Molecular cloning and sequence analysis of the clorobiocin biosynthetic gene cluster: new insights into the biosynthesis of aminocoumarin antibiotics

Florence Pojer, Shu-Ming Li and Lutz Heide

Author for correspondence: Shu-Ming Li. Tel: +49 7071 2976995. Fax: +49 7071 295250.
e-mail: shuming.li@uni-tuebingen.de

Eberhard-Karls-Universität Tübingen, Pharmazeutische Biologie, Auf der Morgenstelle 8, D-72076 Tübingen, Germany

The biosynthetic gene cluster of the aminocoumarin antibiotic clorobiocin was cloned by screening of a cosmid library of *Streptomyces roseochromogenes* DS 12.976 with two heterologous probes from the novobiocin biosynthetic gene cluster. Sequence analysis revealed 27 ORFs with striking similarity to the biosynthetic gene clusters of novobiocin and coumermycin A₁. Inactivation of a putative aldolase gene, *cloR*, by in-frame deletion led to the abolishment of the production of clorobiocin. Feeding of the mutant with 3-dimethylallyl-4-hydroxybenzoic acid (Ring A of clorobiocin) restored clorobiocin production. Here, it is suggested that the formation of Ring A of clorobiocin may proceed via a retro-aldol reaction catalysed by CloR, i.e. by a mechanism different from the previously elucidated benzoic acid biosynthetic pathway in *Streptomyces maritimus*. A comparison of the gene clusters for clorobiocin, novobiocin and coumermycin A₁ showed that the structural differences between the three antibiotics were reflected remarkably well by differences in the organization of their respective biosynthetic gene clusters.

Keywords: coumermycin A₁, novobiocin, aldolase, retro-aldol reaction

INTRODUCTION

Aminocoumarin antibiotics, such as novobiocin, clorobiocin and coumermycin A₁ (Fig. 1) (Berger & Batcho, 1978), are produced by various *Streptomyces* strains and are very potent against Gram-positive pathogenic bacteria, including methicillin-resistant *Staphylococcus* strains. Novobiocin is licensed in the USA as an antibiotic for use in humans (Albamycin; Pharmacia & Upjohn).

Bacterial DNA gyrase is the target of the aminocoumarin antibiotics (Hooper *et al.*, 1982; Maxwell, 1999). X-ray crystallographic examinations (Ali *et al.*, 1993; Lewis *et al.*, 1996; Maxwell, 1993; Tsai *et al.*, 1997; Lafitte *et al.*, 2002) have demonstrated that as well as the aminocoumarin moieties and the substituted deoxysugar moieties, the prenylated 4-hydroxybenzoate moieties of clorobiocin and novobiocin are essential for the binding of these compounds to the B subunit of gyrase.

The affinity of the aminocoumarin antibiotics for

bacterial gyrase is extremely high. The inhibition constant (K_i) values of these antibiotics are in the 10 nM range, i.e. two orders of magnitude lower than those of modern fluoroquinolones. This makes the aminocoumarins very interesting starting materials for the development of new antibacterial compounds.

Previously, our group has identified the biosynthetic gene clusters for novobiocin from *Streptomyces spheroides* NCIMB 11891 (Steffensky *et al.*, 2000a) and for coumermycin A₁ from *Streptomyces rishiriensis* DSM 40489 (Wang *et al.*, 2000). In the present study, we report the identification of the clorobiocin biosynthetic gene cluster from *Streptomyces roseochromogenes* DS 12.976. A comparison of the gene clusters for clorobiocin, novobiocin and coumermycin A₁ showed that the structural differences between the three antibiotics corresponded well to the differences in the organization of their respective biosynthetic gene clusters. Furthermore, the very stringent organization of the biosynthetic genes encoding the three different aminocoumarin antibiotics into 'modules', each of which carries the complete genetic information required for the biosynthesis of the respective antibiotic, offers excellent prospects for
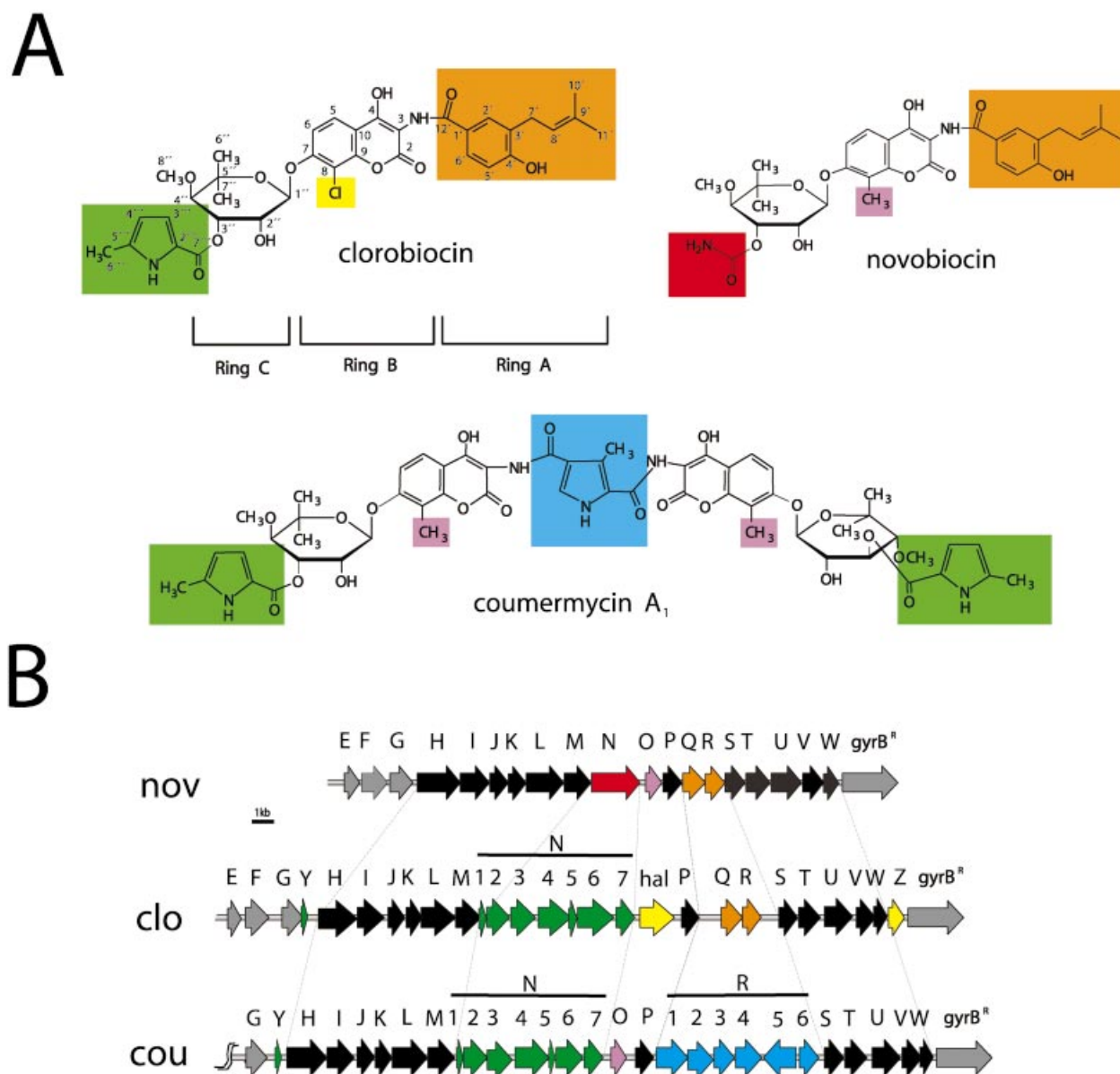
Fig. 1. (A) Structures of the aminocoumarin antibiotics. (B) Map of the clorobiocin biosynthetic gene cluster of *S. roseochromogenes* DS 12.976 (*clo*), compared to the biosynthetic gene clusters of novobiocin (*nov*) and coumermycin A₁ (*cou*). We have previously revised the nomenclature of the genes involved in coumermycin A₁ biosynthesis (GenBank accession no. AF235050).

