Microbiology (2003), **149**, 1633–1645 DOI 10.1099/mic.0.26310-0

Cloning, sequencing and heterologous expression of the medermycin biosynthetic gene cluster of *Streptomyces* sp. AM-7161: towards comparative analysis of the benzoisochromanequinone gene clusters

Koji Ichinose, Makoto Ozawa, Keiko Itou, Kanako Kunieda and Yutaka Ebizuka

Graduate School of Pharmaceutical Sciences, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Correspondence
Koji Ichinose
ichinose@mol.f.u-tokyo.ac.jp

Received 19 February 2003
Revised 24 March 2003
Accepted 26 March 2003

Medermycin is a *Streptomyces* aromatic *C*-glycoside antibiotic classified in the benzoisochromanequinones (BIQs), which presents several interesting biosynthetic problems concerning polyketide synthase (PKS), post-PKS tailoring and deoxysugar pathways. The biosynthetic gene cluster for medermycin (the *med* cluster) was cloned from *Streptomyces* sp. AM-7161. Completeness of the clone was proved by the heterologous expression of a cosmid carrying the entire *med* cluster in *Streptomyces coelicolor* CH999 to produce medermycin. The DNA sequence of the cosmid (36 202 bp) revealed 34 complete ORFs, with an incomplete ORF at either end. Functional assignment of the deduced products was made for PKS and biosynthetically related enzymes, tailoring steps including strereochemical control, oxidation, angolosamine pathway, *C*-glycosylation, and regulation. The *med* cluster was estimated to be about 30 kb long, covering 29 ORFs. An unusual characteristic of the cluster is the disconnected organization of the minimal PKS genes: *med*-ORF23 encoding the acyl carrier protein is 20 kb apart from *med*-ORF1 and *med*-ORF2 for the two ketosynthase components. Secondly, the six genes (*med*-ORF14, 15, 16, 17, 18 and 20) for the biosynthesis of the deoxysugar, angolosamine, are all contiguous. Finally, the finding of a glycosyltransferase gene, *med*-ORF8, suggests a possible involvement of conventional *C*-glycosylation in medermycin biosynthesis. Comparison among the three complete BIQ gene clusters – *med* and those for actinorhodin (*act*) and granaticin (*gra*) – revealed some common genes whose deduced functions are unavailable from database searches (the 'unknowns'). An example is *med*-ORF5, a homologue of *actVI*-ORF3 and *gra*-ORF18, which was highlighted by a recent proteomic analysis of *S. coelicolor* A3(2).

INTRODUCTION

Medermycin is an antibiotic active against Gram-positive bacteria, including staphylococci, isolated from *Streptomyces* sp. K73 (Takano *et al.*, 1976). Its structure was shown (Tatsuta *et al.*, 1990) to be identical with lactoquinomycin A, independently characterized as an anticancer compound from *Streptomyces tanashiensis* (Tanaka *et al.*, 1985; Okabe *et al.*, 1985). Medermycin featured in the first production of hybrid antibiotics by genetic engineering (Hopwood *et al.*, 1985), where a medermycin producer, *Streptomyces* sp. AM-7161, was transformed to produce mederhodins A and B (Ōmura *et al.*, 1986). This was achieved by introduction of a segment of the biosynthetic gene cluster of a structurally related antibiotic, actinorhodin, from *Streptomyces coelicolor* A3(2).

Medermycin and actinorhodin are among a class of *Streptomyces* aromatic antibiotics known as benzoisochromanequinones (BIQs; Ichinose *et al.*, 1998a). BIQs present a number of interesting biosynthetic problems concerning the polyketide synthase (PKS) (Hopwood, 1997), post-PKS

This paper is dedicated to Professor Sir David Hopwood, a pioneer in the field of *Streptomyces* research, on the occasion of his 70th birthday.

Abbreviations: ACP, acyl carrier protein; ARO, aromatase; BIQ, benzoisochromanequinone; CYC, cyclase; DOH, deoxyhexose; GT, glycosyltransferase; KR, ketoreductase; KS, ketosynthase; PKS, polyketide synthase.

The GenBank accession number for the sequence reported in this paper is AB103463.

^{1}H and ^{13}C-NMR spectra of medermycin isolated from *S. coelicolor* CH999/pIK340 are available as supplementary data with the online version of this paper at http://mic.sgmjournals.org.

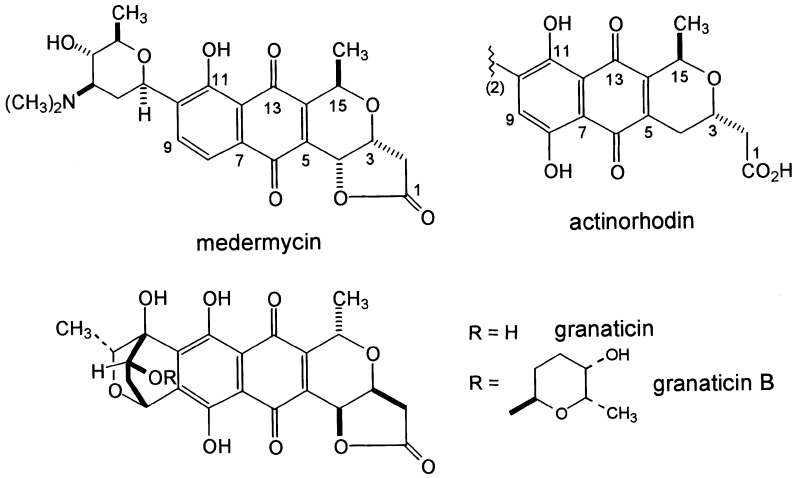


Fig. 1. Structures of medermycin, actinorhodin and granaticins.

modification (Rix *et al.*, 2002) ('tailoring'), and for the glycosylated BIQs, such as medermycin and granaticin, deoxysugar biosynthesis (Trefzer *et al.*, 1999). Of particular interest are the common chemical modifications at C-10 via a C–C bond either by glycosylation (medermcyin and granaticin) or by dimerization (in actinorhodin) (Fig. 1). The biosynthetic gene clusters for the BIQs that have been cloned and sequenced to date are for actinorhodin (the *act* cluster: Hallam *et al.*, 1988; Fernández-Moreno *et al.*, 1991, 1992, 1994; Caballero *et al.*, 1991), frenolicin B (the *fren* [*frn*] cluster: Bibb *et al.*, 1994; GenBank accession no. AF058302), griseusin B (the *gris* cluster: Yu *et al.*, 1994), and granaticin (the *gra* cluster: Sherman *et al.*, 1989; Bechthold *et al.*, 1995; Ichinose *et al.*, 1998b). Sequence analysis of ORFs in these clusters revealed rather diverse genetic organizations, especially for genes encoding tailoring enzymes (Ichinose *et al.*, 1998b). Some common 'unknown' genes are also present, whose deduced functions are unavailable from database searches, but are apparently important for BIQ biosynthesis (Cole *et al.*, 1987; Taguchi *et al.*, 2000a, b). The present paper describes the complete DNA sequence of the medermycin biosynthetic gene cluster of *Streptomyces* sp. AM-7161. This is the third example (after actinorhodin and granaticin) of the cloning of an entire cluster for BIQ biosynthesis, allowing further insight into aromatic polyketide biosynthesis in *Streptomyces* spp. on the basis of comparative genomics.

METHODS

Bacterial strains, plasmids and culture conditions. *Streptomyces* sp. strain AM-7161 was obtained from Professor Satoshi Ōmura of the Kitasato Institute, and maintained as described by Ōmura *et al.* (1986). *S. coelicolor* CH999 (*proA1 argA1 redE60* Δ*act*::*ermE* SCP$^-$ SCP2$^-$) was described by McDaniel *et al.* (1993). *Escherichia coli* DH5α was used for standard cloning experiments. *E. coli* strains (Stratagene) for cosmid manipulations were XL-1 Blue-MR and XL-1 Blue-MRF′. Plasmids were passed through *E. coli* ET12567 (*dam dcm hsdS*) to generate unmethylated DNA before their use to transform *S. coelicolor* CH999 (MacNeil *et al.*, 1992). pBluescript SK$^{-/+}$ (pBS-SK$^{-/+}$) and pT7Blue(R) T-Vector were from Stratagene and Novagen, respectively. Cosmid pKU402 (Pang *et al.*, 1994) was kindly provided by Professor Haruo Ikeda, Kitasato University. Plasmid pTST59.1 was a generous gift from Dr Josef Altenbuchner, University of Stuttgart, Germany.

General genetic manipulation. Genomic DNA was isolated as described by Kieser *et al.* (2000). Restriction mapping and general molecular biology methods were performed as described by Sambrook *et al.* (1989). PCRs were performed with AmpliTaq Gold (Perkin Elmer) and Ex-Taq (Takara). Synthetic oligonucleotides for PCR primers were obtained from Nihon Bioservice (Saitama, Japan). The following primers were used in this study: PS-A-S, 5′-GGVTTCGGSGGSTTCCAGAGCGC-3′; PS-A-A, 5′-CCAGGCGAAS GACTSGTASRCGCTCAC-3′; PS-B-S, 5′-CSGGSGSSGCSGGSTTCAT CGG-3′; PS-B-A, 5′-GGGWRCTGGYRSGGSCCGTAGTTG-3′; Neo-F, 5′-AGACAATCGGCTGCTCTGATG-3′; Neo-R, 5′-TAAAGCACGA GGAAGCGGTCAGCCC-3′.

