffimycin in the extract was analyzed by high-performance liquid chromatography (HPLC) as described below. In liquid cultures, strains were grown as a seed culture in tryptone soy broth (30 ml in a 250-ml Erlenmeyer flask). After 2 days of incubation in a rotary incubator (30°C; 250 rpm), 2.5% (vol/vol) of the cultures were used to inoculate 30 ml of R5A liquid medium. After 5 additional days of incubation with shaking, the cultures were harvested for analysis; 1-ml aliquots of fermentation broth were removed and adjusted to pH ~3 by the addition of 90% formic acid. Ethyl acetate (300  $\mu$ l) was added, and the sample was then mixed vigorously for 30 min. The phases were separated by centrifugation in a microcentrifuge, and then the ethyl acetate was removed by evaporation using a Speed-Vac. Residues were resuspended in methanol and clarified by centrifugation. Analysis of steffimycin production was performed by HPLC in a reversed-phase column (Symmetry C<sub>18</sub>; 4.6 by 250 mm; Waters) with acetonitrile and 0.1% trifluoroacetic acid (TFA) in water as solvents. A linear gradient from 10 to 100% acetonitrile in 30 min, at a flow rate of 1 ml/min, was used. Detection and spectral characterization of peaks were performed with a photodiode array detector and Millennium software (Waters), and bidimensional chromatograms were extracted at 433 nm.

LC-mass spectrometry analysis was performed in a Symmetry C<sub>18</sub> column (2.1 by 150 mm; Waters) using an Alliance chromatographic module coupled with a photodiode array detector 2996 and mass spectrometer ZQ4000 (Waters-Micromass). The chromatographic conditions were as previously described with a flow rate of 0.25 ml/min. Mass analysis was performed by electrospray ionization (ESI) in positive mode, with a capillary voltage of 3 kV and cone voltage of 20 V.

**Isolation and physicochemical properties of the new intermediates of steffimycin biosynthesis.** Compounds 1 and 2 were purified from cultures of *S. albus* carrying pEM4KSEHX-Hyg. The culture conditions were as described above. Cultures were harvested by centrifugation (12,000 rpm; 30 min), and the supernatant was filtered using a 1- $\mu$ m Mini Profile cartridge (Pall). The filtrated broth was solid phase extracted (SepPak C<sub>18</sub>; 10 g; Waters) and then eluted by a linear gradient of methanol and 0.1% TFA in water (0 to 100% in 30 min at a flow rate of 10 ml/min), taking fractions every 5 min. Fractions containing compound 1 or 2 were concentrated by evaporation using a rotavapor after the addition of 10 ml 0.1 M phosphate buffer (pH 7.0). The extracts were dissolved in 5 ml DMSO-methanol (50:50) and then chromatographed in a  $\mu$ Bondapak C<sub>18</sub> cartridge (PrepPak; 25 by 100 mm; Waters) via isocratic elution with acetonitrile and 0.1% TFA in water (55:45) at 10 ml/min. Pure compounds were solid phase extracted to remove the acid; 20.8 mg of compound 1 and 32 mg of compound 2 were recovered from 5-liter cultures.

The structures of the new compounds 1 and 2 were elucidated by nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry. The ESI mass spectra (ESI-MS) were acquired using a Finnigan LCO mass spectrometer. The electron impact (EI) ionization mass spectra were measured using a Finnigan PolarisQ mass spectrometer. The high-resolution EI ionization mass spectra were recorded at 25 eV on a JEOL JMS-700T M station (magnetic-sector instrument) at a resolution greater than 10,000. The UV spectra were recorded on a Varian model CARY50 spectrophotometer. All NMR data were recorded in (*d*<sub>5</sub>-pyridine using either a Varian Mercury 300 or a Varian Inova 400 MHz spectrometer. All NMR assignments were confirmed by heteronuclear single quantum correlation and heteronuclear multiple bond correlation spectra, allowing an unambiguous assignment of all signals.

**Nucleotide sequence accession number.** The sequence of the insert in pEM4KSEH and cosmid Stf2B3 (700-bp overlapping region; 42,877 bp) was deposited in the EMBL database under accession number AM156932.

## RESULTS AND DISCUSSION

**Isolation and characterization of the steffimycin gene cluster.** For the isolation of the steffimycin gene cluster, we took advantage of a DNA sequence from "*S. steffisburgensis*" available from the databases (accession no. ABO24979) coding for part of a KS. BLAST analysis of the deduced partial protein

coded by this DNA sequence showed high similarities to KSs involved in the biosynthesis of different type II aromatic polyketides. In particular, the best match corresponded to Snoa1 (76% identity, 86% similarity), involved in the biosynthesis of nogalamycin (59). With this information, a PCR fragment was amplified from "*S. steffisburgensis*" total DNA, verified by DNA sequencing to contain KS-like amino acid sequences, and cloned into pOJ260 generating pOJKS (Fig. 2A). This construct in a *Streptomyces* suicide vector was used for gene disruption experiments in order to generate a non-producing mutant. After transformation of "*S. steffisburgensis*" protoplasts, apramycin-resistant colonies were selected and analyzed for steffimycin production by LC-mass spectrometry. All isolated colonies were found to correspond to nonproducing mutants, and one of them (M1KS) (Fig. 2B) was selected for further studies. The correct insertion in the "*S. steffisburgensis*" chromosome was verified by Southern hybridization using the PCR product as a probe (data not shown).

Once the region containing the steffimycin gene cluster was identified, digestions and religations of total DNA from the mutant strain (as described in Materials and Methods) allowed the recovery of plasmid pOJ260 containing chromosomal-DNA fragments adjacent to the point of insertion and covering 15 kb in two fragments of approximately 7.5 kb each (pOJKSE and pOJKSH). The inserts in these constructs were cloned together, leading to plasmid pEM4KSEH (Fig. 2C).

Expression of pEM4KSEH in *S. albus* led to the production of a red compound with an HPLC retention time of 17.9 min and a mass of 357 *m/z* ([M+H]<sup>+</sup>), which might correspond to 2-*O*-demethyl-8-demethoxy-10-deoxysteffimycinone (compound 1) (Fig. 3B). Compound 1 was converted to steffimycin when fed to the "*S. steffisburgensis*" M1KS mutant (data not shown), confirming this compound as an intermediate of steffimycin biosynthesis.

Preliminary partial sequencing of the insert pEM4KSEH showed a region of approximately 4 kb presumably not involved in steffimycin biosynthesis at the EcoRI side of the insert. The remaining 11 kb of the insert contained genes possibly involved in the biosynthesis of steffimycin.

The insert in pOJKSH was used as a probe to screen a cosmid library of "*S. steffisburgensis*" chromosomal DNA. Three overlapping positive cosmid clones (Stf2B3, Stf2D7, and Stf3E5) were isolated. The insert in pEM4KSEH and cosmid Stf2B3 (700-bp overlapping region) were sequenced.

Analysis of the sequence for open reading frames (ORFs) and alignment with related sequences in databases revealed the presence of 36 ORFs (one of them incomplete). Of these, 24 ORFs (covering a region of 26.5 kb) are probably involved in steffimycin biosynthesis as structural, regulatory, and resistance genes and are flanked by 12 ORFs containing genes involved in primary metabolism and housekeeping genes (Fig. 1B and Table 1).

Several ORFs (*orf1* to *orf4* and *orf33* to *orf36*) showed by BLAST analysis significant similarities to related proteins from *Streptomyces coelicolor* A3(2) and *Streptomyces avermitilis* MA-4680. Interestingly, they also share the same genetic organization. Other ORFs (*orf29* to *orf32*) did not show similarities to proteins in databases that would suggest a role for these proteins in steffimycin biosynthesis (Table 1 and Fig. 1B).