the production of novel aminocoumarins using a combinatorial biosynthetic approach.

METHODS

Bacterial strains, plasmids and culture conditions. *S. roseochromogenes* var. *oscitans* DS 12.976 (kindly provided by Aventis) was routinely cultivated at 28 °C for 2 days in HA medium containing 1·0 % malt extract, 0·4 % yeast extract, 0·4 % glucose and 1·0 mM $CaCl_2$ (pH 7·3). For the production of clorobiocin and other secondary metabolites, wild-type and mutant strains of *S. roseochromogenes* were cultured in 500 ml

baffled flasks containing 50 ml pre-culture medium (Mancy *et al.*, 1974). After growth of these cultures for 48 h at 33 °C and 210 r.p.m., 5 ml of the cultures were inoculated into 500 ml baffled flasks containing 50 ml production medium (Mancy *et al.*, 1974). After inoculation into the production medium, cells were cultured at 33 °C and 210 r.p.m. for 7 days.

Escherichia coli XL-1 Blue MRF′ (Stratagene) was grown in liquid or on solid Luria–Bertani medium at 37 °C (Sambrook & Russell, 2001). SuperCos-1 was purchased from Stratagene. pBSKT, an integrative vector carrying carbenicillin and thiostrepton resistances, was described by Lombo *et al.* (1997). Carbenicillin (50 µg ml⁻¹) and thiostrepton (50 µg ml⁻¹) were
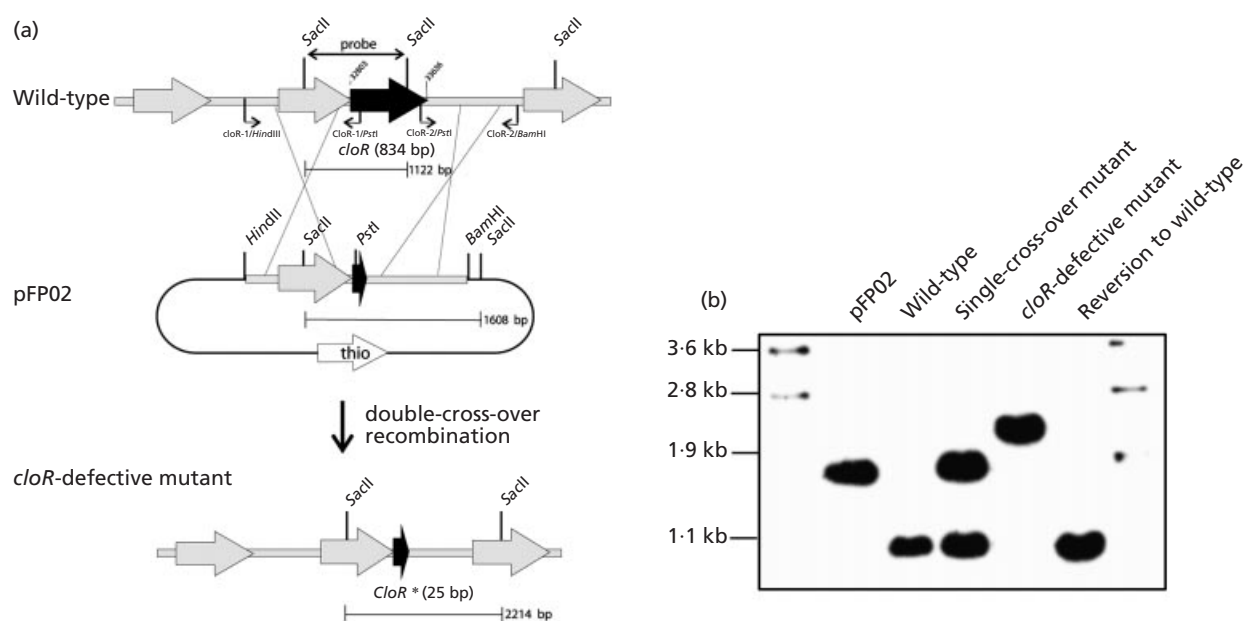
Fig. 2. Inactivation of *cloR* of the clorobiocin biosynthetic gene cluster. (a) Schematic presentation of the gene inactivation experiment. thio, Thiostrepton-resistance gene. (b) Southern blot analysis of the wild-type and mutant strains. Genomic DNA was digested with *Sac*II. The indicated 1122 bp *Sac*II fragment containing part of *cloR* in (a) was used as the probe.

used in the medium for selection of recombinant plasmids and strains.

Genetic procedures. Standard methods for DNA isolation and manipulation were performed as described by Kieser *et al.* (2000). DNA fragments were isolated from agarose gels using a NucleoSpin 2 in 1 extraction kit (Macherey-Nagel). Isolation of cosmids and plasmids was carried out with ion-exchange columns (Nucleobond AX kit; Macherey-Nagel). Genomic DNA was isolated from *Streptomyces* strains by lysozyme treatment and phenol/chloroform extraction.

Construction and screening of the cosmid library. Chromosomal DNA of *S. roseochromogenes* was partially digested with *Sau*3AI, dephosphorylated and then ligated into the *Bam*HI sites of SuperCos-1. The ligation products were packaged with Gigapack III XL (Stratagene, Heidelberg; Germany) and transduced into *E. coli* XL-1 Blue MRF′.

Southern blot analysis was performed on Hybond-N membranes (Amersham) with digoxigenin-labelled probes by using the DIG high prime DNA labelling and detection kit II (Roche Molecular Biochemicals). Two probes, one containing part of the dTDP-glucose 4,6-dehydratase gene *novT* (Steffensky *et al.*, 2000a) and one containing a 1·58 kb *Sph*I–*Bam*HI fragment of the novobiocic acid synthetase gene *novL* (Steffensky *et al.*, 2000b), were used for hybridizations.

DNA sequencing and computer-assisted sequence analysis. Double-stranded sequencing of the entire cosmid K1F2 (carrying an insert of 42 291 bp) was performed by the dideoxynucleotide chain termination method on a LI-COR automated sequencer (MWG-Biotech) using a shotgun library with DNA fragments of approximately 1·5–2·0 kb in length.