PCR amplification and hybridization. Genomic (approx. 0·5 μg template) PCR with the PS-A-S and PS-A-A primers was performed in a final volume of 50 μl for 40 cycles of amplification using a step programme (0·5 min at 94 °C, 0·5 min at 64 °C and 1 min at 72 °C) in the Robocycler Gradient 40 (Stratagene) according to the manufacturer's protocol except for the addition of 5 % DMSO. The PCR product of expected size (approx. 0·45 kb) was subcloned into pT7Blue T-vector and propagated in *E. coli* DH5α. DNA sequencing revealed a unique 467 bp gene. Amplification with PS-B-S and PS-B-A was carried out in an essentially similar manner except for the annealing temperature of 68 °C to afford an approximately 0·55 kb product. Sequence analysis of the product identified a unique 541 bp gene. Both products were DIG-labelled (PKS and DOH probes) following the manufacturer's protocol (Roche Diagnostics). Southern hybridization was performed with Nytran N membranes (Schleicher & Schuell) at 68 °C. Membranes were washed once with 1× SSC in 0·1 % SDS at room temperature for 5 min and twice with 0·1× SSC in 0·1 % SDS at 68 °C for 15 min. Subsequent colorimetric detection was carried out with 4-nitroblue tetrazolium chloride (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (X-phosphate) following the standard protocol (Roche Diagnostics).

Library construction and screening. A genomic library of *Streptomyces* sp. AM-7161 was prepared in cosmid pKU402 as described by Pang *et al.* (1994). A total of 1152 independent clones were screened with the PKS and DOH probes to identify seven positive clones, two of which (pIK130 and pIK133) were positive for the both probes. pIK130 was further characterized by restriction mapping and the overlapping subclones were subsequently generated based on pBS-SK$^{-/+}$ for sequence analysis.

DNA sequencing and computer-assisted sequence analysis. DNA was sequenced on an automated DNA sequencer model 4000L or 4200L (LI-COR). Sequencing reactions were carried out on double-stranded plasmid DNA (pBS-SK$^{+/-}$) with the Thermo Sequenase Cycle Sequencing kit (Amersham-Pharmacia). DNA sequences were analysed using the DNASIS software package (version 3.7; Hitachi Software Engineering). A version of the FRAME program (Bibb *et al.*, 1984), MacFRAME (version 1.2, developed by K. Kendal, Tulane University, LA, USA) was used to identify potential protein-coding regions. Deduced gene products were analysed using a version of BLASTP provided by the DNA Database of Japan (DDBJ, http://www.ddbj.nig.ac.jp). Identity/similarity scores were obtained by the BLAST2 SEQUENCES program available from National Center for Biotechnology Information, National Library of Medicine, NIH (http://www.ncbi.nlm.nih.gov/). Phylogenetic analysis was carried out with a version (http://spiral.genes.nig.ac.jp/homology/clustalw.shtml) of the CLUSTAL W program, provided by DDBJ, based on a neighbour-joining method. A phylogenetic tree was created with the TreeView program (version 1.6.2, freely available from the Taxonomy and Systematics server at the University of Glasgow, UK). The number of bootstrap replications was 1000.

Construction of expression plasmid. pTST59.1 is a derivative of the SUPERCOS1 cosmid (Stratagene) carrying the *att*P site and integrase derived from ΦC31 (Lomovskaya *et al.*, 1980). The vector was digested with *Xba*I followed by calf intestine alkaline phosphatase treatment and *Bam*HI digestion. The entire insert of pIK130 was released by *Dra*I digestion, and ligated with a *Bam*HI–*Sma*I adaptor and p*Sma*I linker (Takara). The resulting fragment was ligated with the foregoing vector, and the mixture was packaged followed by transfection into *E. coli* XL-1 Blue MRF′. The completed expression plasmid was designated as pIK340.

Heterologous expression of the *med* cluster and chemical characterization. Protoplast formation, transformation and regeneration of protoplasts from *S. coelicolor* CH999 were carried out by standard procedures (Kieser *et al.*, 2000). Genome integration was confirmed by colony PCR with the primer set Neo-F and Neo-R for kanamycin-resistant clones as described by Ishikawa *et al.* (2000). Transformants were liquid-cultured as described previously (Taguchi *et al.*, 2000a). The medium separated from the broth by centrifugation was directly subjected to reversed-phase HPLC analysis on a TOSOH 8020 system under the following conditions: column, TSK gel ODS-80TM (4·6 mm i.d. × 150 mm, TOSOH); column temperature, 40 °C; gradient elution, solvent A (0·5 % acetic acid in CH_3CN) and solvent B (0·5 % acetic acid in distilled H_2O), gradient profile (0–5 min, 20 % A; 5–25 min, 20–70 % A; 25–28 min, 70–95 % A; 28–32 min, 95 % A; 32–35 min, 95–20 % A); flow rate, 1·0 ml min^{-1}; photo-diode array detector (PD-8020, TOSOH), 250–600 nm. LC/MS spectra were recorded on a Thermoquest LCQ equipped with a Hewlett Packard HP1100 series LC system under the same chromatographic conditions as used for the HPLC analysis. Atmospheric pressure chemical ionization (APCI) was applied to detect positive and negative ions of samples. Medermycin was also isolated from CH999/pIK340 as follows. Culture broth (1·2 l) was neutralized with 1 M NaOH, and extracted with ethyl acetate. The combined extracts were washed with brine and water, followed by drying over Na_2SO_4. The residue, after evaporation of the solvent, was subjected to chromatography on silica gel (Wako-gel C200) with chloroform/methanol (15 : 1) to give medermycin (5 mg). The purified compound gave NMR data (available as supplementary data with the online version of this paper at http://mic.sgmjournals.org) consistent with the published values for natural medermycin/lactoquinomycin A (Okabe *et al.*, 1985; Tatsuta *et al.*, 1990; Williamson *et al.*, 2002).

RESULTS

Cloning of the medermycin biosynthetic gene cluster

Medermycin biosynthesis consists of polyketide-aglycone formation followed by attachment of the sugar moiety, angolosamine, produced via the deoxyhexose (DOH) pathway. The cloning strategy was based on PCR amplifications of conserved genetic regions encoding the PKS as well as a key DOH enzyme. In anticipation of the clustering of the entire biosynthetic genes on the genome as encountered in the cloning of the granaticin biosynthetic genes (Ichinose *et al.*, 1998b), and those for many other *Streptomyces* antibiotics, the resultant genes were to be used as screening probes of a genomic library of the medermycin producer, *Streptomyces* sp. AM-7161. Genetic organization is conserved among the minimal PKS (type II) components, ketosynthase$_\alpha$ (KS$_\alpha$), ketosynthase$_\beta$ (KS$_\beta$) and acyl carrier protein (ACP). In particular, the universal tandem pair of KS$_\alpha$ and KS$_\beta$, which are normally translationally coupled, led us to design degenerate primers from the C-terminus of KS$_\alpha$ (PS-A-S) to the N-terminus of KS$_\beta$ (PS-A-A) based on the known examples of BIQs, actinorhodin (Fernández-Moreno *et al.*, 1992), frenolicin B (Bibb *et al.*, 1994), griseusin B (Yu *et al.*, 1994) and granaticin (Sherman *et al.*, 1989). A DOH-specific enzyme, NDP-glucose-4,6-dehydratase, was successfully used for designing degenerate primers to clone the genes involved in DOH biosynthesis directed to a number of *Streptomyces* antibiotics. The suggested primers, PS-B-S and PS-B-A (Decker *et al.*, 1996), were applied to this study. PCR using genomic DNA of *Streptomyces* sp. AM-7161 as a template afforded PCR products of the expected sizes (approx. 450 bp for PKS; approx. 550 bp for DOH). Subcloning and sequence analysis of the PCR products afforded 467 bp PKS and 541 bp DOH fragments, likely to encode PKS (KS$_\alpha$-KS$_\beta$) and a DOH enzyme (NDP-glucose-4,6-dehydratase), respectively. Both proteins closely resembled (> 70 % similarity) the proteins used for the PCR primer design. Genomic hybridization analysis (data not shown) with the PCR products as probes revealed a single positive restriction fragment on electrophoresis: PKS (7·7 kb *Bam*HI, 3·3 kb *Pst*I); and DOH (10 kb *Bam*HI, 4·2 kb *Pst*I). A cosmid library was constructed using approximately 40 kb insert DNA prepared by partial *Sau*3AI digestion of the genomic DNA. Library screening with the PCR products as hybridization probes identified seven positive clones out of 1152, with two clones (pIK130, 133) positive to both probes. We chose pIK130 for further characterization. Southern blot analysis suggested that the clone contained the same *Bam*HI fragments as revealed in the genomic blot analysis.