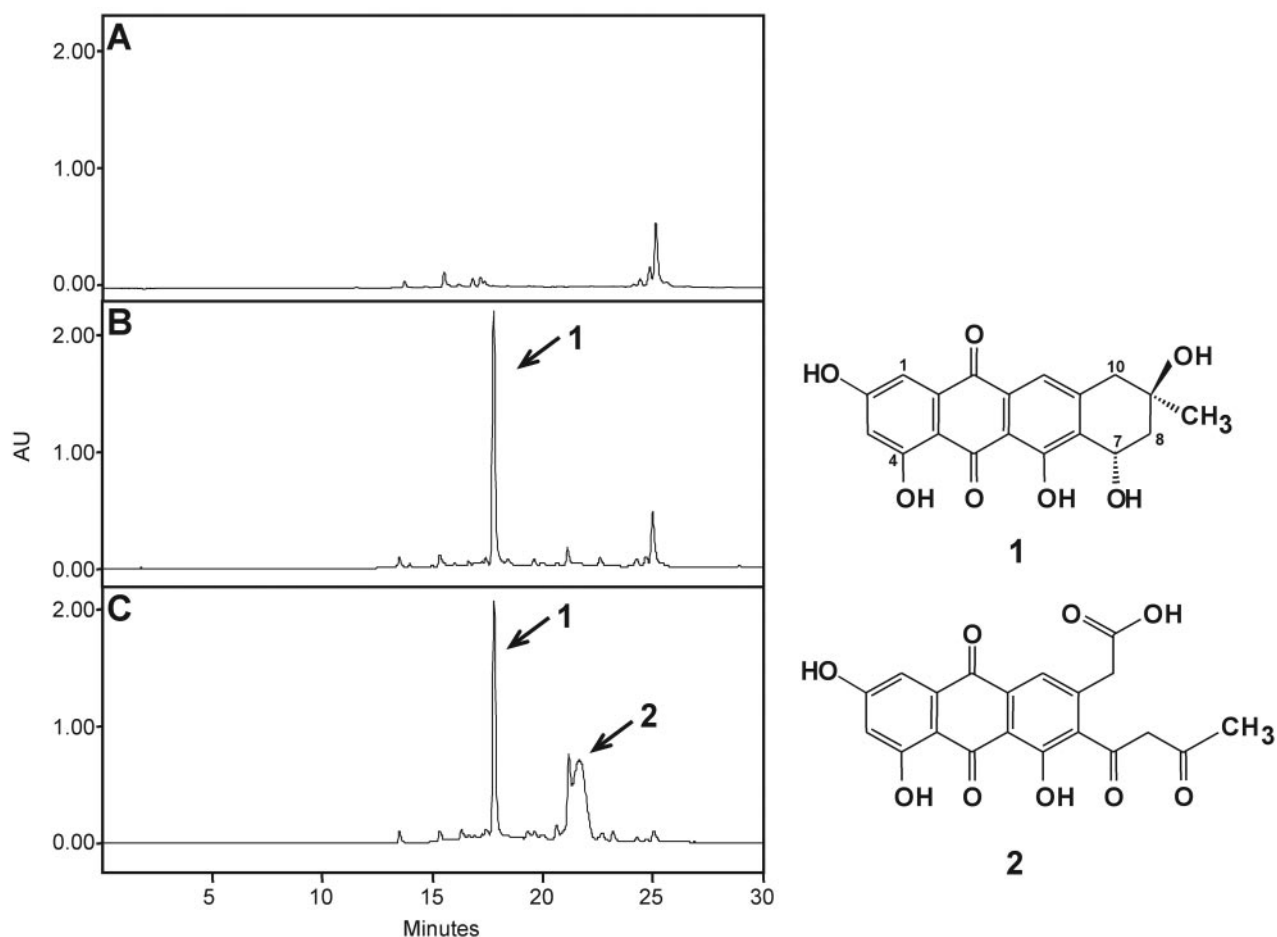


FIG. 3. HPLC analysis of *S. albus* expressing pEM4 (A), pEM4KSEH (B), and pEM4KSEHX-Hyg (C) and the products accumulated (1 and 2). AU, arbitrary units.

**Genes encoding a type II polyketide synthase and related genes.** Biosynthesis of anthracyclines begins with the condensation of small carboxylic acids by the iterative action of a type II PKS composed of two ketosynthases ( $KS_{\alpha}$  and  $KS_{\beta}$ , also known as chain length factor [CLF]) and an acyl carrier protein. Once the polyketide backbone is formed, different keto-reductases (KR), cyclases (CYC), aromatases, and oxygenases take actions to fold and decorate the polyketide into an aromatic compound.

Biosynthesis of steffimycin resembles in the initial steps that of other anthracycline antibiotics previously described. In particular, steffimycin biosynthesis mimics nogalamycin biosynthesis from 10 acetates up until the formation of 2-hydroxyl-nogalonic acid (compound 2) (Fig. 4). However, they differ in the absence of a polyketide reductase, a homologue of *SnoaD*, responsible for the lack of a C-2 hydroxyl group in nogalamycin (23, 17), a residue present in steffimycin (Fig. 1A).

In the steffimycin cluster, *stfP*, *stfK*, and *stfS* code for the minimal PKS. Then, aromatization of the first ring would likely be performed by the *stfQ* product and cyclization of the second and third rings by the *stfY* product, and the *stfOI* gene would code for an oxygenase involved in anthrone oxygenation at C-12. These functions are proposed based on the high degree

of similarity of these products to similar enzymes present in databases and involved in the biosynthesis of different anthracyclines (aclacinomycin, daunorubicin, nogalamycin, and rhodomycin) and related compounds (chartreusin, chromomycin A<sub>3</sub>, mithramycin, and simocyclinone) (Table 1).

At this point, cyclization to the fourth ring takes place, apparently by a different mechanism. In the biosynthesis of nogalamycin and daunorubicin, at least three activities are required to transform nogalonic acid to nogalaviketone or aklanonic acid to aklaviketone. The C-10 carboxyl-methyl transferases *SnoaC/DnrC* (17, 29), the nogalonic-aklanonic acid methyl ester cyclases *SnoaL/DnrD* (51, 20), and homologues of *ActVI-ORFA*, *SnoO/DpsH*, with an unclear function (17, 16), are required for cyclization (Fig. 5). However, in the equivalent steps in steffimycin biosynthesis (to transform compound 2 into compound 1), only one activity is probably required (Fig. 4 and 5). *StfX*, due to its homology to *SnoO* (Table 1), *DnrH*, *MtmX*, and *CmmX*, and also based on disruption experiments (see below), corresponds to a CYC likely involved in the closure of the fourth ring.

Finally, a C-7 KR, coded by *stfT*, similar to *SnoaF/DnrE* (17, 18), reduces the C-7 keto group (Fig. 4 and 5). In daunorubicin biosynthesis, an additional step catalyzed by the *DnrP* esterase