The DNASIS software package (version 2.1; Hitachi Software Engineering) was used for sequence analysis. Amino acid sequence homology searches were carried out in the GenBank database by using the BLAST program (release 2.0).

Construction of pFP02 for in-frame gene inactivation. For

inactivation of *cloR* in *S. roseochromogenes*, the fragments cloR-1 (1282 bp) and cloR-2 (1301 bp) were amplified by PCR. Primer pair cloR-1/*Hin*dIII (5′-GTCACCGG**AAGCTT**TGC-CTG-3′) and cloR-1/*Pst*I (5′-GCATGTT**CTGCAG**AGCC-TTG-3′) was used to amplify cloR-1; primer pair cloR-2/*Pst*I (5′-GCCTGCA**CTGCAG**GCCCCAA-3′) and cloR-2/*Bam*HI (5′-TCGTA**GGATCC**TCCCGTCGTC-3′) was used to amplify cloR-2. Restriction sites introduced into the primers are shown in bold in the aforementioned sequences. cloR-1 was digested with *Hin*dIII and *Pst*I and cloned into the corresponding sites of vector pBSKT, a pBluescript SK(+) derivative containing carbenicillin and thiostrepton resistances, resulting in pFP01. cloR-2 was digested with *Pst*I and *Bam*HI and ligated into the same sites of pFP01 to give pFP02.

Transformation of *S. roseochromogenes* and selection of recombinant mutants. Transformation of *S. roseochromogenes* with pFP02 was carried out by polyethylene glycol-mediated protoplast transformation (Kieser *et al.*, 2000). For the preparation of protoplasts, mycelia of *S. roseochromogenes* were grown in CRM medium containing 10·3 % sucrose, 2·0 % tryptic soy broth, 1·0 % MgCl$_2$.6H$_2$O, 1·0 % yeast extract and 0·75 % glycine (pH 7·0) for 48 h. The mycelia were then harvested and incubated in 5 ml P (protoplast) buffer (g mycelium)$^{-1}$ containing 1 mg lysozyme ml^{-1} for 30–60 min at 30 °C.

For transformation of *S. roseochromogenes*, pFP02 was mixed with 200 μl P-buffer containing 1×10^9 *S. roseochromogenes* protoplasts and 500 μl T (transformation) buffer containing 50 % (w/v) polyethylene glycol 1000 (Roth). The resulting suspension was plated onto R2YE agar. After incubation for 20 h at 30 °C, the plates were overlaid with 3 ml of soft R2YE agar containing a total of 500 μg thiostrepton, for selection of the recombinant mutants.

After the transformation of *S. roseochromogenes* protoplasts with pFP02, thiostrepton-resistant colonies were obtained.

Table 1. Identified ORFs in the biosynthetic gene cluster of clorobiocin

ORF	Size of the product (aa)	Similar entity/entities*	Identity with similar entity/entities (%)	Accession no.
ORF1	306	Fkbl (lysine cyclodeaminase) from *Streptomyces hygroscopicus*	37	AAF86391
ORF2	197	Sarcosine oxidase γ-subunit from *Corynebacterium* sp.	44	Q46338
ORF3	962	Sarcosine oxidase α-subunit from *Corynebacterium* sp.	59	Q46337
ORF4	93	Sarcosine oxidase δ-subunit from *Corynebacterium* sp.	66	Q46336
ORF5	406	Sarcosine oxidase β-subunit from *Corynebacterium* sp.	80	P40875
ORF6	406	Serine hydroxymethyltransferase from *S. coelicolor*	75	O86565
ORF7	218	Putative transcriptional regulator from *S. coelicolor*	45	AL596248.1
ORF8	149	Unknown protein		
ORF9	78	Transposase from *S. coelicolor*	68	AL109949.1
cloE	217	*novE* (217 aa)	**82**	
		lmbU protein from *S. lincolnensis*	43	S44974
cloF	362	*novF* (362 aa)	**73**	
		Prephenate dehydrogenase from *Streptomyces lavendulae*	30	AAK81837
cloG	319	*novG* (318 aa)/*couG* (319 aa)	**79/80**	
		Regulatory protein (StrR) from *Streptomyces glaucescens*	40	S44506
cloY	71	*couY* (71 aa)	**81**	
		Hypothetical protein from *S. coelicolor*	68	T36310
cloH	600	*novH* (600 aa)/*couH* (599 aa)	**75/80**	
		Peptide synthetase-like protein from *Streptomyces antibioticus*	50	AAG34184.1
cloI	407	*novI* (407 aa)/*couI* (407 aa)	**90/95**	
		Cytochrome P450 enzyme from *S. antibioticus*	66	AAG34185.1
cloJ	258	*novJ* (262 aa)/*couJ* (258 aa)	**72/77**	
		3-Oxoacyl reductase from *Bacillus halodurans*	48	BAB06210
cloK	245	*novK* (244 aa)/*couK* (245 aa)	**77/81**	
		Reductase from *S. antibioticus*	35	AAG34182
cloL	527	*novL* (527 aa)/*couL* (529 aa)	**86/86**	
		Amide synthetase from *S. antibioticus*	37	AAG34183
cloM	390	*novM* (379 aa)/*couM* (402 aa)	**78/78**	
		Glycosyltransferase from *Streptomyces argillaceus*	38	AAC64927
cloN1	95	*couN1* (95 aa)/hypothetical protein	**86**	
cloN2	355	*couN2* (355 aa)	**86**	
		dpsC from *Streptomyces peucetius*	32	1093565
cloN3	376	*couN3* (373 aa)	**81**	
		Acyl-CoA dehydrogenase (*pltE*) from *P. fluorescens*	45	AAD24879
cloN4	501	*couN4* (501 aa)	**87**	
		Acyl-CoA synthetase (*pltF*) from *P. fluorescens*	45	AAD24881
cloN5	89	*couN5* (89 aa)	**91**	
		Hypothetical protein (*pltL*) from *P. fluorescens*	34	T17418
cloN6	561	*couN6* (560 aa)	**94**	
		Hypothetical protein		
cloN7	278	*couN7* (281 aa)	**82**	
		Putative hydrolase from *S. coelicolor*	34	CAB95984
clo-hal	524	Non-haem halogenase from *S. lavendulae*	35	AAK81830
cloP	277	*novP* (262 aa)/*couP* (276 aa)	**91/89**	
		O-Methyltransferase III from *Streptomyces olivaceus*	58	CAC32469
cloQ	324	*novQ* (323 aa)/hypothetical protein	**84**	
cloR	277	*novR* (270 aa)	**95**	
		Putative aldolase *from S. coelicolor*	47	CAB82026
cloS	288	*novS* (288 aa)/*couS* (288 aa)	**84/87**	
		4-Ketoreductase from *S. antibioticus*	50	AAD55455
cloT	336	*novT* (336 aa)/*couT* (336 aa)	**82/87**	
		dNDP-glucose-4,6-dehydratase from *Streptomyces globisporus*	66	AAF13998

Table 1 (continued).