DNA sequence analysis

Restriction mapping was performed for the pIK130 insert, and its overlapping subclones were sequenced to cover the complete insert (36 202 bp, overall G+C content 72·8 mol%). Probable ORFs were identified by the FRAME

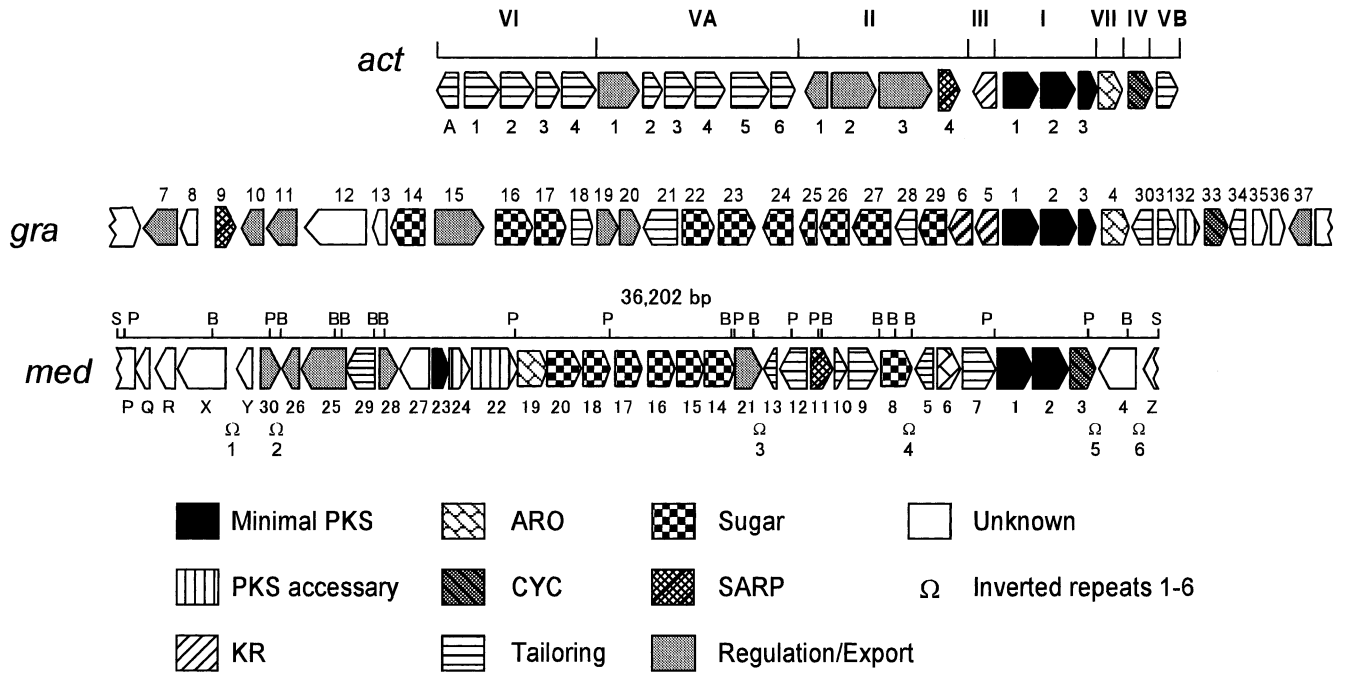


Fig. 2. Organization of the *med* cluster of *Streptomyces* sp. AM-7161 and, for comparison, the *act* cluster of *S. coelicolor* A3(2) and the *gra* cluster of *S. violaceoruber* Tü22. Restriction sites used for mapping: B, *Bam*HI; P: *Pst*I; S, *Sau*3AI. The minimal PKS consists of ketosynthase subunits and acyl carrier protein (ACP). KR, ketoreductase; ARO, aromatase; CYC, cyclase; SARP, *Streptomyces* antibiotic regulatory protein.

program (Bibb *et al.*, 1984), and by potential ribosome-binding sites (Strohl, 1992). The deduced ORFs were functionally designated based on database searches as shown in Fig. 2. We identified 34 complete ORFs (Table 1), with an incomplete ORF at either end, and several inverted repeats in intergenic regions. The cluster was designated as the *med* cluster. Following is the description of the *med*-ORFs with the corresponding *act* genes, if any, as a reference.

PKS and closely related ORFs

The early biosynthetic enzymes closely associated with the minimal PKS include a ketoreductase (KR) functioning at C-9, an aromatase (ARO) to aromatize the first ring by dehydration, and a cyclase (CYC) controlling second ring formation between C-5 and C-14. Taken together they produce the bicyclic intermediate of octaketide origin, which is the common intermediate for medermycin, actinorhodin and granaticin (Fig. 3). Nearly all the known examples of the minimal PKS (Hopwood, 1997) include a gene encoding the ACP immediately downstream of a set of KS_α and KS_β. Surprisingly, the ACP gene (*med*-ORF23) was discovered 20 kb upstream of the *med*-ORF2 (KS_β). Such an unusual organization is known for the biosynthetic gene clusters for daunorubicin (Grimm *et al.*, 1994), R1128 (Marti *et al.*, 2000) and griseorhodin A (Li *et al.*, 2002). Downstream of *med*-ORF23 are a pair of genes, *med*-ORF24 and 22, possibly related to polyketide starter formation and chain extension. The jadomycin B biosynthetic gene cluster from *Streptomyces venezuelae* contains a similar organization for *jadM* and *jadN* (Wang *et al.*, 2001). The *med*-ORF24 protein is similar to the product of *jadM*, which was proposed to encode a phosphopantetheinyl transferase, functioning as a holo-ACP synthase for jadomycin B biosynthesis. Other *med*-ORF24 homologues (similarity/identity at amino acid level) include *pgaX* from *Streptomyces* sp. PGA64 (65/53 %, GenBank accession no. AY034378) and *mtaA* from *Stigmatella aurantiaca* (48/35 %, GenBank accession no. AF188287). *gra*-ORF32, previously assigned as 'unknown' (Ichinose *et al.*, 1998b), now belongs to this family, encoding a protein with 38 % similarity to the *med*-ORF24 product. Very high homology was found for *med*-ORF22 with *jadN*, and also with *lanP* from *Streptomyces cyanogenus*, probably involved in landomycin biosynthesis (Westrich *et al.*, 1999). The homologues both resemble acyl-CoA decarboxylase genes. The *med*-ORF6 protein is an apparent homologue of the KR encoded by *actIII*, controlling C-9 keto-reduction. The ARO gene, *med*-ORF19, was found over 15 kb apart from the KS_α gene, *med*-ORF1, rather closer to the ACP gene, *med*-ORF23. *med*-ORF3, encoding CYC, is downstream of *med*-ORF2, with putative translational coupling. Thus, the genes required for the formation of the bicyclic intermediate were identified as expected.

Post-PKS tailoring steps in aglycone formation

Stereochemical control. The bicyclic intermediate lies at the biosynthetic branch point towards various BIQ

Table 1. Deduced functions of ORFs

ORF	Size (aa)	Putative functional category	Homologues*				
			Gene	Deduced role	SM/ID (%)†	Origin	Nucleotide accesssion no.
P	217+	No relation?	*cysD*	Sulfate adenylyltransferase	88/84	*S. coelicolor*	AL939126
Q	177	No relation?	*cysC*	Adenylylsulfate kinase	90/84	*S. coelicolor*	AL939126
R	235	No relation?	*cysH*	Phosphoadenosine phosphosulfate reductase	91/85	*S. coelicolor*	AL939126
X	564	No relation?	SCBAC1A6.26c	Nitrile/sulphite reductase	89/82	*S. coelicolor*	AL939126
Y	189	No relation?	SCBAC1A6.27c	Acetyltransferase	81/73	*S. coelicolor*	AL939126
30	226	Regulation	*jadR1*	Repressor-response regulator	63/48	*S. venezuelae*	U24659
26	206	Unknown	*gra*-ORF19	Putative disulphide-forming protein	53/41	*S. violaceoruber*	AJ011500
25	525	Export	SMa0185	Transmembrane transport protein	53/40	*Sinorhizobium meliloti*	AE007204
29	341	Tailoring	*actVI-2*	Enoyl reductase	37/28	*S. coelicolor*	X62373
28	213	Regulation	SC5F1.06c	Tet-R family transcriptional regulator	44/36	*S. coelicolor*	AL939132
27	348	Unknown	*pgaK*	Unknown	57/44	*S.* sp. *PGA64*	AY034378
23	91	PKS	*actI*-ORF3	Acyl carrier protein (ACP)	82/74	*S. coelicolor*	X63449
24	238	Starter	*jadM*	Phosphopantetheinyl transferase	62/55	*S. venezuelae*	AF222693
22	538	Starter	*jadN*	Decarboxylase	88/83	*S. venezuelae*	AF026363
19	318	PKS	*actVII*	First ring aromatase (ARO)	53/42	*S. coelicolor*	X63449
20	376	DOH	*dnrJ*	NDP-deoxyhexose 3-aminotransferase	77/70	*S. peucetius*	M80237
18	354	DOH	*gra*-ORF16	NDP-1-glucose synthase	68/58	*S. violaceoruber*	AJ011500
17	333	DOH	*gra*-ORF17	NDP-glucose-4,6-dehydratase	76/64	*S. violaceoruber*	AJ011500
16	479	DOH	*gra*-ORF27	NDP-deoxyglucose-2, 3-dehydratase	77/68	*S. cyanogenus*	AJ011500
15	240	DOH	*snogX*	*N*-Methyl transferase	63/52	*S. nogalater*	AL224512
14	325	DOH	*gra*-ORF22	NDP-4-keto-6-deoxyhexose reductase	45/35	*S. violaceoruber*	AJ011500
21	338	Regulation	SC6G10.31c	Putative carbohydrate kinase	64/53	*S. coelicolor*	AL939111
13	167	Tailoring	*actVB*	NADH : FMN oxidoreductase	51/36	*S. coelicolor*	X63449
12	324	Tailoring	*actVI-1*	Stereospecific ketoreductase	65/57	*S. coelicolor*	X62373
11	258	Regulation	*actII-4*	Pathway-specific transcriptional activator	63/43	*S. coelicolor*	M64683
10	146	Tailoring	*actVI*-ORFA	Unknown	62/54	*S. coelicolor*	X62373
9	339	Tailoring	*actVI-2*	Enoyl reductase	57/48	*S. coelicolor*	X62373
8	376	DOH/tailoring	*urdGT2*	*C*-Glycosyl transferase	58/45	*S. fradiae*	AF16490
5	214	Tailoring	*actVI-3*	Pyran ring cyclase/possible exporter	57/39	*S. coelicolor*	X62373
6	261	PKS	*actIII*	Keto reductase (KR)	76/67	*S. coelicolor*	M19536
7	379	Tailoring	*actVA-5*	Oxygenase/hydroxylase	59/49	*S. coelicolor*	X58833
1	422	PKS	*actI*-ORF1	Keto-acyl synthase α (KSα)	81/72	*S. coelicolor*	X63449
2	408	PKS	*actI*-ORF2	Keto-acyl synthase β (KSβ)	72/63	*S. coelicolor*	X63449
3	302	PKS	*actIV*	Second ring cyclase (CYC)	71/57	*S. coelicolor*	X63449
4	409	No relation?	*csn*	Putative secreted chitosanase	76/66	*S. coelicolor*	AL939106
Z	165+	No relation?	SCBAC1A6.29c	Putative regulatory protein	68/59	*S. coelicolor*	AL939126

*Homologues are representative proteins; listing priority is given to those from the BIQ clusters, if any.
†SM/ID, % similarity/identity of amino acid sequences. PKS, polyketide synthase and related proteins; DOH, deoxyhexose pathway.