TABLE 1. Proposed functions of deduced steffimycin gene cluster products

ORF	No. of aa <sup>a</sup>	Proposed function in steffimycin biosynthesis	Closest similar protein, origin, (% identity/similarity), accession no.
<i>orf1</i>	483		Putative IMP dehydrogenase SCL6.18c, <i>S. coelicolor</i> A3(2), (91/96), CAB76883
<i>orf2</i>	117		Hypothetical protein SCL6.19c, <i>S. coelicolor</i> A3(2), (78/89), CAB76884
<i>orf3</i>	347		Putative transcriptional regulator SCL6.20c, <i>S. coelicolor</i> A3(2), (96/98), CAB76885
<i>orf4</i>	228		Ribulose-phosphate 3-epimerase SCL6.21c, <i>S. coelicolor</i> A3(2), (95/98), CAB76886
<i>sfrA</i>	712	Transmembrane efflux protein	Putative membrane protein similar to ActII-3 SAV6879, <i>Streptomyces avermitilis</i> MA-4680, (76/86), BAC74590
<i>stfS</i>	86	Acyl carrier protein	Acyl carrier protein NcnC, <i>Streptomyces arenae</i> DSM40737, (64/80), AAD20269
<i>stfK</i>	406	Chain length factor	Ketosynthase $\beta$ Snoa2, <i>Streptomyces nogalater</i> ATCC 27451, (72/81), CAA12018
<i>stfP</i>	422	Ketoacyl synthase	Ketosynthase, " <i>Streptomyces steffisburgensis</i> " JCM 4833T, (98/98), BAA92281
<i>stfOI</i>	111	Monoxygenase	Oxygenase SnoaB, <i>Streptomyces nogalater</i> ATCC 27451, (47/65), CAA12015
<i>stfT</i>	274	C-7 ketoreductase	Ketoreductase ChaL, <i>Streptomyces chartreuses</i> HKI-249, (46/72), CAH10174
<i>stfX</i>	146	Fourth-ring cyclase	SnoO, <i>Streptomyces nogalater</i> ATCC 27451, (52/62), AAF01807
<i>stfY</i>	259	Second- and third-ring cyclase	Putative polyketide cyclase SnoaM, <i>Streptomyces nogalater</i> ATCC 27451, (74/83), AAF01818
<i>stfRI</i>	290	SARP family regulator	Putative SARP family pathway-specific regulatory protein ORF71, <i>Streptomyces rochei</i> 7434AN4, (55/73), BAC76529
<i>stfRII</i>	196	Putative transcriptional regulator	Putative two-component response regulator Gra-orf10, <i>Streptomyces violaceoruber</i> Tu22, (49/69), CAA09631
<i>stfQ</i>	309	Aromatase	Putative aromatase SAV2383, <i>Streptomyces avermitilis</i> MA-4680, (51/63), BAC70094
<i>stfMI</i>	225	O-Methyltransferase	Putative O-methyltransferase EncK, <i>Streptomyces maritimus</i> , (40/59), AAF81726
<i>stfC</i>	187	Unknown	SWIM Zn finger Franean1DRAFT_6984, <i>Frankia</i> sp. strain EAN1pec, (56/70), ZP_00567311
<i>stfD</i>	618	Unknown	SNF2-related domain:helicase, C-terminal Franean1DRAFT_6983, <i>Frankia</i> sp. strain EAN1pec, (72/81), ZP_00567310
<i>sfrB</i>	502	Transmembrane efflux protein	Putative transmembrane efflux protein SCE22.16c, <i>Streptomyces coelicolor</i> A3(2), (41/61), CAB90983
<i>stfRIII</i>	116	Putative transcriptional regulator	Hypothetical DNA-binding protein YcgE, <i>Bacillus licheniformis</i> ATCC 14580, (31/59), AAU23698
<i>stfPI</i>	402	Monoxygenase	Cytochrome P450 Franean1DRAFT_4258, <i>Frankia</i> sp. strain EAN1pec, (46/59), ZP_00570217
<i>stfE</i>	526	Dehydrogenase	BusJ, <i>Saccharopolyspora pogona</i> NRRL 30141, (55/68), AAY88927
<i>stfOII</i>	387	Monoxygenase	Possible dioxygenase Mb3186c, <i>Mycobacterium bovis</i> AF2122/97, (51/64), NP_856831
<i>stfMII</i>	391	O-Methyltransferase	Putative O-methylase SnogY, <i>Streptomyces nogalater</i> ATCC 27451, (54/68), CAA12021
<i>stfF</i>	113	Unknown	5-Chloromuconolactone dehalogenase ClcF, <i>Rhodococcus opacus</i> 1CP, (38/66), CAD28145
<i>stfG</i>	420	Glycosyltransferase	Putative glycosyltransferase SnogE, <i>Streptomyces nogalater</i> ATCC 27451, (62/74), AAF01809
<i>stfPII</i>	366	Cytochrome P450-like	Protein similar to DnmQ AknT, <i>Streptomyces galilaeus</i> ATCC 31615, (36/47), AAF73456
<i>stfMIII</i>	392	O-Methyltransferase	Methyltransferase MycE, <i>Micromonospora griseorubida</i> , (42/58), BAC57026
<i>orf29</i>	1,404		Hypothetical protein SAV4434, <i>S. avermitilis</i> MA-4680, (37/49), NP_825611
<i>orf30</i>	374		Hypothetical protein FG11565.1, <i>Gibberella zeae</i> PH-1, (41/57), EAA78433
<i>orf31</i>	344		Trypsin-like serine proteases Npun02003978, <i>Nostoc punctiforme</i> PCC 73102, (35/51), ZP_00110193
<i>orf32</i>	415		Uncharacterized protein conserved in bacteria PphaDRAFT_1502, <i>Pelodictyon phaeoclathratiforme</i> BU-1, (41/57), ZP_00590365
<i>orf33</i>	475		Conserved hypothetical Sun-family protein SCL6.29c, <i>S. coelicolor</i> A3(2), (90/95), CAB76894
<i>orf34</i>	310		Methionyl-tRNA formyltransferase SCL6.30c, <i>S. coelicolor</i> A3(2), (89/95), CAB76895
<i>orf35</i>	183		Hypothetical protein SAV6876, <i>S. avermitilis</i> MA-4680, (88/91), BAC74587
<i>orf36</i>	422 <sup>b</sup>		Putative primosomal protein N' SCL6.32c, <i>S. coelicolor</i> A3(2), (86/90), CAB76897

<sup>a</sup> Number of amino acids of the deduced protein.

<sup>b</sup> Incomplete.

is required to remove the carboxymethyl group at C-10 after glycosylation (29, 10).

Since 2-O-demethyl-8-demethoxy-10-deoxysteffimycinone (compound 1) was isolated from *S. albus* expressing pEM4KSEH, which contains the *stfPKS* (minPKS), *stfQ* (aromatase), *stfY* (CYC), *stfOI* (oxygenase), *stfX*, and *stfT* (C7 KR) genes (Fig. 3 and 4), this implies that activities encoded by those genes are sufficient for biosynthesis of compound 1, with StfX as the unique candidate for cyclization of the fourth ring. In addition, it demonstrates that methylesterification of the carboxyl group is not

required to activate the cyclization reaction, as in daunorubicin biosynthesis (20).

**Post-PKS genes.** The next step in the biosynthesis of steffimycin is a C-2-O methylation accomplished by an O-methyltransferase, leading to 8-demethoxy-10-deoxysteffimycinone (compound 3), which might be the substrate for the glycosyltransferase (Fig. 4). There are three genes in the cluster coding for O-methyltransferases, *stfMI*, *stfMII*, and *stfMIII*. Based on the similarities shown in Table 1, StfMII could be the methyltransferase responsible for methylation of the C-2' hydroxyl

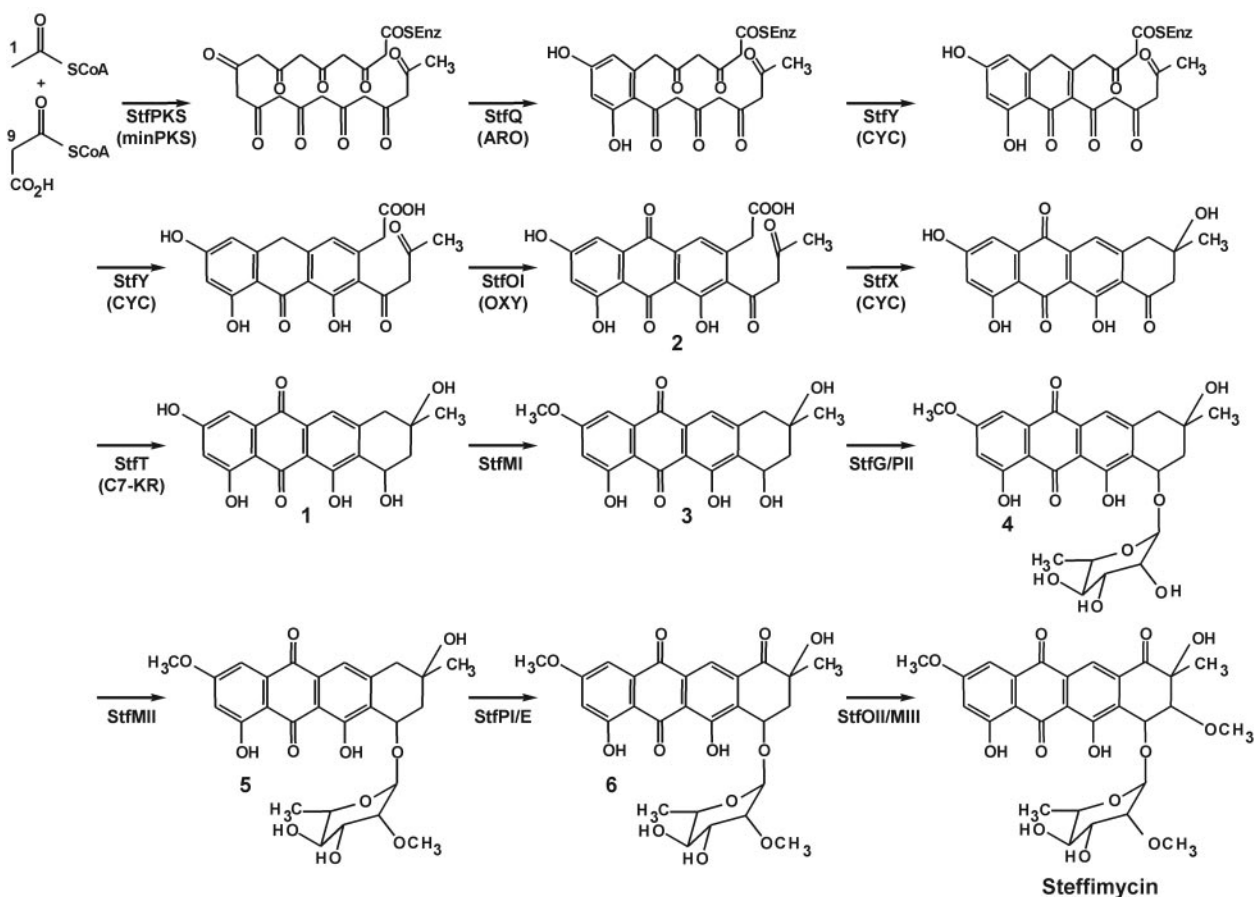


FIG. 4. Proposed pathway for the biosynthesis of steffimycin.

group in L-rhamnose, since its close homologue, SnogY, is involved in the methylation of the deoxysugar L-nogalose (56). StfMI and StfMIII might be involved in O methylation of the 2- and 8-hydroxyl groups, respectively, since coexpression of pEM4KSEH (*stfMI* incomplete) and cosmid Stf3E5 (Fig. 1B) in *S. albus* leads to production of compound 3 (Fig. 4), while coexpression of pEM4KSEH and cosmid Stf2B3 (*stfMI* incomplete) (Fig. 1B) leads to production of compound 1 (Fig. 4 and data not shown).