ORF	Size of the product (aa)	Similar entity/entities*	Identity with similar entity/entities (%)	Accession no.
cloU	420	*novU* (420 aa)/*couU* (420 aa)	**88/90**	
		D-Mycarose 3-C-methyltransferase from *S. argillaceus*	48	CAB96549
cloV	296	*novV* (297 aa)/*couV* (296 aa)	**89/92**	
		dTDP-glucose synthase from *Streptomyces spectabilis*	74	AAD31796
cloW	198	*novW* (207 aa)/*couW* (198 aa)	**86/91**	
		dTDP-4-keto-6-deoxyglucose 3,5-epimerase from *S. glaucescens*	54	S44236
cloZ	253	Hypothetical protein		
gyrB[R]	Partial sequence	***gyrB*[R]-*nov*** (novobiocin cluster)/***gyrB*[R]-*cou*** (coumermycin cluster)		

* *nov* genes are from the novobiocin biosynthetic gene cluster of *S. spheroides* (accession no. AF170880); *cou* genes are from the coumermycin A_1 biosynthetic gene cluster of *S. rishiriensis* (accession no. AF235050).

The single-cross-over mutant RSCO2 was grown in the absence of thiostrepton, allowed to sporulate and then examined for loss of resistance to thiostrepton due to double-cross-over events. Two mutants, named RDCO30 and RDCO32, were examined further. Chromosomal DNA from wild-type *S. roseochromogenes*, as well as from mutants RSCO2, RDCO30 and RDCO32, was digested with *Sac*II and hybridized with a probe containing part of *cloR*. A band of approximately 1·1 kb in size was detected upon hybridization of the wild-type *S. roseochromogenes* DNA with the probe, whereas hybridization of the chromosomal DNA from mutant strain RDCO30 with the probe produced the expected 2·2 kb band, which corresponded to the in-frame deletion of *cloR* (Fig. 2).

Complementation of the *cloR*-defective mutant with Ring A. 3-Dimethylallyl-4-hydroxybenzoic acid (=Ring A of clorobiocin and novobiocin) was obtained by hydrolysis of novobiocin, as described by Kominek & Sebek (1974). The *cloR*-defective mutant strain RDCO30 was grown in 50 ml pre-culture medium (see above) supplemented with 1 mg Ring A. After 48 h incubation in the pre-culture medium, 5 ml of the culture were inoculated into 50 ml production medium (see above) supplemented with 1 mg Ring A. After 7 days cultivation of the strain in the production medium, secondary metabolites were analysed.

Analysis of secondary metabolites. The bacterial culture (20 ml) was acidified to pH 4 with HCl and extracted twice with an equal volume of ethyl acetate. After centrifugation, the solvent was evaporated and the dried extract was resuspended in 1 ml methanol. Metabolites were analysed by HPLC using a Multosphere RP18-5 column (250 × 4 mm, 5 μm) with a linear gradient from 60 to 100 % methanol in 1 % aqueous formic acid and detection at 340 nm. Authentic clorobiocin (Aventis) was used as the standard.

For preparative isolation of the secondary metabolites, the fractions from HPLC analysis were collected and the solvent was evaporated. The product was analysed by MS and ¹H-NMR. Negative fast-atom bombardment (FAB) mass spectra were recorded on a TSQ70 spectrometer (Finnigan) using methanol as the solvent; peaks A and B gave identical isotopic peaks, characteristic for substances with one chlorine atom, at m/z 697 and 695 [M-H]⁻ that corresponded to the clorobiocin standard. The ¹H-NMR spectrum was measured on an AMX 400 spectrometer (Bruker), and peak A gave signals corresponding to those of the clorobiocin standard: $\delta_{\mathrm{p.p.m.}}$ (CD$_3$OD, 400 MHz), 7·90 (d, $J=9\cdot2$ Hz, H-5), 7·76 (d, $J=2\cdot5$ Hz, H-2′), 7·72 (dd, $J=8\cdot4$ Hz, 2·5 Hz, H-6′), 7·33 (d, $J=9\cdot2$ Hz, H-6), 6·90 (d, $J=3\cdot6$ Hz, H-3′′′), 6·84 (d, $J=8\cdot4$ Hz, H-5′), 5·94 (d, $J=3\cdot6$ Hz, H-4′′′), 5·73 (d, $J=1\cdot8$ Hz, H-1′′), 5·71 (dd, $J=10\cdot3$ Hz, $J=2\cdot9$ Hz, H-3′′), 5·35 (broad t, $J=7\cdot1$ Hz, H-8′), 4·34 (t, $J=2\cdot7$ Hz, H-2′′), 3·72 (d, $J=10\cdot3$ Hz, H-4′′), 3·52 (s, 3H-8′′), 3·34 (d, $J=7\cdot1$ Hz, 2H-7′), 2·29 (s, 3H-6′′′), 1·75 (s, 3H-11′), 1·74 (s, 3H-10′), 1·35 (s, 3H-7′′), 1·18 (s, 3H-6′′).

RESULTS

Cloning and sequencing of the clorobiocin biosynthetic gene cluster

We have previously cloned the novobiocin (Steffensky *et al.*, 2000a) and coumermycin A_1 (Wang *et al.*, 2000) biosynthetic gene clusters (Fig. 1). In the novobiocin cluster, *novT* encodes a dTDP-glucose 4,6-dehydratase involved in the biosynthesis of the deoxysugar moiety of novobiocin; *novL* encodes the novobiocic acid synthetase that catalyses the formation of the amide bond between Ring A and Ring B of novobiocin (Fig. 1) (Steffensky *et al.*, 2000b). Similar reactions to these were expected to be involved in clorobiocin biosynthesis. Therefore, Southern hybridizations of genomic DNA from the clorobiocin producer *S. roseochromogenes* were carried out with probes for *novT* and *novL*; each hybridization resulted in a single band.

A cosmid library from *S. roseochromogenes* was constructed in SuperCos-1 and screened with the *novT* and *novL* probes. The hybridizing cosmids were analysed by restriction mapping. Four different but overlapping cosmids were obtained which covered a continuous 55 kb region of the chromosome. Cosmid K1F2 was sequenced on both strands. From the sequence of this cosmid, 36 ORFs were identified; 27 of these showed striking similarity to genes of the novobiocin and/or coumermycin A_1 biosynthetic gene clusters. In addition, a partial sequence of the aminocoumarin-resistance gene (*gyrB*[R]) was found at the 3′ end of the clorobiocin gene cluster. Strikingly, in the gene clusters of coumermycin A_1, novobiocin and clorobiocin, the corresponding

Fig. 3. Hypothetical biosynthetic pathway for clorobiocin.
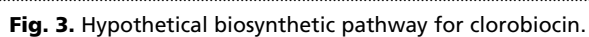
ORFs were arranged in exactly the same order and oriented in the same direction (Fig. 1).

Table 1 lists the homologies found between the genes in cosmid K1F2 and the genes of the novobiocin and coumermycin A_1 clusters, as well as the homologies of the K1F2 genes to other sequences found in GenBank. The sequence of cosmid K1F2 has been deposited in GenBank under accession no. AF329398.