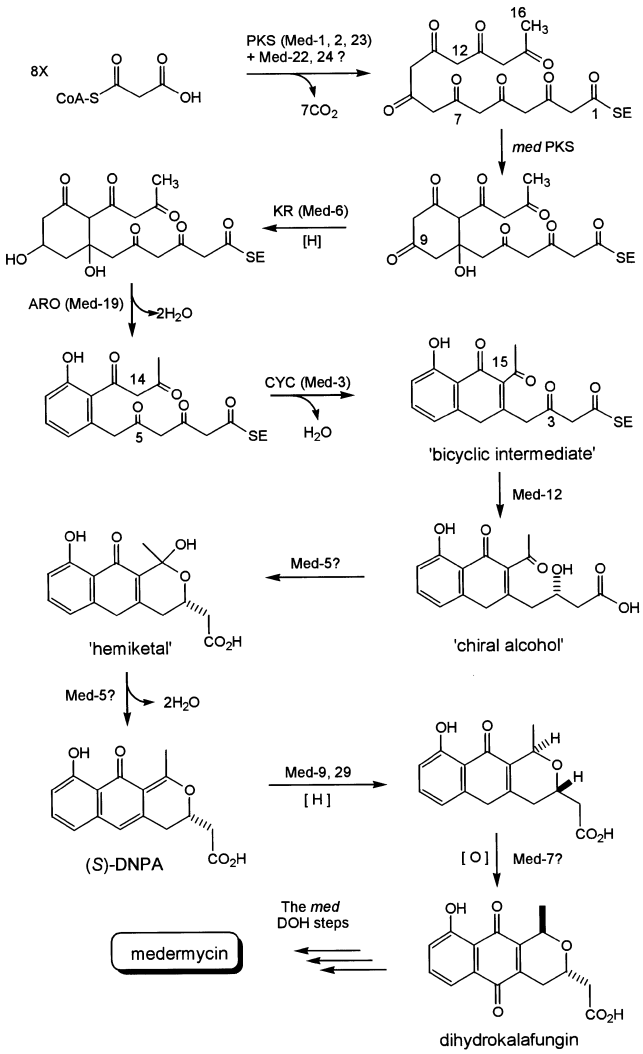


Fig. 3. Proposed biosynthetic pathway of the BIQ chromophore leading to medermycin and functional assignment of the *med* proteins. Biosynthetic intermediates are tentatively depicted as enzyme-bound (R-SE) to the bicyclic intermediate. (?) indicates tentative assignments. [H] and [O] show reduction and oxidation, respectively.

derivatives. The next step is stereospecific reduction at C-3 to afford either the 3*S*- or the 3*R*-configuration (Ichinose *et al.*, 2001; Taguchi *et al.*, 2001). In actinorhodin biosynthesis, *actVI*-ORF1 was proved to encode the reductase responsible (Ichinose *et al.*, 1999). Since medermycin possesses the same stereochemistry concerning the two chiral centres (3*S*, 15*R*) as actinorhodin, an *actVI*-ORF1 homologue was expected. A homologue, *med*-ORF12, was indeed found. The product ('chiral alcohol') of the C-3 reduction would cyclize to the hemiketal, followed by dehydration to afford (*S*)-DNPA (Cole *et al.*, 1987) (Fig. 3). Our earlier experiments (Taguchi *et al.*, 2000a) suggested the *actVI*-ORF2 protein, an enoyl reductase homologue, to function as the second reductase at C-15 to complete the (3*S*, 15*R*) configuration. The presence of another homologue, *actVI*-ORF4, prompted us to propose that the product might assist an unusual mode of reduction of (*S*)-DNPA (Taguchi *et al.*, 2000a). The reduction was proposed to proceed in two coupled steps, with stereospecific reduction (by ActVI-ORF2) of the double bond between C-14 and C-15 followed by isomerization (by ActVI-ORF4) to afford the α,β-unsaturated carbonyl structure, thus accomplishing the 1,4-reduction concerning C-15 and C-6. Here, we encountered the same situation as for the *act* cluster: two *actVI*-ORF2 homologues, *med*-ORF9 and *med*-ORF29 (37 % similar to each other), are present, in agreement with our assumption.

Oxygenation and hydroxylation. The oxygenase/hydroxylase would be involved in the introduction of oxygen functionalities at C-6 and C-8 of the BIQs. In actinorhodin biosynthesis, the *actVA*-ORF6 protein was proposed (Fernández-Moreno *et al.*, 1994) to oxidize C-6, and later biochemical proof was given for its catalytic activity (Kendrew *et al.*, 1997). This enzyme performs oxygenation without any prosthetic group, metal ion, or cofactor: its novel catalytic mechanism was revealed by a recent X-ray crystallographic study (Sciara *et al.*, 2003). The *med* cluster lacks an obvious homologue, and instead includes a gene, *med*-ORF7, similar to *actVA*-ORF5, deduced to encode a hydroxylase (Caballero *et al.*, 1991). *actVA*-ORF5 and its close homologue, *gra*-ORF21, were reasonably assigned to encode a hydroxylase at C-8 based on their significant similarity (>70 %) to *pheA*, encoding a phenol hydroxylase from *Bacillus stearothermophilus* (Kim *et al.*, 1995), correlating with the presence of the hydroxyl group at C-8 of actinorhodin and granaticin. The absence of an oxygen at C-8 of medermycin might allow us to assume the *med*-ORF7 protein to be an oxygenase at C-6. *med*-ORF13 is a clear homologue of *actVB*, encoding a flavin : NADH oxidoreductase (Kendrew *et al.*, 1995), and *gra*-ORF34 (Ichinose *et al.*, 1998b). This finding supports our proposal (Ichinose *et al.*, 1998b) of their possible roles in establishing the redox change at C-10 required for the substitution steps, dimerization for actinorhodin and *C*-glycosylation for granaticin, since medermycin is *C*-glycosylated at C-10.

***actVI*-ORFA homologue.** *actVI*-ORFA homologues constitute a family of genes widely found in the gene clusters for *Streptomyces* aromatic polyketides, the members being *fren*-ORFX (Bibb *et al.*, 1994), *gra*-ORF31 (Ichinose *et al.*, 1998b), *mtmX* from the mithramycin cluster of *Streptomyces argillaceus* (GenBank accession no. X89899), and *dpsH* from the daunorubicin clusters of *Streptomyces peucetius* (Gerlitz *et al.*, 1997) and *Streptomyces* sp. strain C5 (GenBank accession no. U43704). The *actVI*-ORFA disruptant produced reduced amounts of actinorhodin, together with biosynthetic intermediates and shunt products (Taguchi *et al.*, 2000a, b). A possible involvement of *dpsH* in polyketide chain cyclization was also postulated (Gerlitz *et al.*, 1997). Although the gene function is still speculative, it apparently controls the

efficiency of biosynthetic enzymes in a given cluster. *med*-ORF10 belongs to this family.

***actVI*-ORF3 homologue.** The phenotypic and chemical characterization of an *actVI*-ORF3 disruptant (Fernández-Moreno *et al.*, 1992; Taguchi *et al.*, 2000a) suggested that the product might catalyse chemically spontaneous processes such as the hemiketal formation and dehydration to afford (*S*)-DNPA (Fig. 3). Recent proteomic analysis (Hesketh *et al.*, 2002) of the *act* enzymes in *S. coelicolor* A3(2) suggested an interesting possibility for a role in exporting actinorhodin (see Discussion). Unlike the *actVI*-ORFA family, the *actVI*-ORF3 homologues are restricted to the clusters for BIQs. *gra*-ORF18 and the gene herein discovered, *med*-ORF5, are the only other family members.