Two genes are involved in the attachment of the sugar moiety L-rhamnose, *stfG* and *stfP11*. StfG homologues are well-known glycosyltransferases involved in the biosynthesis of other anthracyclines. StfP11 resembles cytochromes P450 that lack the conserved Cys necessary to bind the heme prosthetic group. The roles of several enzymes belonging to that class of cytochromes, AknT (27), EryCII (60), and DesVIII (5), have been recently unraveled as proteins needed to activate their counterpart glycosyltransferases (AknS, EryCI, and DesVII) during the process of glycosylation.

Another aspect of the glycosylation process is the total absence of sugar biosynthesis genes in the cluster. This seems to be a general feature of all gene clusters containing L-rhamnose so far characterized, such as elloramycin (A. Ramos, unpublished results) and spinosyn (58). Whether those genes are shared for steffimycin and cell wall biosyntheses, as in *S. spinosa* (30), remains to be addressed.

After glycosylation of compound 3, late steps in steffimycin biosynthesis involve 2'-O methylation of L-rhamnose (probably by StfMII) and oxidation of C-10 to a keto group (probably by cytochrome P450 StfPI and dehydrogenase StfE), followed by introduction of the 8-methoxy group (likely by StfOII and StfMIII) (Fig. 4). Although there is no experimental evidence of those steps being catalyzed by the enzymes described, the sequence of events might be as shown, since compounds 8-demethoxy-2'-demethyl-steffimycin D (compound 4), 8-demethoxy-steffimycin D (compound 5), and 8-demethoxy-steffimycin (compound 6) (Fig. 2B and 4) have been previously characterized (24). Those compounds were also detected by LC-mass spectrometry from cultures of "*S. steffisburgensis*" during the development of this work (data not shown).

**Regulatory and resistance genes.** Three putative regulatory genes are located in the steffimycin gene cluster (*stfRI*, *stfRII*, and *stfRIII*).

StfRI is similar to a variety of pathway-specific *Streptomyces* antibiotic regulatory proteins (SARP) (Table 1), including DnrI from *Streptomyces peucetius* (54). StfRII resembles several putative regulatory proteins similar to two-component response regulators (Table 1), such as DnrN from *S. peucetius*. StfRII might be essential for transcription of StfRI, as DnrN controls the expression of the pathway-specific transcriptional factor DnrI involved in daunorubicin biosynthesis (15). As in *S. peucetius* (40, 15), no putative sensor kinase gene (usually

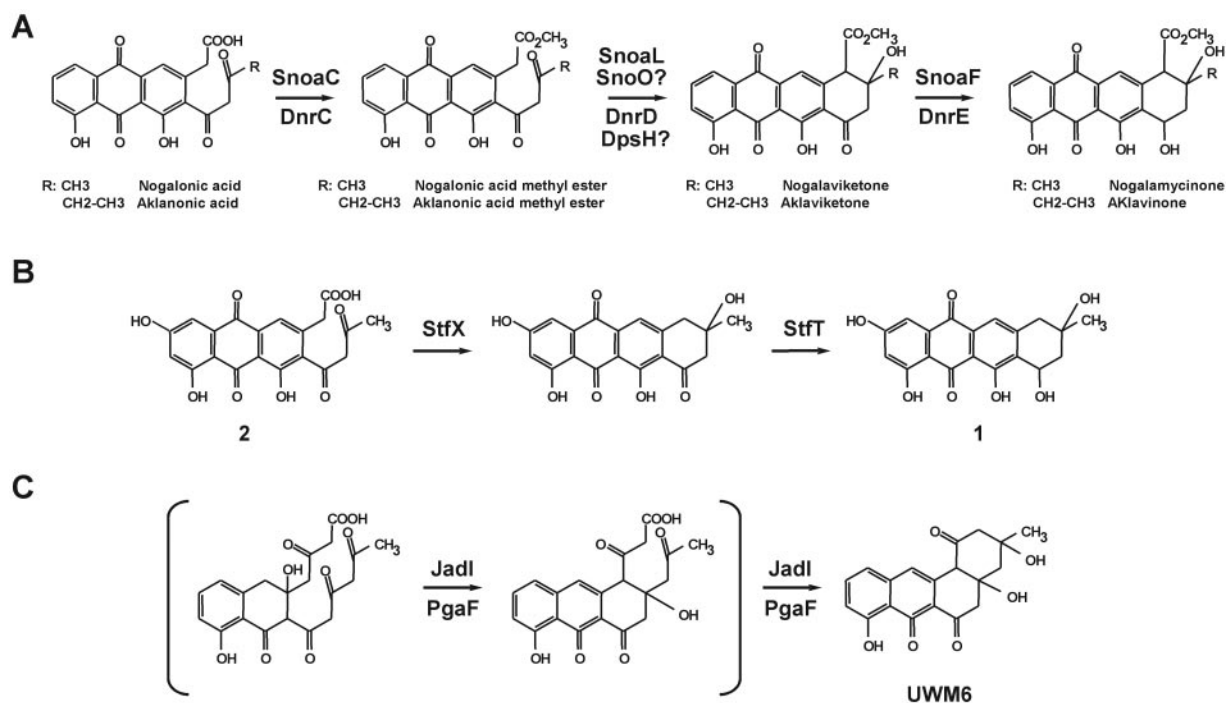


FIG. 5. Comparison of cyclization of the fourth ring in the biosyntheses of different aromatic polyketides. (A) Nogalamycin (56) and daunorubicin (18). (B) Steffimycin. (C) UWM6 in jadomycin and rabelomycin biosynthesis (22, 35).

located near the response regulator gene) was found in the steffimycin gene cluster. However, an upper level of regulation by the pleiotropic regulatory gene *blaA* that encodes tRNA-Leu (24) might be involved in steffimycin production, since there is one TTA codon in *stfRI* and two in *stfRII*.

A third regulatory protein, StfRIII, with similarity to DNA-binding proteins (Table 1) and to putative regulatory proteins of the MarR family (pfam01047) from *S. coelicolor* and *S. avermitilis*, is also present in the cluster. Whether StfRIII is involved in regulation of steffimycin biosynthesis or resistance remains to be determined, since MarR proteins are usually involved in antibiotic resistance as repressors and are also multidrug-binding proteins whose repressor activities are affected by their interactions with different compounds (50).

Two genes in the cluster (*sfrA* and *sfrB*) are candidates to be involved in steffimycin resistance. SfrA and SfrB are similar to integral membrane proteins of the MMPL family (pfam03176) and permeases of the major facilitator superfamily (MFS; COG0477). They contain 12 and 14 putative membrane-spanning helices, respectively, and are probably involved in the export of and self-resistance to steffimycin.

"*S. steffisburgensis*" shows resistance to at least 50  $\mu\text{g/ml}$  steffimycin. *S. albus* expressing Stf2B3 (*sfrB* is present) is resistant to 25  $\mu\text{g/ml}$  steffimycin, while *S. albus* expressing pEM4KSEH (*sfrA* is present) shows no resistance to steffimycin at all (data not shown). This result points to SfrB as the real resistance determinant. However, cooperation between SfrB and SfrA to confer the level of resistance observed in "*S. steffisburgensis*" cannot be excluded.

**Genes of unknown function.** Three genes in the cluster code for proteins whose alignment with proteins from databases

does not allow a suggestion of their roles in steffimycin biosynthesis. The *stfF* gene product belongs to muconolactone delta-isomerases (pfam02426), and its close homologue ClcF (Table 1) is a 5-chloromuconolactone dehalogenase that converts 5-chloromuconolactone to *cis*-dienelactone (36).

The *stfC* gene product shows similarity to hypothetical proteins containing SWIM-like zinc-chelating domains (31). StfD homologues are proteins of the SNF2 family defined by the presence of a conserved motif called the SNF2 domain (pfam00176). This domain is found in ATP-dependent proteins involved in a variety of processes, including transcription, regulation, DNA repair, DNA recombination, and chromatin unwinding. It has been mentioned that in prokaryotes lacking a SWIM domain in SWI2/SNF2 ATPases, a gene encoding a SWIM domain-containing protein is located directly upstream of the gene for a SWI2/SNF2 ATPase (31), as is the case for *stfD-stfC* (SNF2 and SWIM proteins, respectively) in the steffimycin cluster (Fig. 1B).

The involvement of StfD and StfC in resistance to steffimycin, contributing as some kind of repair system analogous to the UV repair system found in daunorubicin (25), mithramycin (13), and chromomycin (34) gene clusters, might be possible and remains to be addressed.