Genes involved in the biosynthesis of the deoxysugar moiety (Ring C)

At the 3′ end of the clorobiocin gene cluster, five ORFs with high homology to genes involved in the deoxysugar biosynthesis were discovered (*cloSTUVW*) (Table 1). Homologous genes are found in the same position in the gene clusters of novobiocin (*novSTUVW*) and coumermycin A_1 (*couSTUVW*) (Fig. 1). Based upon their homology to known genes involved in deoxysugar biosynthesis, we previously assigned these genes to the five steps required for the biosynthesis of the deoxysugar moiety (Fig. 3); functional proof for this hypothesis was provided by an inactivation experiment with *novT* in the novobiocin producer *S. spheroides* (Steffensky *et al.*, 2000a). The presence of these genes in the clorobiocin gene cluster provides additional support to our previous functional assignment of these genes.

O-Methylation at position 4 of the deoxysugar moiety is regarded as the last step in aminocoumarin biosynthesis (Fig. 3) (Queener *et al.*, 1978). The genes *cloP*, *novP* and *couP*, situated in the same relative position of their respective gene clusters, show homology to genes encoding *O*-methyltransferases and may be assigned to this reaction. CouP has recently been proven to be responsible for the methylation of 4-OH of the deoxysugar (Li *et al.*, 2002).

Genes presumably involved in the formation of the pyrrole ring

Clorobiocin and coumermycin A_1 contain pyrrole carboxylic acid rings attached to position 3 of their deoxysugar moieties (Fig. 1). Novobiocin contains a carbamyl group at the corresponding position.

These structural similarities and differences between the three antibiotics are reflected in the organization of their respective gene clusters (Fig. 1). Downstream of the glycosyltransferase gene *novM*, the novobiocin cluster contains a gene (*novN*) with homology to those encoding carbamoyl transferases, whereas in the same relative position of the clorobiocin and coumermycin clusters, a group of seven genes is found (*cloN1–N7* or *couN1–N7*, respectively) which show very high homology between the two clusters (on average 87 % identity). These genes can be assigned to pyrrole biosynthesis, which has been elucidated on the genetic and biochemical level: *cloN3*, *cloN4* and *cloN5* show sequence similarities to *pltE*, *pltF* and *pltL*, respectively,

which are involved in the biosynthesis of the pyrrole moiety of pyoluteorin in *Pseudomonas fluorescens* Pf-5 (Nowak-Thompson *et al.*, 1999) and to *redW*, *redM* and *redO*, respectively, which are involved in the biosynthesis of the pyrrole moiety of undecylprodiginin in *Streptomyces coelicolor* (Cerdeno *et al.*, 2001). PltF and RedM convert L-proline into its acyl adenylate (Fig. 3) and the small proteins PltL and RedO act as peptidyl carrier proteins (PCPs) (Thomas *et al.*, 2002). Therefore, the same functions may be assigned to the homologous proteins CloN4 and CloN5, respectively. CloN3, like CouN3, PltE and RedW, shows homology to flavine-dependent acyl-coenzyme A dehydrogenases. PltE and RedW catalyse the dehydrogenation of the PCP-bound proline (Thomas *et al.*, 2002). The resulting pyrroline derivative (presumably Δ^2) undergoes spontaneous oxidation to the aromatic pyrrole derivative (Fig. 3).

CloN6 (CouN6) belongs to the BchE-like/methyltransferase subgroup of radical SAM proteins, which has recently been identified using bioinformatic techniques (Sofia *et al.*, 2001). We suggest that CloN6 catalyses the transfer of a methyl group to position 5 of the pyrrole-2-carboxylic acid. *cloN2* (*couN2*) shares homology with *dpsC*, which encodes an enzyme with acyltransferase activity. *cloN7* and/or *cloN2* may be involved in the transfer of the activated pyrrole-2-carboxylic acid to the 3-OH of the deoxysugar moiety. The small ORF *cloN1* (encoding 95 aa) does not show homology to other database entries, and its function remains unknown at present.

Genes presumably involved in the biosynthesis of the aminocoumarin ring (Ring B)

The genes for the biosynthesis of the characteristic aminocoumarin ring must be present in all three gene clusters; hence, a comparison of the three gene clusters presents an obvious method for identifying possible candidate genes involved in the biosynthesis of this ring. In the clorobiocin cluster, *cloHIJK* showed, on average, 85 % homology to the corresponding genes in the novobiocin (*novHIJK*) and coumermycin A_1 (*couHIJK*) clusters (Table 1). It appears likely that the products of *these genes* are involved in the formation of the aminocoumarin ring from tyrosine (see Discussion).

Clorobiocin contains a chlorine atom at position 8 of the aminocoumarin ring, whereas novobiocin and coumermycin A_1 contain a methyl group at the same position (Fig. 1). This structural difference of the antibiotics is perfectly reflected in the organization of their gene clusters: the novobiocin and coumermycin clusters contain a C-methyltransferase gene, *novO* or *couO*, respectively, whereas in the clorobiocin cluster, *clo–hal*, a homologue of non-haem halogenase genes, is found at the same relative position.

Genes presumably involved in the biosynthesis of the 3-dimethylallyl-4-hydroxybenzoic acid (Ring A)

Clorobiocin and novobiocin contain a prenylated 4-hydroxybenzoate moiety (Ring A). Coumermycin A_1
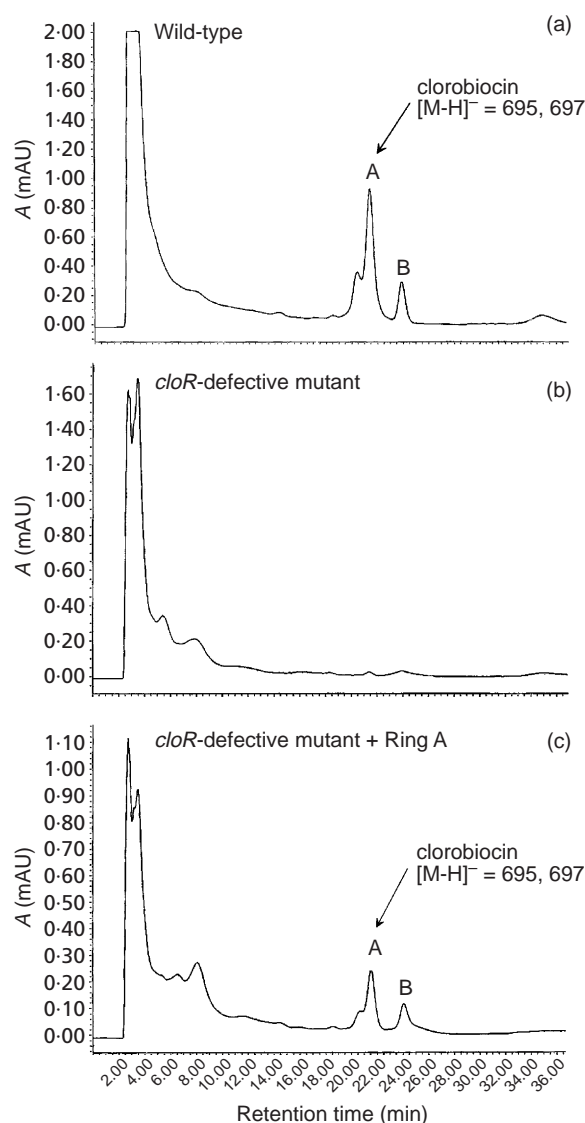
Fig. 4. HPLC analysis of the secondary metabolites produced by (a) wild-type *S. roseochromogenes* DS 12.976, (b) the *cloR*-defective mutant and (c) the *cloR*-defective mutant complemented with Ring A (3-dimethylallyl-4-hydroxybenzoic acid). Absorbance values are in milli arbitrary units (mAU).

contains a pyrrole dicarboxylic acid moiety instead, linking the two aminocoumarin rings of this molecule (Fig. 1). The aromatic nucleus of Ring A of clorobiocin and novobiocin is derived from tyrosine (Bunton *et al.*, 1963; Kominek & Sebek, 1974), but the exact reaction sequence is unknown.