DOH pathway. The biosynthetic genes for the angolosamine moiety were found to be contiguous and transcribed in the same direction. Examples of such clustering for DOH pathways in *Streptomyces* glycoside antibiotics are the D-olivose/L-rhodinose genes for urdamycin A formation in *Streptomyces fradiae* (Hoffmeister *et al.*, 2000), L-mycarose for tylosin in *S. fradiae* (GenBank accession no. AF147704), L-digitoxose for jadomycin B in *S. venezuelae* (GenBank accession no. AY026363), and L-desosamine for pikromycin/methymycin (GenBank accession no. AF079138). The first three steps are common with those for the DOH groups (4-keto-2,6-dideoxyglucose and L-rhodinose) of granaticin (Bechthold *et al.*, 1995; Ichinose *et al.*, 1998b): NDP-glucose synthase (encoded by *gra*-ORF16), NDP-glucose-4,6-dehydratase (encoded by *gra*-ORF17) and NDP-glucose-2,3-dehydratase (encoded by *gra*-ORF27), all of which were reasonably assigned to the respective gene products of *med*-ORF18, *med*-ORF17 and *med*-ORF16 based on significant similarity scores (>70 %). The *med*-ORF20 protein resembles a class of aminotransferases involved in the biosynthesis of *Streptomyces* antibiotics containing aminodeoxysugars. Representative proteins are DnrJ for daunosamine formation in daunorubicin biosynthesis in *S. peucetius* (Madduri *et al.*, 1995) and DesV (with 54 % sequence similarity) for desosamine production related to macrolide antibiotics in *S. venezuelae* (GenBank accession no. AF079138). The *med*-ORF20 protein was reasonably assigned as an aminotransferase at C-3′. All of the foregoing genes appear to be translationally coupled. The remaining steps, which may be interchangeable, are catalysed by 4′-keto-reductase and *N*-methyltransferase. The genes responsible were found downstream of *med*-ORF16 as a translationally coupled pair: *med*-ORF15 and *med*-ORF14. 4′-Keto reduction is required for granaticin B (L-rhodinose) biosynthesis: the assigned gene, *gra*-ORF22 (Tornus *et al.*, 2001), encodes a protein 45 % similar to the *med*-ORF15 product. Other homologues include *urdZ3* for urdamycin A (47 % similarity, Hoffmeister *et al.*, 2000), *lanZ3* for landomycin (50 % similarity, Westrich *et al.*, 1999) and *snoG* for nogalamycin biosynthesis in *Streptomyces nogalater* (45 % similarity, GenBank accession no. AF187532). Strong similarity was found for *med*-ORF15 to a family of *N*-methyltransferase genes, *snoX* for nogalamycin (Ylihonko *et al.*, 1996) and *rdmD* for rhodomycin biosynthesis in *Streptomyces purpurascens* (56 %, GenBank accession no. U10405). Functional assignment of the DOH genes was thus completed (Fig. 4).

***C*-glycosylation.** *C*-glycosylation is one of the characteristic features of the BIQs. For example, granaticin is derived from the attachment of 4-keto-2,6-dideoxy-D-glucose to the BIQ aglycone via two C–C bonds. The *gra* cluster contains a single apparent glycosyltransferase (GT) gene, *gra*-ORF14 (Ichinose *et al.*, 1998b). The *gra*-ORF14 protein resembles the product of *med*-ORF8. Extensive sequence comparison of *med*-ORF8 with other *Streptomyces* GT genes indicated its probable role in *C*-glycosylation, which is required for angolosamine transfer in medermycin biosynthesis (see Discussion).

Regulatory and antibiotic export genes

The pathway-specific activator genes, now known as *Streptomyces* antibiotic regulatory proteins, SARPs (Wietzorrek *et al.*, 1997), are a growing family of genes found mainly in the gene clusters for *Streptomyces* aromatic polyketides. *actII*-ORF4 (Fernández-Moreno *et al.*, 1991) is the best-characterized member. *med*-ORF11 is a clear homologue. *actII*-ORF4 contains a rare triplet codon, TTA, for leucine, involved in a key regulatory role in actinorhodin production (Fernández-Moreno *et al.*, 1991), but the *med* protein contains no TTA codon. In this it resembles *redD* in the undecylprodigiosin pathway of *S. coelicolor*, which is controlled by a second regulator, *redZ*, which contains a TTA codon (Guthrie *et al.*, 1998).

The product of *med*-ORF26 resembles those of *gra*-ORF19 (Ichinose *et al.*, 1998b) and *frnE* (GenBank accession no. AF058302), becoming 'common' proteins in the BIQ clusters. We proposed the function to be a redox-sensor by forming a disulphide bridge (Ichinose *et al.*, 1998b). All of the three homologues contain a key motif of CXXC (^{11}CPWC^{16} for Med-ORF26) found in the glutaredoxin/thioredoxin superfamily. The *med*-ORF30 protein is homologous with the product of *jadR1* in the jadomycin B cluster, which is deduced to encode an essential transcriptional regulator (GenBank accession no. U24659).

The only gene found for a possible antibiotic transporter is *med*-ORF25, whose product is homologous with a probable transmembrane-transport protein, SMa0185, from *Sinorhizobium meliloti* (See Table 1). Other regulatory genes in the *med* cluster are *med*-ORF21 (a possible kinase gene) and *med*-ORF28 (a putative Tet-R family transcriptional regulator gene) (see Table 1).

'Unknown' and unassigned genes

Downstream of the essential PKS gene, *med*-ORF3, is *med*-ORF4. The product resembles a family of chitosanases,

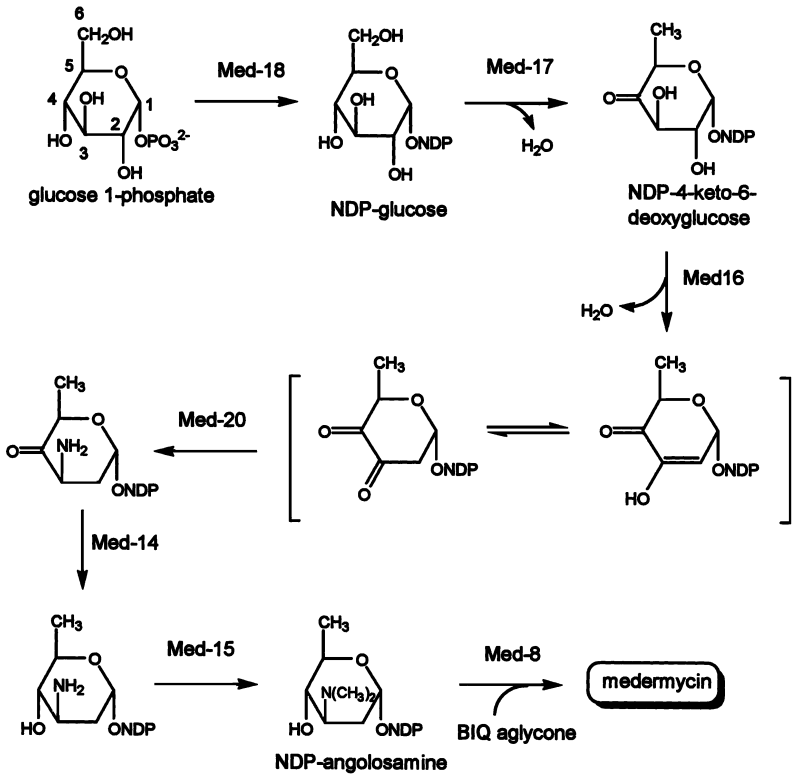


Fig. 4. Proposed biosynthetic pathway of angolosamine and functional assignment of the *med* proteins. NDP, nucleotide diphosphate (N: A, T, U, C or G).

which catalyse the hydrolysis of chitosan. Typical protein length is about 300 aa, whereas Med-ORF4 is 100 aa longer (409 aa). Its region of homology to a putative chitosanase gene, *csn*, from *S. coelicolor* (Bentley *et al.*, 2002) was localized towards the C-terminus. Since chitosanases are catabolic enzymes, *med*-ORF4 appears to be unrelated to medermycin biosynthesis. A lower G+C (approx. 65 mol%) region was present upstream of *med*-ORF23. FRAME analysis suggested a possible protein-coding region, designated as *med*-ORF27. The deduced product showed partial similarity (approx. 50 %) to the products of 'unknown' genes found in *Streptomyces* antibiotic gene clusters, including *pgaK* from the rubromycin B-producing *Streptomyces* sp. PGA64 (GenBank accession no. AY034378), ORF12 from the tylosin producer *S. fradiae* (Bate *et al.*, 1999) and an incomplete ORF, *aur1*O, from the auricin-producing *Streptomyces aureofaciens* (Novakova *et al.*, 2002). All other ORFs (P, Q, R, X, Y and Z) identified were highly homologous with hypothetical proteins revealed by the genome sequencing project (Bentley *et al.*, 2002) of *S. coelicolor* (Table 1).

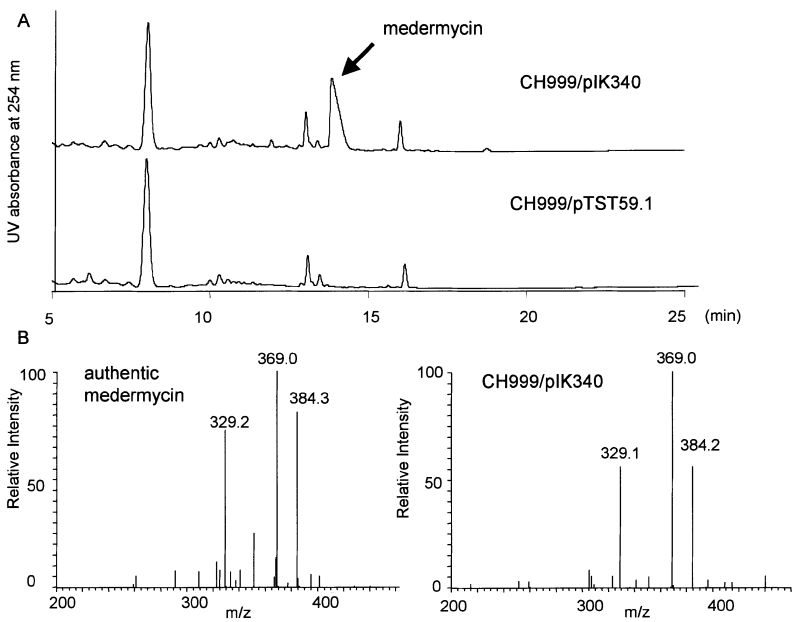


Fig. 5. LC-APCIMS analysis of *S. coelicolor* CH999 recombinants. (A) HPLC profiles (UV absorption at 254 nm) of CH999/pTST59.1 and CH999/pIK340. (B) MS/MS spectra (precursor ion at *m/z* 458 $[M+H]^+$) of authentic medermycin and CH999/pIK340 product.