**Heterologous expression of the steffimycin gene cluster in *S. albus*.** To definitively verify that the isolated cluster led to the production of steffimycin, a new construct, pEM4STFa, was generated in which all genes proposed to be involved in steffimycin biosynthesis were included. Since no L-rhamnose biosynthesis genes were identified in the sequenced region, pEM4STFa was introduced in *S. albus*, together with a second plasmid, pRHAM, which is known to lead to the production of



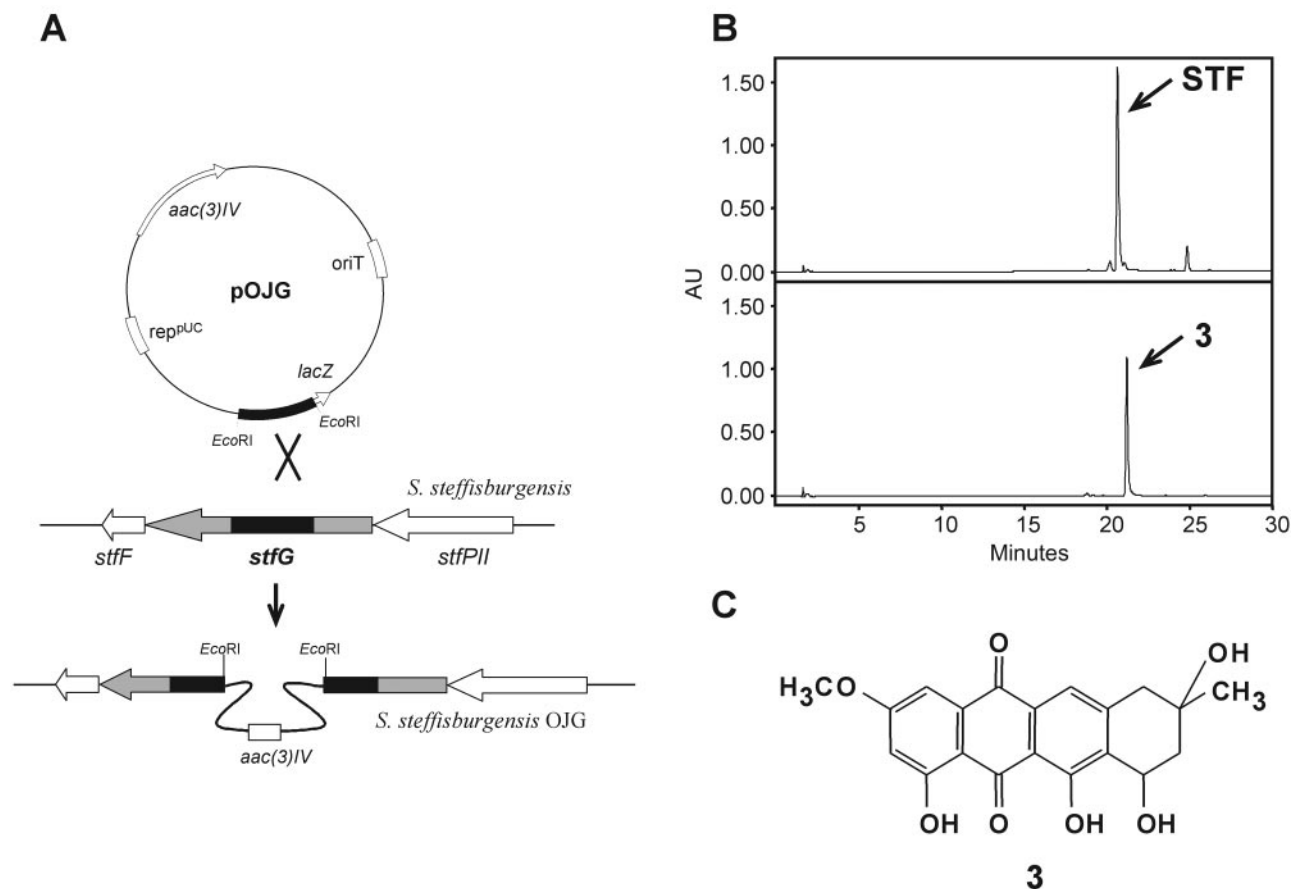


FIG. 6. (A) Strategy to generate a mutant in *stfG*. (B) HPLC analysis of “*S. steffisburgensis*” (top) and “*S. steffisburgensis*” OJG (bottom). (C) Product accumulated by the mutant. *aac(3)IV*, apramycin resistance gene; AU, arbitrary units.

L-rhamnose (46), thus generating strain STF<sub>a</sub>RHAM. In a control experiment, pEM4STF<sub>a</sub> and pEM4 were simultaneously introduced in *S. albus*, generating strain STF<sub>a</sub>C.

Analysis of products accumulated by *S. albus* STF<sub>a</sub>C (Fig. 1C) showed the production of a red compound with an HPLC retention time of 21.25 min and a mass of 371 *m/z* ( $[M+H]^+$ ), which might correspond to 8-demethoxy-10-deoxysteffimycinone, compound 3 (Fig. 4), probably steffimycin aglycon. No glycosylated compounds were detected.

Cultures of *S. albus* STF<sub>a</sub>RHAM (Fig. 1C) showed the production of a mixture of red compounds, the major one of which corresponded to steffimycin, showing an HPLC mobility of 20.7 min and a mass of 575 *m/z* ( $[M+H]^+$ ). In addition, other minor compounds were identified by LC-mass spectrometry as compounds 4, 5, and 6 (Fig. 1C and 4) with mobilities of 18.7, 20.3, and 20.1 min and masses of 517, 531, and 545 *m/z* ( $[M+H]^+$ ), respectively. Compounds 4, 5, and 6 were previously detected by LC-mass spectrometry, as mentioned above, in cultures of “*S. steffisburgensis*” (Fig. 2B). These expression experiments clearly show the presence in the sequenced steffimycin cluster of all the genes required for the biosynthesis of the polyketide, the tailoring steps, and decoration of the sugar moiety, with the exception of genes coding for the enzymes involved in the biosynthesis of L-rhamnose. In addition, plasmid pEM4STF<sub>a</sub> represents a

useful tool for the future development of steffimycin derivatives containing different sugar moieties.

**Inactivation of *stfX* and *stfG*.** Two gene inactivation experiments were performed, first, to determine if *stfX* is involved in cyclization of the fourth ring of 2-hydroxy-nogalonic acid (compound 2) (Fig. 4), and second, to isolate the substrate for StfG-mediated glycosylation and to verify its correspondence to the previously identified compound 3 in cultures of *S. albus* strain STF<sub>a</sub>C (Fig. 1C).

*stfX* was inactivated in pEM4KSEH by insertion of a copy of the *hyg* gene into a unique SnaBI site. pEM4KSEHX-Hyg was expressed in *S. albus*, and the production of red compounds was monitored in both liquid and solid media by LC-mass spectrometry. In both media, the formation of compound 1 previously observed by expression of pEM4KSEH in *S. albus* was detected. In addition, a new compound 2 with an HPLC retention of 21.59 min and a mass of 399 *m/z* ( $[M+H]^+$ ), which might correspond to 2-hydroxy-nogalonic acid (Fig. 3C), was also identified. Compound 2 was also detected as a minor compound in cultures of *S. albus* pEM4KSEH and of “*S. steffisburgensis*,” but in *S. albus* pEM4KSEHX-Hyg, the production rose 30-fold while that of compound 1 remained constant.

These results are in agreement with previous results observed by disruption of *actVI*-ORFA, involved in actinorhodin biosynthesis by *S. coelicolor* (53), where it leads to the accu-