Sequencing of the clorobiocin gene cluster revealed two genes that are also present in the novobiocin cluster, but which are not present in the coumermycin cluster, i.e. *cloQ* and *cloR*. This fact drew us to the hypothesis that these genes may be involved in Ring A biosynthesis.

cloQ and *cloR*, like *novQ* and *novR*, show transcriptional coupling (i.e. the stop codon of *cloQ* is fused with the start codon of *cloR*) and are likely to be transcribed

as a single operon. Unusually large intergenic regions are found upstream and downstream of *cloQR* (1001 and 830 bp, respectively). CloR has 47% identity to a putative aldolase from *S. coelicolor*. *cloQ* did not show homologies to other genes in the database, with the exception of *novQ*.

Generation of a *cloR*-defective mutant

To test whether *cloR* was indeed involved in Ring A biosynthesis, a gene inactivation experiment was carried out. An inactivation vector carrying a thiostrepton-resistance gene (pFP02) was constructed, in which the structural gene *cloR* was disrupted by an in-frame deletion (Fig. 2). The deletion mutant, RDCO30, was subsequently cultured and the ethyl acetate extract of the culture was examined by HPLC for secondary metabolites. As shown in Fig. 4, the production of clorobiocin was abolished in this mutant. Another thiostrepton-sensitive strain obtained in the screening for double-cross-over mutants, RDCO32, represented a reversion to the wild-type (Fig. 2), and showed clorobiocin production identical to that of the wild-type strain (data not shown).

Feeding of Ring A to the *cloR*-defective mutant

To restore clorobiocin biosynthesis in the *cloR*-defective mutant strain RDCO30, Ring A (3-dimethylallyl-4-hydroxybenzoic acid) was added to the culture medium of this strain. This led to the reconstitution of clorobiocin production (Fig. 4), to one-third the level of the wild-type strain. The identity of this product was confirmed by MS and ^{1}H-NMR, in a comparison with authentic clorobiocin.

Besides the major peak for clorobiocin (peak A in Fig. 4), a minor peak of identical mass (peak B) was detected both in the wild-type and the complemented mutant strain. This substance is likely to represent an isomer of clorobiocin, possibly carrying the pyrrole carboxylic acid moiety in position 2 instead of position 3 of the deoxysugar. Such isomers have been reported previously for novobiocin (Hinman *et al.*, 1957).

Genes involved in the linkage of Rings A, B and C of clorobiocin

Attachment of the deoxysugar to the 7-OH group of the aminocoumarin ring requires very similar glycosyl transferases in clorobiocin, novobiocin and coumermycin A_1 biosynthesis; indeed, three very similar putative glycosyltransferase genes, *cloM*, *novM* and *couM*, are found at the same relative position in all three gene clusters.

In clorobiocin and novobiocin, the aminocoumarin moiety (Ring B) and the prenylated 4-hydroxybenzoate moiety (Ring A) are linked by an amide bond (Fig. 1). It has been demonstrated that the enzyme NovL catalyses

this reaction, i.e. the adenylation of the substituted benzoyl moiety and its transfer to the amino group (Steffensky *et al*., 2000b). *cloL* shows high homology to *novL* and is most probably involved in the formation of the amide bond of clorobiocin.

Resistance and regulatory genes

Downstream of the deoxysugar biosynthesis genes *cloSTUVW* in the clorobiocin cluster, and similarly at the corresponding position of the novobiocin and coumermycin clusters, a gene encoding an amino-coumarin-resistant gyrase B subunit ($gyrB^R$) is located. This gene has previously been identified as the principal novobiocin-resistance gene in the novobiocin producer *S. spheroides* (Thiara & Cundliffe, 1988).

cloG, *novG* and *couG* are homologous to *strR*, a regulatory gene from the streptomycin cluster. Streptomycin biosynthesis is known to be regulated by γ-butyrolactones (Horinouchi & Beppu, 1995). It may, therefore, be speculated that γ-butyrolactones are involved in the regulation of the biosynthesis of clorobiocin and other aminocoumarin antibiotics.

cloE has homology to the *lmbU* gene of the lincomycin biosynthetic gene cluster of *Streptomyces lincolnensis* 78-11. It was suggested that LmbU may have a regulatory function, but no experimental evidence is available so far to support this (Peschke *et al*., 1995).

Genes with unknown function

At present, no function can be suggested for the small ORFs *cloY* and *cloN1*, which have homologues in the coumermycin A$_1$ cluster, and for *cloZ*, which has no homologues in other clusters.

DISCUSSION

In this study, we have cloned and sequenced the clorobiocin biosynthetic gene cluster which spans approximately 31 kb and comprises 27 ORFs. The suggested function of these ORFs in clorobiocin biosynthesis is depicted in Fig. 3.

The primary metabolic genes found at the 5′ end of the clorobiocin cluster (Table 1) suggest that *cloE* represents the 5′ border of the cluster. The gene adjacent to *cloE*, i.e. *ORF9*, encodes a putative transposase, and it may be speculated that this gene is related to the introduction of the cluster into the *S. roseochromogenes* genome. Also at the 3′ end, downstream of $gyrB^R$, primary metabolic genes were found, suggesting that the sequence depicted in Fig. 1 comprises all biosynthetic genes of the clorobiocin cluster. However, it cannot be excluded that additional biosynthetic enzymes for clorobiocin formation are encoded at different loci of the genome.