Heterologous expression of the *med* cluster

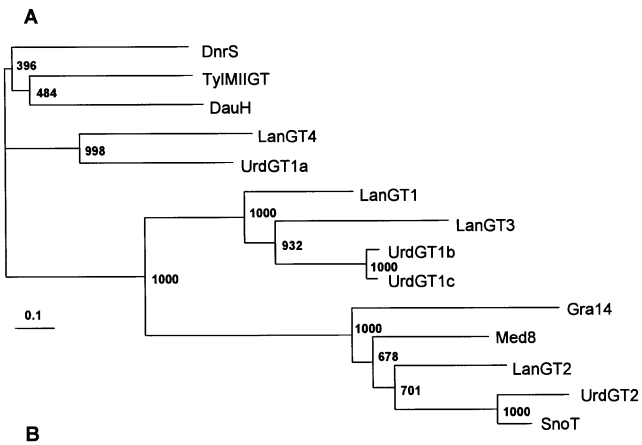
Deduction of all essential biosynthetic steps to medermycin from the deduced *med* protein functions induced us to attempt heterologous expression of the entire *med* cluster. An integrative expression vector (pTST59.1) was chosen for stable replication of the large DNA fragment derived from a cosmid clone. The whole insert of pIK130 was used to construct pIK340, which was then introduced into *S. coelicolor* CH999 (McDaniel *et al.*, 1993). Transformants on plates gave brownish pigmentation, and they were subjected to liquid culture followed by HPLC analysis. The recombinant, CH999/pIK340, gave a peak corresponding to standard medermycin, which was not observed in the profile derived from the control recombinant, CH999/pTST59.1 (Fig. 5A). LC/MS analysis suggested its molecular mass to be 457 Da, which is consistent with the formula of medermycin ($C_{24}H_{27}O_8N$). MS/MS spectra (the precursor ion at *m*/*z* 458) for the sample from CH999/pIK340 and authentic medermycin were identical (Fig. 5B). Final confirmation using NMR spectra (available as supplementary data with the online version of this paper at http://mic.sgmjournals.org) indicated clearly that the sample was medermycin. This result provides definite proof that the *med* cluster described above is indeed sufficient for medermycin production.

DISCUSSION

We have cloned and sequenced a 36 kb genomic fragment of *Streptomyces* sp. AM-7161. Its heterologous expression in *S. coelicolor* CH999 resulted in the production of medermycin, unambiguously proving that the fragment carries the whole biosynthetic gene cluster for medermycin (the *med* cluster). The cluster contains 34 complete ORFs, together with two incomplete ORFs at either end. What are the possible limits of the *med* cluster? *med*-ORF30 at the left-hand end shows obvious similarity to known essential genes in *Streptomyces* antibiotic gene clusters. The right-hand terminus could be *med*-ORF3, because *med*-ORF4 and ORFZ are deduced to encode proteins unrelated to medermycin production. We therefore propose that the *med* cluster spans the region from ORF30 to ORF3 (29 ORFs) over 30 kb. Compared with the other complete BIQ clusters (the *act* cluster has 21 ORFs over 22 kb; the *gra* cluster has 32 ORFs over 33 kb), the size is reasonable.

Mechanistically, medermycin shares the early biosynthetic pathway with actinorhodin and granaticin. The genes encoding the PKS and cyclizing enzymes (ARO and CYC) were found, two of which (*med*-ORF23 for ACP and *med*-ORF19 for ARO) have unusual locations upstream of those for the core PKS components, KS_α and KS_β. Post-PKS tailoring steps for aglycone formation were expected to follow the actinorhodin pathway. The *med* cluster provides the genes for the key reductases for stereochemical control, *med*-ORFs 12, 9 and 29. A structural difference of the BIQ chromophore between medermycin and actinorhodin concerns the hydroxyl group at C-8. Earlier functional assignment of ActVA-ORF5 (Caballero *et al.*, 1991) as a hydroxylase at C-8 is open to discussion based on the present finding of *med*-ORF7. On the other hand, the *med* cluster lacks an *actVA*-ORF6 homologue, encoding an essential monooxygenase to introduce oxygen at C-6 of actinorhodin (Kendrew *et al.*, 1997; Sciara *et al.*, 2003). This situation is reasonably explained by our tentative assignment of Med-ORF7 as an oxygenase at C-6.

An increasing number of genes for DOH biosynthesis have been cloned (Trefzer *et al.*, 1999). The angolosamine biosynthetic genes were found to be contiguous in the *med*



B

Med8	275	PTCDLLVHHGGGVTGMTALAAGTPQLLLS-RWDIFASTMRRTAER--GAA	321
Gra14	275	RTCDLLVHRAGGNTMLHAIVCGVPQLVIP-AMPKQVGMSARLAEY--GAA	321
UrdGT2	265	PTCDLLVHHAGGVSTLTGLSAGEPQLLIP-KGSVLEAPARRVADY--GAA	311
LanGT2	275	PTCDVVVHHAGGVTALTAMNAGVPQLIVS-QGGNFVEAGLRISDF--GAA	321
SnoT	274	PTCDVLVHHGGGVSTLTGLNAGVPQLLIP-RGAVLEKPALRVADH--GAA	319
UrdGT1a	319	PTCSAVVHHGGGGTFAAAVAHQVPQLI-PAP--GEGGDRVAFARYVEQRG	364
UrdGT1b	284	DGCAAVVHHGGAGTTMTAGAFGLPQLVLP-QLADQFGHGDRVSE--VGAG	330
UrdGT1c	283	DGCAAVVHHGGAGTTMTAGAFGLPQLVLP-QLADHFGHGDRVSE--VGAG	329
LanGT1	285	GSCHAVVHHGGSGTAMTASAFGLPQLVLP-QLADQFWHGDALAR--VGAA	331
LanGT3	294	RTADAVVHHGGAGTTMSAGYAGLPQLVLP-QFADTFASAGQLVR--AGAG	340
LanGT4	311	PTCSAIIHHGGGGTFAAAVAHQVPQLVTPMPFWGEAAT----AQYVADNG	356
TylMII	338	PTCSAVVHHGGAGTCFTATLNGLPQIVVAAL-WDAPLKGAQLAE--AGAG	384
DnrS	316	PSCAAVVHHGGAGTWATAALHGVPQLALAWQ-WDDVFRAGQLEK--LGAG	362
DauH	318	PSCSGIIHHGGSGTFMTALAHATPQLIVPDMMWDAMEKAHGLAR----SG	362

Fig. 6. (A) Phylogenetic tree of selected *Streptomyces* glycosyltransferases. See below for abbreviations of clone designations. The indicated scale represents 0·1 amino acid substitution per site. The bootstrap values are shown at the branch points. (B) Sequence alignment of the conserved regions of the *Streptomyces* glycosyltransferases involved in antibiotic biosynthesis. Conserved residues are in bold, and the Gly-rich motifs are indicated with a solid bar. The strictly conserved Pro (#) and conserved Asp (*) are marked. Entries of the sequence (source, protein accession number): Gra14 (*S. violaceoruber*, CAA09635); SnoT (*S. nodosus*, CAA01753); UrdGT1a, UrdGT1b, UrdGT1c and UrdGT2 (*S. fradiae*, AAF00214, AAF00215, AAF00217, AAF00209); LanGT1, LanGT2, LanGT3 and LanGT4 (*S. cyanogenus*, AAD13555, AAD13562, AAD13559, AAD13562); TylMII (*S. fradiae*, CAA57472); DnrS (*S. peucetius*, AAD15267); DauH, *Streptomyces* sp. C5 (AAB08020). The numbers given are the amino acid positions.

cluster, possibly providing a useful gene cassette of DOH production for metabolic engineering (Tang *et al.*, 2001). One of our initial interests was *C*-glycosylation in the BIQs. The *gra* cluster for granaticin biosynthesis contains a single GT gene, *gra*-ORF14, although the producer *Streptomyces violaceoruber* produces an additional granaticin derivative, granaticin B, which is the *O*-rhodinoside of granaticin. To reconcile two glycosylation steps with the presence of a single GT, two alternatives were proposed: (a) Gra-ORF14 catalyses the rhodinosyl transfer, and the *C*-glycosylation is catalysed by an 'unknown' enzyme; (b) Gra-ORF14 catalyses both glycosylations. Since the reaction centres of an acceptor molecule for *O*- and *C*-glycosylations are quite different, it would be hard to imagine that Gra-ORF14 could have such dual activities. We then proposed (Ichinose *et al.*, 1998b) that the first sugar attachment might proceed in rather unusual fashion, with nucleophilic substitution of the C-10 of an aglycone with the keto group at C′-4, giving a plausible explanation to support possibility (a). In the meantime, a *C*-GT gene, *urdGT2* for urdamycin biosynthesis (Faust *et al.*, 2000), was characterized. Its *C*-glycosylating activity for D-olivose was demonstrated by gene inactivation to result in the accumulation of aglycone derivatives, urdamycins I, J and K (Künzel *et al.*, 1999). The UrdGT2 protein shows strong similarities to Gra-ORF14 (53 %) and Med-ORF8 (58 %). Phylogenetic analysis of *Streptomyces* GTs indicated that the three proteins, along with two others (LanGT2 and SnoT), occupy a distinct position from those of *O*-GTs (Fig. 6A). Recent crystallographic studies shed light on the catalytic mechanism of GTs (Ünligil *et al.*, 2000). Particularly noteworthy is the structural study of the UDP-glucosyltransferase, GtfB, involved in the biosynthesis of vancomycin-group antibiotics (Mulichak *et al.*, 2001), where the highly conserved Gly-rich sequence HHGGAGT, the strictly conserved Pro, and the highly conserved Asp (the potential catalytic base) were highlighted. Sequence alignment of the entries used for the foregoing phylogenetic analysis showed that the Gly-rich motif and the Pro residue were universally aligned, whereas conservation of the possible base residue, Asp, was mainly restricted to *O*-GTs (Fig. 6B). Since the general base is supposed to abstract a proton from an acceptor hydroxyl group, this residue might determine the specificity for *C*-glycosylation. Taking these results together, we favour the idea that Gra-ORF14 and Med-ORF8 both function as *C*-GTs.