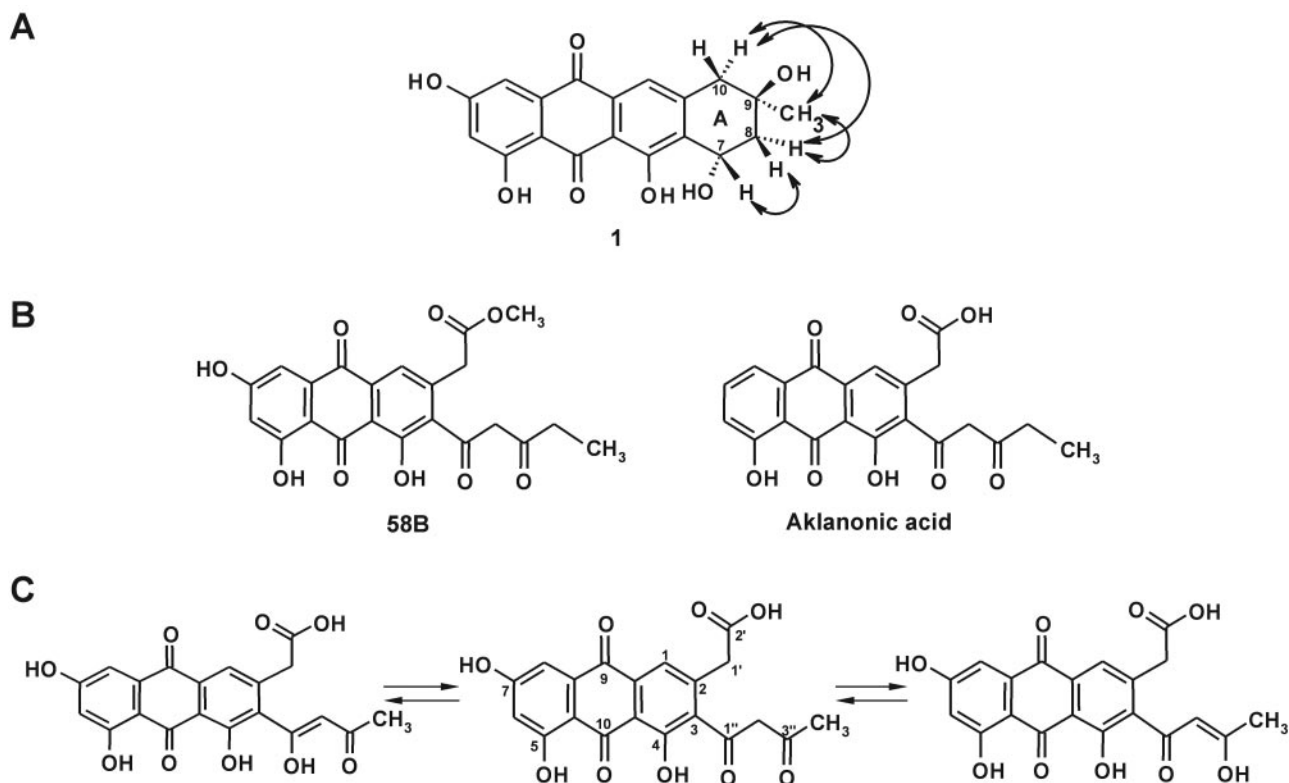


FIG. 7. (A) Nuclear Overhauser enhancement effects observed in ring A of structure 1 confirm the relative stereochemistry of the centers in ring A, as shown. (B) Chemical structures of the closest relatives of 2-hydroxynogalonic acid (compound 2), 58B (55) and aklanonic acid (11). (C) Tautomeric mixture of 2-hydroxynogalonic acid (compound 2) as evident from the  $^1\text{H}$  NMR spectrum (data for the minor tautomers are not listed). Compounds 1 and 2 correspond to compounds in Fig. 4.

mulation of possible biosynthetic intermediates with a gross reduction in actinorhodin biosynthesis. The authors suggested that ActVI-ORFA played a general role in stabilizing the multicomponent type II PKS complex, rather than in catalyzing the ring formation itself (53). Our results point to a spontaneous cyclization of the fourth ring, with the StfX protein facilitating the process.

In the biosynthesis of the angucycline antibiotics jadomycin and rabelomycin, cyclases JadI and PgaF catalyze an intramolecular aldol condensation to form the fourth angular ring concomitant with the removal of a carboxyl group (22, 35), which leads to the common intermediate UWM6 (Fig. 5). Apart from the similarities between the cyclization mechanisms of steffimycin biosynthesis and UWM6 formation, there is no significant similarity between StfX and JadI or PgaF, probably due to the structures of the products of the enzymatic reactions (anthracycline versus angucycline) and because JadI and PgaF are also involved in third-ring cyclizations.

*stfG* was inactivated by gene disruption using plasmid pOJG (Fig. 6). Cultures of “*S. steffisburgensis*” OJG revealed the production, as expected, of compound 3 with a mobility of 21.25 min and a mass of 371  $m/z$  ( $[\text{M}+\text{H}]^+$ ), which might correspond to 8-demethoxy-10-deoxysteffimycinone (Fig. 4 and 6C). This compound was previously observed in cultures of *S. albus* strain STFAC, in which no genes coding for enzymes involved in L-rhamnose biosynthesis are present (Fig. 1C). No glycosylated compounds were detected by LC-mass spectrom-

etry analysis. For further confirmation of compound 3 as an intermediate in the biosynthesis of steffimycin, it was fed to “*S. steffisburgensis*” strain M1KS, which was able to convert it to steffimycin (data not shown). In addition, disruption of *stfG* was complemented in *trans* by *stfG* expressed in pEM4, restoring steffimycin production.

**Structure elucidation of 2-O-demethyl-8-demethoxy-10-deoxysteffimycinone (compound 1) and 2-hydroxy-nogalonic acid (compound 2).** Compound 1 showed a molecular ion peak in low-resolution EI-MS at  $m/z$  356  $[\text{M}]^+$ . The ESI-MS of compound 1 displayed in the negative ion mode a molecular ion of  $m/z$  355  $[\text{M}-\text{H}]^-$ , suggesting a molecular mass of 356 g/mol. The molecular formula was determined from high-resolution EI-MS to be  $\text{C}_{19}\text{H}_{16}\text{O}_7$ . The presence of the *peri*-hydroxy-anthraquinone chromophore in compound 1 is in agreement with the observed UV maximum at a  $\lambda_{\text{max}}$  of 435 nm in the UV/visible-light spectrum. The  $^1\text{H}$  NMR spectrum of compound 1 showed two typical signals of chelated OH groups at  $\delta$  13.11 and 12.49, one singlet at  $\delta$  7.77, and two *meta*-coupled aromatic protons at  $\delta$  7.69 and 6.98. Furthermore, one proton-coupled oxygen-bound carbon at  $\delta$  5.80, two pairs of methylene protons at  $\delta$  3.18 and 3.12 and  $\delta$  2.71 and 2.42, and one methyl group at  $\delta$  1.71 were identified. The  $^{13}\text{C}$  NMR of compound 1 revealed the presence of 19 carbon atoms. The quaternary carbon signals at  $\delta$  191.2 and 182.3 could be assigned to two carbonyl groups of a quinone system. The spectra showed three aromatic methines at  $\delta$  121.6, 110.1, and 109.3 and nine

aromatic or olefinic quaternary carbon atoms, three of which, at  $\delta$  167.8, 166.2, and 162.1, could be assigned to aromatic carbons connected to oxygen atoms. Furthermore, two signals at  $\delta$  69.1 and 64.5 for aliphatic quaternary carbons connected to oxygen atoms, two methylenes at  $\delta$  46.1 and 45.8, and one methyl group at  $\delta$  30.8 were delivered. Searches in the Chemical Abstract database produced no results, pointing to a new metabolite. The structure of compound 1 was established, taking into consideration  $^1J_{C-H}$  and  $^3J_{C-H}$  long-range couplings observed in the heteronuclear single quantum correlation and the constant time inverse-detected gradient accordion resolved-heteronuclear multiple bond correlation spectra, respectively, and by comparison of the spectral data with the structurally related 8-demethoxy-2'-demethyl-steffimycin D (compound 4) (Fig. 4) (23). The stereochemistry of ring A was deduced to be (7*S*, 9*S*), because of a missing nuclear Overhauser effect correlation between 7-H and 9-CH<sub>3</sub>, while 9-CH<sub>3</sub> showed nuclear Overhauser effects with each of the H atoms of the 8- and the 10-methylene groups (Fig. 7A). The absolute stereochemistry was established through comparison with compound 4. Schmidt-Bäse et al. determined the absolute configuration of steffimycin B by X ray (49). In this study, the configurations of the stereocenters in ring A were established to be 7*R*, 8*S*, and 9*S*. The ring adopted a half-chair conformation without an extra stabilizing intramolecular H bond from 9-OH to 7-O (see the supplemental material).

Compound 2 accommodates the similarity of aklanonic acid's UV spectrum ( $\lambda_{\max}$  in methanol, 435, 287, and 258 nm.). It showed in the negative ion mode of ESI a mass spectrum of an [M-H]<sup>-</sup> peak at  $m/z$  397, corresponding to a molecular mass of 398 g/mol. The <sup>1</sup>H NMR spectrum of compound 2 revealed one sharp singlet at  $\delta$  8.21, two *meta*-coupled aromatic protons at  $\delta$  7.69 (2.1 Hz) and 6.99 (2.1 Hz), and one olefinic proton at  $\delta$  6.31. Furthermore, one methylene group at  $\delta$  4.25 and CH<sub>3</sub> protons at  $\delta$  2.07 were observed. The <sup>13</sup>C NMR spectrum of compound 2 exhibited five carbonyl carbons at  $\delta$  192.8, 190.7, 186.1, 182.0, and 173.0, with the chemical shift being consistent with a COOH group. Furthermore, quaternary carbon atoms attached to oxygen were observed at  $\delta$  168.3, 168.2, and 166.4, with four *sp*<sup>2</sup> methine carbons at  $\delta$  122.0, 110.9, 109.4, and 105.0, respectively, and six *sp*<sup>2</sup> quaternary carbon atoms. Additionally, one methylene group at  $\delta$  40.6 and one methyl signal at  $\delta$  25.0 were observed.