The clorobiocin cluster contains the genes *cloHIJK*, for which homologues exist in the novobiocin cluster (*novHIJK*) and the coumermycin cluster (*couHIJK*). NovH has recently been shown to activate tyrosine by covalent binding to the 4-phosphopantetheinyl cofactor, and the P450 enzyme NovI catalyses the β-hydroxylation of the activated tyrosine (Chen & Walsh, 2001). A central, unresolved question in aminocoumarin biosynthesis is how activated β-hydroxy tyrosine is then converted to the coumarin ring, especially how the ring oxygen is introduced.

cloJ and *cloK*, the genes immediately downstream of *cloH* and *cloI*, are homologous to *novJK* and *couJK* of the novobiocin and coumermycin clusters, respectively. The detection of *cloJ* and *cloK* in the clorobiocin cluster, and the homologous *simJ1* and *simK* genes in the biosynthetic gene cluster of the aminocoumarin antibiotic simocyclinone (Galm *et al*., 2002; Trefzer *et al*., 2002), now strongly supports the hypothesis that these genes are indeed related to aminocoumarin biosynthesis. *cloJ* shows homology to 3-oxo-acyl-[ACP] reductases and may likely be involved in the oxidation of a β-hydroxy-tyrosyl to a β-keto-tyrosyl intermediate (Fig. 3). Also, *cloK* shows homology to oxidoreductases, but this homology is not very high (mean of 35% on the amino acid level). Chen & Walsh (2001) speculated that NovJ and NovK may act together to oxidize the β-hydroxyl function to a keto group. The unresolved step in the postulated Ring B biosynthesis, however, is the hydroxylation of the activated tyrosyl derivative in position 2 of the aromatic nucleus (Fig. 3). Bunton *et al*. (1963) had reported that the ring oxygen of the aminocoumarin may be derived from the carboxyl group of tyrosine rather than from molecular oxygen. This was recently disproven by Holzenkämpfer & Zeeck (2002), who showed that the ring oxygen of the aminocoumarin moiety of simocyclinone is in fact derived from molecular oxygen. Therefore, coumarin ring formation most likely proceeds via the 2-hydroxylation of a tyrosine derivative. Chen & Walsh (2001) had speculated that the predicted flavine dioxygenase NovC, encoded by a gene near the novobiocin cluster, may catalyse this reaction. An important finding of our study is that no *novC* homologue was detected in or near the clorobiocin gene cluster. Likewise, no *novC* homologue was detected in the simocyclinone cluster (Galm *et al*., 2002; Trefzer *et al*., 2002). We therefore suggest that *novC* is not related to aminocoumarin biosynthesis. The enzyme responsible for the 2-hydroxylation of the β-keto-tyrosyl intermediate remains unknown at present. Whether *cloK* is involved in this or another reaction of aminocoumarin biosynthesis has yet to be demonstrated.

Halogenation of the aminocoumarin ring may occur after ring formation, as depicted in Fig. 3, or at an earlier stage.

The prenylated 4-hydroxybenzoate moiety (Ring A) of clorobiocin and novobiocin is formed from tyrosine (Kominek & Sebek, 1974) and an isoprenoid precursor, but this reaction sequence is also unknown. The biosynthesis of this moiety requires: (i) the assembly of the isoprenoid precursor (probably dimethylallyl diphosphate) via the methylerythritol phosphate pathway

(Li *et al*., 1998); (ii) the conversion of the phenylpropanoid compound tyrosine to a benzoic acid derivative; and (iii) the prenylation of the aromatic nucleus in a prenyltransferase reaction. The conversion of the phenylpropanoid intermediate to a benzoic acid derivative may proceed by a mechanism analogous to the oxidation of fatty acids, as demonstrated in *Streptomyces maritimus* (Hertweck & Moore, 2000). Alternatively, this conversion may occur by retro-aldol cleavage of a 3-hydroxylated phenylpropanoid compound, as found in *P. fluorescens* (Gasson *et al*., 1998) and *Amycolatopsis* sp. (Achterholt *et al*., 2000). Retro-aldol cleavage would result in a benzaldehyde derivate, which would subsequently be oxidized to the benzoic acid derivative.

In the clorobiocin cluster, we could not detect genes similar to those for the β-oxidation of fatty acids. We did find, however, the gene *cloR* which showed homology to a putative aldolase from *S. coelicolor*. Comparison of the gene clusters of clorobiocin, novobiocin and coumermycin A$_1$ led to the hypothesis that this gene may be involved in Ring A biosynthesis. Inactivation of the enzyme led to an abolishment of clorobiocin production. When the mutant was complemented with the prenylated 4-hydroxybenzoate moiety (=Ring A), clorobiocin production was restored. This proved that *cloR* is indeed involved in Ring A biosynthesis. We suggest that the formation of Ring A of clorobiocin may proceed via a retro-aldol reaction catalysed by CloR, i.e. by a mechanism different from the elucidated benzoic acid biosynthesis in *S. maritimus* (Hertweck & Moore, 2000).

The substrate of CloR may be an enzyme-bound prenylated β-hydroxytyrosine, as suggested by Chen & Walsh (2001) and depicted in Fig. 3. Alternatively, prenylated 4-hydroxyphenylpyruvate may be the substrate of CloR, since this compound was detected as the product of a prenyltransferase of the novobiocin producer *S. spheroides* (Steffensky *et al*., 1998). Further studies are now in progress to investigate these steps in clorobiocin biosynthesis.

The cloning and sequencing of the clorobiocin gene cluster has completed the genetic information on the biosynthesis of three 'classical' aminocoumarin antibiotics, namely novobiocin, clorobiocin and coumermycin A$_1$. Comparison of the three gene clusters revealed a strikingly stringent correspondence between the structures of the antibiotics and the organization of the biosynthetic genes, unprecedented so far in any class of natural products outside the polyketide and the peptide antibiotics. For each structural moiety of the aminocoumarin antibiotics, the biosynthetic genes are grouped together, resulting in a 'modular' structure of the clusters. The orders of the modules and the order of the genes within each module are perfectly identical for the three 'classical' aminocoumarins, and nearly all of the genes within the clusters are orientated in the same direction. The comparison of the three clusters greatly facilitates the prediction of functions for the different genes. As an example, *cloR* was recognized by such a comparison as a candidate gene for the biosynthesis of

the prenylated 4-hydroxybenzoate moiety of clorobiocin, and this was experimentally proven by an inactivation and complementation experiment. The similarity between the three clusters also provides excellent opportunities for the production of hybrid aminocoumarins by genetic methods.

ACKNOWLEDGEMENTS

REFERENCES

Achterholt, S., Priefert, H. & Steinbuchel, A. (2000). Identification of *Amycolatopsis* sp. strain HR167 genes, involved in the bioconversion of ferulic acid to vanillin. *Appl Microbiol Biotechnol* **54**, 799–807.

Ali, J. A., Jackson, A. P., Howells, A. J. & Maxwell, A. (1993). The 43-kilodalton N-terminal fragment of the DNA gyrase B protein hydrolyzes ATP and binds coumarin drugs. *Biochemistry* **32**, 2717–2724.

Berger, J. & Batcho, A. D. (1978). Coumarin-glycoside antibiotics. *J Chromatogr Libr* **15**, 101–159.

Bunton, C. A., Kenner, G. W., Robinson, M. J. T. & Webster, B. R. (1963). Experiments related to the biosynthesis of novobiocin and other coumarins. *Tetrahedron* **19**, 1001–1010.

Cerdeno, A. M., Bibb, M. J. & Challis, G. L. (2001). Analysis of the prodiginine biosynthesis gene cluster of *Streptomyces coelicolor* A3(2): new mechanisms for chain initiation and termination in modular multienzymes. *Chem Biol* **8**, 817–829.

Chen, H. & Walsh, C. T. (2001). Coumarin formation in novobiocin biosynthesis: β-hydroxylation of the aminoacyl enzyme tyrosyl-S-NovH by a cytochrome P450 NovI. *Chem Biol* **8**, 301–312.