This study presents the third example of a complete biosynthetic gene cluster for BIQ biosynthesis, allowing comparative analysis of the homologous genes. Actinorhodin has served as one of the pioneering model compounds for understanding the PKS, tailoring steps and regulatory aspects of antibiotic production (Hopwood, 1997). The recent completion (Bentley *et al.*, 2002) of the genome sequencing project for the actinorhodin producer, *S. coelicolor* A3(2), provided a new direction of post-genome research, such as proteomics. Proteomic analysis (Hesketh *et al.*, 2002) of mycelial extracts suggested that ActVI-ORF3 is exported to the cell wall matrix, with its N-terminal signal peptide truncated. This provides an intriguing possibility for dual activities of ActVI-ORF3, extracellular as well as intracellular, leading to the efficient biosynthesis of actinorhodin. Since no definite roles of the

Table 2. Sequence-based comparison of the genes in the BIQ clusters

Functionally assigned genes are listed. (–) indicates the absence of a homologue.

PKS and related genes												
act	I-1	I-2	I-3	III	VII	IV	–	–				
gra	1	2	3	5, 6*	4	33	–	32				
med	1	2	23	6	19	3	22	24				
Tailoring steps												
act	VI-A	VI-1	VI-2, 4	VI-3	VA-2	VA-3	VA-4	VA-5	VA-6	VB		
gra	31	–	–	18	–	28, 30	–	21	–	34		
med	10	12	9, 29	5	–	–	–	7	–	13		
DOH pathway												
act	–	–	–	–	–	–	–	–	–	–	–	–
gra	14	16	17	22	23	24	25	26	27	29	–	–
med	8	18	17	14	–	–	–	–	16	–	15	20
Regulation and export												
act	II-1	II-2	II-3	II-4	VA-1	–	–	–	–	–	–	–
gra	–	15	–	9	–	10	11	19	20	–	–	–
med	–	–	–	11	–	–	–	26	–	21	25	30

**gra-6* encodes the C-3 reductase, which is a tailoring enzyme (Taguchi *et al.*, 2001).

'unknowns' such as ActVI-ORF3 are available from database analysis, combined proteomics and comparative genomics would be the most efficient approach to highlight the critically important genes in a given metabolic pathway. The finding of an *actVI*-ORF3 homologue, *med*-ORF5, clearly contributes to this approach. What other genes are common among the three complete BIQ clusters? Table 2 shows the sequence-based comparison of the cluster genes. Apart from the functionally defined genes, the *actVI*-A, *actVA*-ORF5 and *actVB* families each contain members in all three clusters. Priority should be given to these genes for further characterization in the quest for comprehensive understanding of the biosynthesis of BIQ antibiotics.

ACKNOWLEDGEMENTS

We are grateful to Professor Satoshi Ōmura, The Kitasato Institute, and Professor Haruo Ikeda, Kitasato University, for the generous gifts of strain and materials. We thank Dr Josef Altenbuchner, University of Stuttgart, for valuable technical advice. This work was supported by a Grant-in-Aid for Scientific Research on Priority Areas 'Genome Biology' from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (14014210), to K. I. We thank David Hopwood for critical reading of the manuscript.

REFERENCES

Bate, N., Butler, A. R., Gandecha, A. R. & Cundliffe, E. (1999). Multiple regulatory genes in the tylosin biosynthetic cluster of *Streptomyces fradiae*. *Chem Biol* **6**, 617–624.

Bechthold, A., Sohng, J. K., Smith, T. M., Chu, X. & Floss, H. G. (1995). Identification of *Streptomyces violaceoruber* Tü22 genes involved in the biosynthesis of granaticin. *Mol Gen Genet* **248**, 610–620.

Bentley, S. D., Chater, K. F. Cerdeno-Tarraga, A. M. & 40 other authors (2002). Complete genome sequence of the model actinomycete *Streptmyces coelicolor* A3(2). *Nature* **417**, 141–147.

Bibb, M. J., Findlay, P. R. & Johnson, M. W. (1984). The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. *Gene* **30**, 157–166.

Bibb, M. J., Sherman, D. H., Ōmura, S. & Hopwood, D. A. (1994). Cloning, sequencing and deduced functions of a cluster of *Streptomyces* genes probably encoding biosynthesis of the polyketide antibiotic frenolicin. *Gene* **142**, 31–39.

Caballero, J. L., Martinez, E., Malpartida, F. & Hopwood, D. A. (1991). Organisation and functions of the *act*VA region of the actinorhodin biosynthetic gene cluster of *Streptomyces coelicolor*. *Mol Gen Genet* **230**, 401–412.

Cole, S. P., Rudd, B. A. M., Hopwood, D. A., Chang, C.-J. & Floss, H. G. (1987). Biosynthesis of the antibiotic actinorhodin. Analysis of blocked mutants of *Streptomyces coelicolor*. *J Antibiot* **40**, 340–347.

Decker, H., Gaiser, S., Pelzer, S., Scheider, P., Westrich, L., Wohlleben, W. & Bechthold, A. (1996). A general approach for cloning and characterizing dNDP-glucose dehydratase genes from actinomycetes. *FEMS Microbiol Lett* **141**, 195–201.

Faust, B., Hoffmeister, D., Weitnauer, G. & 7 other authors (2000). Two new tailoring enzymes, a glycosyltransferase and an oxygenase, involved in biosynthesis of the angucycline antibiotic urdamycin A in *Streptomyces fradiae* Tü22. *Microbiology* **146**, 147–154.

Fernández-Moreno, M. A., Caballero, J. L., Hopwood, D. A. & Malpartida, F. (1991). The *act* cluster contains regulatory and antibiotic export genes, direct targets for transcriptional control by the *bld*A tRNA gene of *Streptomyces*. *Cell* **66**, 769–780.

Fernández-Moreno, M. A., Martinez, E., Boto, L., Hopwood, D. A. & Malpartida, F. (1992). Nucleotide sequence and deduced functions of a set of co-transcribed genes of *Streptomyces coelicolor* A3(2) including the polyketide synthase for the antibiotic actinorhodin. *J Biol Chem* **267**, 19278–19290.

Fernández-Moreno, M. A., Martinez, E., Caballero, J. L., Ichinose, K., Hopwood, D. A. & Malpartida, F. (1994). DNA sequence and functions of the *act*VI region of the actinorhodin biosynthetic gene cluster *of Streptomyces coelicolor* A3(2). *J Biol Chem* **269**, 24854–24863.

Gerlitz, M., Meurer, G., Wedt-Pienkowski, E., Madduri, K. & Hutchinson, C. R. (1997). Effects of the daunorubicin *dps*H 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.

Grimm, A., Madduri, K., Ali, A. & Hutchinson, C. R. (1994). Characterization of the *Streptomyces peucetius* ATCC 29050 genes encoding doxorubicin polyketide synthase. *Gene* **151**, 1–10.

Guthrie, E. P., Flaxman, C. S., White, J., Hodgson, D. A., Bibb, M. J. & Chater, K. F. (1998). A response-regulator-like activator of antibiotic synthesis from *Streptomyces coelicolor* A3(2) with an amino-terminal domain that lacks a phosphorylation pocket. *Microbiology* **144**, 727–738.

Hallam, S. E., Malpartida, F. & Hopwood, D. A. (1988). Nucleotide sequence, transcription and deduced function of a gene involved in polyketide antibiotic synthesis in *Streptomyces coelicolor*. *Gene* **74**, 305–320.

Hesketh, A., Chandra, G., Shaw, A. D., Rowland, J. J., Kell, M. J., Bibb, M. J. & Chater, K. F. (2002). Primary and secondary metabolism, and post-translational protein modifications, as portrayed by proteomic analysis of *Streptomyces coelicolor*. *Mol Microbiol* **46**, 917–932.

Hoffmeister, D., Ichinose, K., Domann, S. & 9 other authors (2000). The NDP-sugar co-substrate concentration and the enzyme expression level influence the substrate specificity of glycosyltransferases: cloning and characterization of deoxysugar biosynthetic genes of the urdamycin biosynthetic gene cluster. *Chem Biol* **7**, 821–831.

Hopwood, D. A. (1997). Genetic contributions to understanding polyketide synthases. *Chem Rev* **97**, 2465–2497.

Hopwood, D. A., Malpartida, F., Kieser, H. M., Ikeda, H., Duncan, J., Fujii, I., Rudd, B. A. M., Floss, H. G. & Ōmura, S. (1985). Production of 'hybrid' antibiotics by genetic engineering. *Nature* **314**, 642–644.

Ichinose, K., Taguchi, T., Ebizuka, Y. & Hopwood, D. A. (1998a). Biosynthetic gene clusters of benzoisochromanequinone antibiotics in *Streptomyces* spp. – identification of genes involved in post-PKS tailoring steps. *Actinomycetologica* **12**, 99–109.

Ichinose, K., Bedford, D. J., Tornus, D., Bechthold, A., Bibb, M., Revill, W. P., Floss, H. G. & Hopwood, D. A. (1998b). The granaticin biosynthetic gene cluster of *Streptomyces violaceoruber* Tü22: sequence analysis and expression in a heterologous host. *Chem Biol* **5**, 647–659.