Searching the resulting structure 2 gave no hits in Chemical Abstracts. Therefore, compound 2 is new. Structure 2 was also confirmed by comparison of the spectral data with the structurally related aklanonic acid and compound 58B (Fig. 7B) (11, 55). Compound 2 (2-hydroxy-nogalonic acid) differs from aklanonic acid and from compound 58B in being a free acid versus a methylcarboxylate and through its acetyl starter group instead of a propionyl starter group (lower side chain). The structure elucidation was complicated by the fact that compound 2 exists in an equilibrium of three tautomeric forms (Fig. 7C). However, as is well known from simple 1,3-diketones, as well as from aklanonic acid, the keto-enol equilibrium lies far on the enol side. This is also evident for compound 2 from its <sup>1</sup>H NMR data.

In conclusion, the isolation and characterization of the steffimycin gene cluster and its heterologous expression in *S. albus* provide a new tool for the generation of novel anthracycline derivatives by combinatorial biosynthesis. The characterization

of this cluster has also provided new insights into the enzymatic mechanisms for polyketide ring cyclization. In contrast to other anthracyclines, for which two enzymes are required for cyclization of the fourth ring, in the case of steffimycin, this cyclization event can probably occur spontaneously, with the presence of the *stfX* product facilitating the process.

#### ACKNOWLEDGMENTS

This research was supported by the Spanish Ministry of Education and Science (BMC2003-00478), the Plan Regional de I+D+I del Principado de Asturias (grant GE-MED01-05), the U.S. National Institutes of Health (grant CA 102102), and Red Temática de Investigación Cooperativa de Centros de Cáncer (Ministry of Health, Spain). We thank Obra Social Cajastur for financial support to C.O. and the Spanish Ministry of Education and Science for a Ph.D. student fellowship for S.G.

The NMR core center of the University of Kentucky is acknowledged for the use of their spectrometers. We also thank Jack Goodmann (University of Kentucky Mass Spectrometry Facility) for the high-resolution mass spectra.

#### REFERENCES

- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403–410.
- Bergy, M. E., and F. Reusser. 1967. A new antibacterial agent (U-20,661) isolated from a *Streptomyces* strain. *Experientia* **23**:254–255.
- Bierman, M., R. Logan, K. O'Brien, E. T. Seno, N. Rao, and B. E. Schoner. 1992. Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene* **116**:43–49.
- Bols, M., L. Binderup, J. Hansen, and P. Rasmussen. 1992. Inhibition of collagenase by aranciamycin and aranciamycin derivatives. *J. Med. Chem.* **35**:2768–2771.
- Borisova, S. A., L. Zhao, C. E. Melancon III, C. L. Kao, and H. W. Lium. 2004. Characterization of the glycosyltransferase activity of *desVII*: analysis of and implications for the biosynthesis of macrolide antibiotics. *J. Am. Chem. Soc.* **126**:6534–6535.
- Brodasky, T. F., and F. Reusser. 1974. Steffimycin B, a new member of the steffimycin family: isolation and characterization. *J. Antibiot.* **27**:809–813.
- Brodasky, T. F., S. Mizsak, and J. R. Hoffstetter. 1985. Steffimycin C, a new member of the steffimycin anthracyclines. Isolation and structural characterization. *J. Antibiot.* **38**:849–855.
- Chater, K. F., and L. C. Wilde. 1980. *Streptomyces albus* G mutants defective in the SalGI restriction-modification system. *J. Gen. Microbiol.* **116**:323–334.
- Devereux, J., P. Haerberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**:387–395.
- Dickens, M. L., N. D. Priestley, and W. R. Strohl. 1997. In vivo and in vitro bioconversion of epsilon-rhodomycolone glycoside to doxorubicin: functions of DauP, DauK, and DoxA. *J. Bacteriol.* **179**:2641–2650.
- Eckardt, K., D. Tresselt, G. Schumann, W. Ihn, and C. Wagner. 1985. Isolation and chemical structure of aklanonic acid, an early intermediate in the biosynthesis of anthracyclines. *J. Antibiot.* **38**:1034–1039.
- Erdal, H., M. Berndtsson, J. Castro, U. Brunk, M. C. Shoshan, and S. Linder. 2005. Induction of lysosomal membrane permeabilization by compounds that activate p53-independent apoptosis. *Proc. Natl. Acad. Sci. USA* **102**:192–197.
- Fernández, E., F. Lombo, C. Méndez, and J. A. Salas. 1996. An ABC transporter is essential for resistance to the antitumor agent mithramycin in the producer *Streptomyces argillaceus*. *Mol. Gen. Genet.* **251**:692–698.
- Fernández, E., U. Weissbach, C. Sánchez Reillo, A. F. Braña, C. Méndez, J. Rohr, and J. A. Salas. 1998. Identification of two genes from *Streptomyces argillaceus* encoding glycosyltransferases involved in transfer of a disaccharide during biosynthesis of the antitumor drug mithramycin. *J. Bacteriol.* **180**:4929–4937.
- Furuya, K., and C. R. Hutchinson. 1996. The DnrN protein of *Streptomyces peucetius*, a pseudo-response regulator, is a DNA-binding protein involved in the regulation of daunorubicin biosynthesis. *J. Bacteriol.* **178**:6310–6318.
- Gerlitz, M., G. Meurer, E. Wendt-Pienkowski, K. Madduri, and C. R. Hutchinson. 1997. The effect of the daunorubicin *dpsH* gene on the choice of starter unit and cyclization pattern reveals that type II polyketide synthases can be unfaithful yet intriguing. *J. Am. Chem. Soc.* **119**:7392–7393.
- Hautala, A., S. Torkkell, K. Raty, T. Kunnari, J. Kantola, P. Mantsala, J. Hakala, and K. Ylihonko. 2003. Studies on a second and third ring cyclization in anthracycline biosynthesis. *J. Antibiot.* **56**:143–153.
- Hutchinson, C. R. 1997. Biosynthetic studies of daunorubicin and tetracenomycin C. *Chem. Rev.* **97**:2525–2536.
- Keller-Schierlein, W., J. Sauerbier, U. Vogler, and H. Zahner. 1970. Aranciamycin. *Helv. Chim. Acta* **53**:779–789.