Galm, U., Schimana, J., Fiedler, H. P., Schmidt, J., Li, S.-M. & Heide, L. (2002). Cloning and analysis of the simocyclinone biosynthetic gene cluster of *Streptomyces antibioticus* Tü 6040. *Arch Microbiol* **178**, 102–114.

Gasson, M. J., Kitamura, Y., McLauchlan, W. R., Narbad, A., Parr, A. J., Parsons, E. L. H., Payne, J., Rhodes, M. J. C. & Walton, N. J. (1998). Metabolism of ferulic acid to vanillin. A bacterial gene of the enoyl-SCoA hydratase/isomerase superfamily encodes an enzyme for the hydration and cleavage of a hydroxycinnamic acid SCoA thioester. *J Biol Chem* **273**, 4163–4170.

Hertweck, C. & Moore, B. S. (2000). A plant-like biosynthesis of benzoyl-CoA in the marine bacterium *Streptomyces maritimus*. *Tetrahedron* **56**, 9115–9120.

Hinman, J. W., Caron, E. L & Hoeksema, H. (1957). Novobiocin. V. Carbamyl migration and isonovobiocin. *J Am Chem Soc* **79**, 5321–5322.

Holzenkämpfer, M. & Zeeck, A. (2002). Biosynthesis of simocyclinone D8 in an $^{18}O_2$-rich atmosphere. *J Antibiot (Tokyo)* **55**, 341–342.

Hooper, D. C., Wolfson, J. S., McHugh, G. L., Winters, M. B. & Swartz, M. N. (1982). Effects of novobiocin, coumermycin A1, clorobiocin, and their analogs on *Escherichia coli* DNA gyrase and bacterial growth. *Antimicrob Agents Chemother* **22**, 662–671.

Horinouchi, S. & Beppu, T. (1995). Autoregulators. *Biotechnology* **28**, 103–119.

Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. (2000). *Practical Streptomyces Genetics*. Norwich: John Innes Foundation.

Kominek, L. A. & Sebek, O. K. (1974). Biosynthesis of novobiocin and related coumarin antibiotics. *Dev Ind Microbiol* **15**, 60–69.

Lafitte, D., Lamour, V., Tsvetkov, P. O., Makarov, A. A., Klich, M., Deprez, P., Moras, D., Briand, C. & Gilli, R. (2002). DNA gyrase interaction with coumarin-based inhibitors: the role of the hydroxybenzoate isopentenyl moiety and the 5′-methyl group of the noviose. *Biochemistry* **41**, 7217–7223.

Lewis, R. J., Singh, O. M., Smith, C. V., Skarzynski, T., Maxwell, A., Wonacott, A. J. & Wigley, D. B. (1996). The nature of inhibition of DNA gyrase by the coumarins and the cyclothialidines revealed by X-ray crystallography. *EMBO J* **15**, 1412–1420.

Li, S.-M., Hennig, S. & Heide, L. (1998). Biosynthesis of the dimethylallyl moiety of novobiocin via a non-mevalonate pathway. *Tetrahedron Lett* **39**, 2717–2720.

Li, S.-M., Westrich, L., Schmidt, J., Kuhnt, C. & Heide, L. (2002). Methyltransferase genes in *Streptomyces rishiriensis*: new coumermycin derivatives from gene-inactivation experiments. *Microbiology* **148**, 3317–3326.

Lombo, F., Siems, K., Brana, A. F., Mendez, C., Bindseil, K. & Salas, J. A. (1997). Cloning and insertional inactivation of *Streptomyces argillaceus* genes involved in the earliest steps of biosynthesis of the sugar moieties of the antitumor polyketide mithramycin. *J Bacteriol* **179**, 3354–3357.

Mancy, D., Ninet, L. & Preud Homme, J. (1974). Antibiotic 18. 631 RP. United States Patent Office, no. 3793147.

Maxwell, A. (1993). The interaction between coumarin drugs and DNA gyrase. *Mol Microbiol* **9**, 681–686.

Maxwell, A. (1999). DNA gyrase as a drug target. *Biochem Soc Trans* **27**, 48–53.

Nowak-Thompson, B., Chaney, N., Wing, J. S., Gould, S. J. & Loper, J. E. (1999). Characterization of the pyoluteorin biosynthetic gene cluster of *Pseudomonas fluorescens* Pf-5. *J Bacteriol* **181**, 2166–2174.

Peschke, U., Schmidt, H., Zhang, H. Z. & Piepersberg, W. (1995). Molecular characterization of the lincomycin-production gene cluster of *Streptomyces lincolnensis* 78-11. *Mol Microbiol* **16**, 1137–1156.

Queener, S. W., Sebek, O. K. & Vezina, C. (1978). Mutants blocked in antibiotic synthesis. *Annu Rev Microbiol* **32**, 593–636.

Sambrook, J. & Russell, D. (2001). *Molecular Cloning: a Laboratory Manual*, 3rd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sofia, H. J., Chen, G., Hetzler, B. G., Reyes-Spindola, J. F. & Miller, N. E. (2001). Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Res* **29**, 1097–1106.

Steffensky, M., Li, S.-M., Vogler, B. & Heide, L. (1998). Novobiocin biosynthesis in *Streptomyces spheroides*: identification of a dimethylallyl diphosphate:4-hydroxyphenylpyruvate dimethylallyl transferase. *FEMS Microbiol Lett* **169**, 69–74.

Steffensky, M., Mühlenweg, A., Wang, Z. X., Li, S.-M. & Heide, L. (2000a). Identification of the novobiocin biosynthetic gene cluster of *Streptomyces spheroides* NCIB 11891. *Antimicrob Agents Chemother* **44**, 1214–1222.

Steffensky, M., Li, S.-M. & Heide, L. (2000b). Cloning, over-expression, and purification of novobiocic acid synthetase from *Streptomyces spheroides* NCIMB 11891. *J Biol Chem* **275**, 21754–21760.

Thiara, A. S. & Cundliffe, E. (1988). Cloning and characterization of a DNA gyrase B gene from *Streptomyces sphaeroides* that confers resistance to novobiocin. *EMBO J* **7**, 2255–2259.

Thomas, M. G., Burkart, M. D. & Walsh, C. T. (2002). Conversion of L-proline to pyrrolyl-2-carboxyl-S-PCP during undecylprodigiosin and pyoluteorin biosynthesis. *Chem Biol* **9**, 171–184.

Trefzer, A., Pelzer, S., Schimana, J., Stockert, S., Bihlmaier, C., Fiedler, H. P., Welzel, K., Vente, A. & Bechthold, A. (2002). Biosynthetic gene cluster of simocyclinone, a natural multihybrid antibiotic. *Antimicrob Agents Chemother* **46**, 1174–1182.

Tsai, F. T., Singh, O. M., Skarzynski, T. & 9 other authors (1997). The high-resolution crystal structure of a 24-kDa gyrase B fragment from *E. coli* complexed with one of the most potent coumarin inhibitors, clorobiocin. *Proteins* **28**, 41–52.

Wang, Z. X., Li, S.-M. & Heide, L. (2000). Identification of the coumermycin A(1) biosynthetic gene cluster of *Streptomyces rishiriensis* DSM 40489. *Antimicrob Agents Chemother* **44**, 3040–3048.