20. **Kendrew, S. G., K. Katayama, E. Deutsch, K. Madduri, and C. R. Hutchinson.** 1999. DnrD cyclase involved in the biosynthesis of doxorubicin: purification and characterization of the recombinant enzyme. *Biochemistry* **38**: 4794–4799.
21. **Kieser, T., M. J. Bibb, M. J. Buttner, K. F. Chater, and D. A. Hopwood.** 2000. Practical *Streptomyces* genetics. The John Innes Foundation. Norwich, United Kingdom.
22. **Kulowski, K., E. Wendt-Pienkowski, L. Han, K. Yang, L. C. Vining, and C. R. Hutchinson.** 1999. Functional characterization of the *jadI* gene as a cyclase forming angucyclinones. *J. Am. Chem. Soc.* **121**:1786–1794.
23. **Kunnari, T., J. Tuikkanen, A. Hautala, J. Hakala, K. Ylihanko, and P. Mantsala.** 1997. Isolation and characterization of 8-demethoxy steffimycins and generation of 2,8-demethoxy steffimycins in *Streptomyces steffisburgensis* by the nogalamycin biosynthesis genes. *J. Antibiot.* **50**:496–501.
24. **Leskiw, B. K., E. J. Lawlor, J. M. Fernández-Abalos, and K. F. Chater.** 1991. TTA codons in some genes prevent their expression in a class of developmental, antibiotic-negative, *Streptomyces* mutants. *Proc. Natl. Acad. Sci. USA* **88**:2461–2465.
25. **Lomovskaya, N., S. K. Hong, S. U. Kim, L. Fonstein, K. Furuya, and C. R. Hutchinson.** 1996. The *Streptomyces peucetius* *drC* gene encodes a UvrA-like protein involved in daunorubicin resistance and production. *J. Bacteriol.* **178**:3238–3245.
26. **Lozano, M. J., L. L. Rensing, L. M. Quirós, A. F. Braña, E. Fernández, C. Sánchez, C. Méndez, J. Rohr, and J. A. Salas.** 2000. Characterization of two polyketide methyltransferases involved in the biosynthesis of the antitumor drug mithramycin by *Streptomyces argillaceus*. *J. Biol. Chem.* **275**:3065–3074.
27. **Lu, W., C. Leimkuhler, G. J. Gatto, Jr., R. G. Kruger, M. Oberthur, D. Kahne, and C. T. Walsh.** 2005. AknS is an activating protein for the glycosyltransferase AknS in L-aminodeoxysugar transfer to the aglycone of aclacinomycin A. *Chem. Biol.* **12**:527–534.
28. **MacNeil, D. J., K. M. Gewain, C. L. Ruby, G. Dezeny, P. H. Gibbons, and T. MacNeil.** 1992. Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* **111**:61–68.
29. **Madduri, K., and C. R. Hutchinson.** 1995. Functional characterization and transcriptional analysis of a gene cluster governing early and late steps in daunorubicin biosynthesis in *Streptomyces peucetius*. *J. Bacteriol.* **177**:3879–3884.
30. **Madduri, K., C. Waldron, and D. J. Merlo.** 2001. Rhamnose biosynthesis pathway supplies precursors for primary and secondary metabolism in *Saccharopolyspora spinosa*. *J. Bacteriol.* **183**:5632–5638.
31. **Makarova, K. S., L. Aravind, and E. V. Koonin.** 2002. SWIM, a novel Zn-chelating domain present in bacteria, archaea and eukaryotes. *Trends Biochem. Sci.* **27**:384–386.
32. **Matsuzawa, Y., T. Oki, T. Takeuchi, and H. Umezawa.** 1981. Structure-activity relationships of anthracyclines relative to cytotoxicity and effects on macromolecular synthesis in L1210 leukemia cells. *J. Antibiot.* **34**:1596–1607.
33. **Méndez, C., and J. A. Salas.** 2003. On the generation of novel anticancer drugs by recombinant DNA technology: the use of combinatorial biosynthesis to produce novel drugs. *Comb. Chem. High Throughput Screen.* **6**:513–526.
34. **Menéndez, N., M. Nur-E-Alam, A. F. Braña, J. Rohr, J. A. Salas, and C. Méndez.** 2004. Tailoring modification of deoxysugars during biosynthesis of the antitumor drug chromomycin A by *Streptomyces griseus* ssp. *griseus*. *Mol. Microbiol.* **53**:903–915.
35. **Metsä-Ketela, M., K. Palmu, T. Kunnari, K. Ylihanko, and P. Mantsala.** 2003. Engineering anthracycline biosynthesis toward angucyclines. *Antimicrob. Agents Chemother.* **47**:1291–1296.
36. **Moiseeva, O. V., I. P. Solyanikova, S. R. Kaschabek, J. Groning, M. Thiel, L. A. Golovleva, and M. Schlomann.** 2002. A new modified ortho cleavage pathway of 3-chlorocatechol degradation by *Rhodococcus opacus* 1CP: genetic and biochemical evidence. *J. Bacteriol.* **184**:5282–5292.
37. **Monneret, C.** 2001. Recent developments in the field of antitumor anthracyclines. *Eur. J. Med. Chem.* **36**:483–493.
38. **Murray, N. E., W. J. Brammar, and K. Murray.** 1977. Lambdoid phages that simplify the recovery of in vitro recombinants. *Mol. Gen. Genet.* **150**:53–61.
39. **Olano, C., B. Wilkinson, C. Sánchez, S. J. Moss, R. Sheridan, V. Math, A. J. Weston, A. F. Braña, C. J. Martín, M. Oliyynyk, C. Méndez, P. F. Leadlay, and J. A. Salas.** 2004. Biosynthesis of the angiogenesis inhibitor borrelidin by *Streptomyces parvulus* Tu4055: cluster analysis and assignment of functions. *Chem. Biol.* **11**:87–97.
40. **Otten, S. L., J. Ferguson, and C. R. Hutchinson.** 1995. Regulation of daunorubicin production in *Streptomyces peucetius* by the *dnrR2* locus. *J. Bacteriol.* **177**:1216–1224.
41. **Quirós, L. M., I. Aguirrezabalaga, C. Olano, C. Méndez, and J. A. Salas.** 1998. Two glycosyltransferases and a glycosidase are involved in oleandomycin modification during its biosynthesis by *Streptomyces antibioticus*. *Mol. Microbiol.* **28**:1177–1185.
42. **Raty, K., T. Kunnari, J. Hakala, P. Mantsala, and K. Ylihanko.** 2000. A gene cluster from *Streptomyces galilaeus* involved in glycosylation of aclarubicin. *Mol. Gen. Genet.* **264**:164–172.
43. **Raty, K., J. Kantola, A. Hautala, J. Hakala, K. Ylihanko, and P. Mantsala.** 2003. Cloning and characterization of *Streptomyces galilaeus* aclacinomycins polyketide synthase (PKS) cluster. *Gene* **293**:115–122.
44. **Reusser, F.** 1967. Mode of action of antibiotic U-20,661. *J. Bacteriol.* **93**:65–70.
45. **Richardson, M. A., S. Kuhstoss, P. Solenberg, N. A. Schaus, and R. N. Rao.** 1987. A new shuttle cosmid vector, pKC505, for *streptomycetes*: its use in the cloning of three different spiramycin-resistance genes from a *Streptomyces ambofaciens* library. *Gene* **61**:231–241.
46. **Rodríguez, L., C. Oelkers, I. Aguirrezabalaga, A. F. Braña, J. Rohr, C. Méndez, and J. A. Salas.** 2000. Generation of hybrid elloramycin analogs by combinatorial biosynthesis using genes from anthracycline-type and macroside biosynthetic pathways. *J. Mol. Microbiol. Biotechnol.* **2**:271–276.
47. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
48. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
49. **Schmidt-Bäse, K., M. Noltemeyer, E. Egert, E. Eigelt, and A. Zeeck.** 1993. Structure of the anthracycline antibiotic aranciamycin. *Acta Crystallogr. C* **49**:250–253.
50. **Schumacher, M. A., and R. G. Brennan.** 2002. Structural mechanisms of multidrug recognition and regulation by bacterial multidrug transcription factors. *Mol. Microbiol.* **45**:885–893.
51. **Sultana, A., P. Kallio, A. Jansson, J. S. Wang, J. Niemi, P. Mantsala, and G. Schneider.** 2004. Structure of the polyketide cyclase SnoaL reveals a novel mechanism for enzymatic aldol condensation. *EMBO J.* **23**:1911–1921.
52. **Suzukake-Tsuchiya, K., Y. Moriya, K. Yamazaki, M. Hori, N. Hosokawa, T. Sawa, H. Iinuma, H. Naganawa, C. Imada, and M. Hamada.** 1990. Screening of antibiotics preferentially active against *ras* oncogene-expressed cells. *J. Antibiot.* **43**:1489–1496.
53. **Taguchi, T., K. Itou, Y. Ebizuka, F. Malpartida, D. A. Hopwood, C. M. Surti, K. I. Booker-Milburn, G. R. Stephenson, and K. Ichinose.** 2000. Chemical characterisation of disruptants of the *Streptomyces coelicolor* A3(2) *actVI* genes involved in actinorhodin biosynthesis. *J. Antibiot.* **53**:144–152.
54. **Tang, L., A. Grimm, Y. X. Zhang, and C. R. Hutchinson.** 1996. Purification and characterization of the DNA-binding protein DnrI, a transcriptional factor of daunorubicin biosynthesis in *Streptomyces peucetius*. *Mol. Microbiol.* **22**:801–813.
55. **Tobe, H., A. Yoshimoto, T. Ishikura, H. Naganawa, T. Takeuchi, and H. Umezawa.** 1982. New anthracycline metabolites from two blocked mutants of *Streptomyces galilaeus* MA144-M1. *J. Antibiot.* **35**:1641–1645.
56. **Torkkell, S., T. Kunnari, K. Palmu, P. Mantsala, J. Hakala, and K. Ylihanko.** 2001. The entire nogalamycin biosynthetic gene cluster of *Streptomyces nogalater*: characterization of a 20-kb DNA region and generation of hybrid structures. *Mol. Genet. Genom.* **266**:276–288.
57. **Wiley, P. F., D. W. Elrod, and D. E. Harper.** 1988. Chemical modification of steffimycin B. *J. Antibiot.* **41**:343–351.
58. **Waldron, C., P. Matsushima, P. R. Rosteck, Jr., M. C. Broughton, J. Turner, K. Madduri, K. P. Crawford, D. J. Merlo, and R. H. Baltz.** 2001. Cloning and analysis of the spinosad biosynthetic gene cluster of *Saccharopolyspora spinosa*. *Chem. Biol.* **8**:487–499.
59. **Ylihanko, K., J. Tuikkanen, S. Jussila, L. Cong, and P. Mantsala.** 1996. A gene cluster involved in nogalamycin biosynthesis from *Streptomyces nogalater*: sequence analysis and complementation of early-block mutations in the anthracycline pathway. *Mol. Gen. Genet.* **251**:113–120.
60. **Yuan, Y., H. S. Chung, C. Leimkuhler, C. T. Walsh, D. Kahne, and S. Walker.** 2005. *In vitro* reconstitution of EryCIII activity for the preparation of unnatural macrolides. *J. Am. Chem. Soc.* **127**:14128–14129.