Ichinose, K., Surti, C., Taguchi, T., Malpartida, F., Booker-Milburn, K. I., Stephenson, G. R., Ebizuka, Y. & Hopwood, D. A. (1999). Proof that the *act*VI genetic region of *Streptomyces coelicolor* A3(2) is involved in stereospecific pyran ring formation in the biosynthesis of actinorhodin. *Bioorg Med Chem Lett* **9**, 395–400.

Ichinose, K., Taguchi, T., Bedford, D. J., Ebizuka, Y. & Hopwood, D. A. (2001). Functional complementation of pyran ring formation in actinorhodin biosynthesis in *Streptomyces coelicolor* A3(2) by ketoreductase genes for granaticin biosynthesis. *J Bacteriol* **183**, 3247–3250.

Ishikawa, J., Tsuchizaki, N., Yoshida, M., Ishiyama, D. & Hotta, K. (2000). Colony PCR for detection of specific DNA sequences in actinomycetes. *Actinomycetologica* **14**, 1–5.

Kendrew, S. G., Harding, S. E., Hopwood, D. A. & Marsh, E. N. G. (1995). Identification of a flavin : NADH oxidoreductase involved in the biosynthesis of actinorhodin. *J Biol Chem* **270**, 17339–17343.

Kendrew, S. G., Hopwood, D. A. & Marsh, E. N. G. (1997). Identification of a monoxygenase from *Streptomyces coelicolor* A3(2) involved in biosynthesis of actinorhodin: purification and characterization of the recombinant enzyme. *J Bacteriol* **179**, 4305–4310.

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

Kim, I. C. & Oriel, P. J. (1995). Characterization of the *Bacillus stearothermophilus* BR219 phenol hydroxylase gene. *Appl Environ Microbiol* **61**, 1252–1256.

Künzel, E., Faust, B., Oelers, C., Weißbach, U., Bearden, D., Weitnauer, G., Westrich, L. & Bechthold, A. (1999). The inactivation of the *urdGT2* gene, which encodes a glycosyltransferase responsible for the *C*-glycosyltransfer of activated D-olivose, leads to the formation of the three novel urdamycins I, J and K. *J Am Chem Soc* **121**, 11058–11062.

Li, A. & Piel, J. (2002). A gene cluster from a marine *Streptomyces* encoding the biosynthesis of the aromatic spiroketal polyketide griseorhodin A. *Chem Biol* **9**, 1017–1026.

Lomovskaya, N. D., Chater, K. F. & Mkrtumian, N. M. (1980). Genetics and molecular biology of *Streptomyces* bacteriophages. *Microbiol Rev* **44**, 206–229.

MacNeil, D. J., Gewain, K. M., Ruby, C. L., Dezeny, G., Gibbons, P. H. & MacNeil, T. (1992). Analysis of *Streptomyces avermitilis* genes required for avermection biosynthesis utilizing a novel integration vector. *Gene* **111**, 61–68.

Madduri, K. & Hutchinson, C. R. (1995). Functional characterization and transcriptional analysis of the *dnrR1* locus, which controls daunorubicin biosynthesis in *Streptomyces peucetius*. *J Bacteriol* **177**, 1208–1215.

Marti, T., Hu, Z., Phol, N. L., Shah, A. N. & Khosla, C. (2000). Cloning, nucleotide sequence, and heterologous expression of the biosynthetic gene cluster for R1128, a non-steroidal estrogen receptor antagonist. *J Biol Chem* **275**, 33443–33448.

McDaniel, R., Ebert-Khosla, S., Hopwood, D. A. & Khosla, C. (1993). Engineered biosynthesis of novel polyketides. *Science* **262**, 1546–1550.

Mulichak, A. M., Losey, H. C., Walsh, C. T. & Garavito, R. M. (2001). Structure of the UDP-glucosyltransferase GtfB that modifies the heptapeptide aglycone in the biosynthesis of vancomycin group antibiotics. *Structure* **9**, 547–557.

Novakova, R., Bisrakova, J., Homerova, D., Rezuchova, B. & Kormanec, J. (2002). Cloning and characterization of a polyketide synthase gene cluster involved in biosynthesis of a proposed angucycline-like poleketide auricin in *Streptomyces aureofaciens* CCM 3239. *Gene* **297**, 197–208.

Okabe, T., Nomoto, K., Funabashi, H., Okuda, S., Suzuki, H. & Tanaka, N. (1985). Lactoquinomycin A, a novel anticancer antibiotic. II. Physico-chemical properties and structure assignment. *J Antibiot* **38**, 1333–1336.

Ōmura, S., Ikeda, H., Malpartida, F., Kieser, H. M. & Hopwood, D. A. (1986). Production of new hybrid antibiotics, mederhodins A and B, by a genetically engineered strain. *Antimicrob Agents Chemother* **29**, 13–19.

Pang, C.-H., Shiiyama, M., Ikeda, H., Tanaka, H. & Ōmura, S. (1994). Cosmid vector for cloning and analysis of *Streptomyces* DNA. *Actinomycetologica* **8**, 21–25.

Rix, U., Fischer, C., Remsing, L. L. & Rohr, J. (2002). Modification of post-PKS tailoring steps through combinatorial biosynthesis. *Nat Prod Rep* **19**, 542–580.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sciara, G., Kendrew, S. G., Miele, A. E., Marsh, N. G., Federici, L., Malatesta, F., Schimperna, G., Savino, C. & Vallone, B. (2003). The structure of ActVA-Orf6, a novel type of monooxygenase involved in actinorhodin biosynthesis. *EMBO J* **22**, 205–215.

Sherman, D. H., Malpartida, F., Bibb, M. J., Kieser, H. M., Bibb, M. J. & Hopwood, D. A. (1989). Structure and deduced function of the granaticin-producing polyketide synthase gene cluster of *Streptomyces violaceoruber* Tü22. *EMBO J* **8**, 2717–2725.

Strohl, W. R. (1992). Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. *Nucleic Acids Res* **20**, 961–974.

Taguchi, T., Itou, K., Ebizuka, Y., Malpartida, F., Hopwood, D. A., Surti, C. M., Booker-Milburn, K. I., Stephenson, G. R. & Ichinose, K. (2000a). Chemical characterization of disruptants of the *Streptomyces coelicolor* A3(2) *act*VI genes involved in actinorhodin biosynthesis. *J Antibiot* **53**, 144–152.

Taguchi, T., Ebizuka, Y., Hopwood, D. A. & Ichinose, K. (2000b). Identification of a novel shunt product produced by a disruptant of the *act*VI-ORFA gene involved in the biosynthesis of actinorhodin in *Streptomyces coelicolor* A3(2). *Tetrahedron Lett* **41**, 5253–5356.

Taguchi, T., Ebizuka, Y., Hopwood, D. A. & Ichinose, K. (2001). A new mode of stereochemical control revealed by analysis of the biosynthesis of dihydrogranaticin in *Streptomyces violaceoruber* Tü22. *J Am Chem Soc* **123**, 11376–11380.

Takano, S., Hasuda, K., Ito, A., Koide, Y., Ishii, F., Haneda, I., Chihara, S. & Koyama, Y. (1976). A new antibiotic, medermycin. *J Antibiot* **29**, 765–768.

Tanaka, N., Okabe, T., Isono, F., Kashiwagi, M., Nomoto, K., Takahashi, M., Shimazu, A. & Nishimura, T. (1985). Lactoquinomycin, a novel anticancer antibiotic. I. Taxonomy, isolation, and biological activity. *J Antibiot* **38**, 1327–1332.

Tang, L. & McDaniel, R. (2001). Construction of desosamine containing polyketide libraries using a glycosyltransferase with broad substrate specificity. *Chem Biol* **8**, 547–555.

Tatsuta, K., Ozeki, H., Yamaguchi, M., Tanaka, M. & Okui, T. (1990). Enantioselective total synthesis of medermycin (lactoquinomycin). *Tetrahedron Lett* **31**, 5495–5498.

Tornus, D. & Floss, H. G. (2001). Identification of four genes from the granaticin biosynthetic gene cluster of *Streptomyces violaceoruber* Tü22 involved in the biosynthesis of L-rhodinose. *J Antibiot* **54**, 91–101.

Trefzer, A., Salas, J. A. & Bechthold, A. (1999). Genes and enzymes involved in deoxysugar biosynthesis in bacteria. *Nat Prod Rep* **16**, 283–299.

Ünligil, U. M. & Rini, J. M. (2000). Glycosyltransferase structure and mechanism. *Curr Opin Struct Biol* **10**, 510–517.

Wang, L., McVey, J. & Vining, L. C. (2001). Cloning and functional analysis of a phosphopantetheinyl transferase superfamily gene associated with jadomycin biosynthesis in *Streptomyces venezuelae* ISP5230. *Microbiology* **147**, 1535–1545.

Westrich, L., Domann, S., Faust, B., Bedford, D., Hopwood, D. A. & Bechthold, A. (1999). Cloning and characterization of a gene cluster

from *Streptomyces cyanogenus* S136 probably involved in landomycin biosynthesis. *FEMS Microbiol Lett* **170**, 381–387.

Wietzorrek, A. & Bibb, M. J. (1997). A novel family of proteins that regulates antibiotic production in streptomycetes appear to contain an OmpR-like DNA-binding fold. *Mol Microbiol* **25**, 1177–1184.

Williamson, R. T., McDonald, L. A., Barbieri, L. R. & Carter, G. T. (2002). In support of the original medermycin/lactoquinomycin A structure. *Org Lett* **4**, 4659–4662.

Ylihonko, K., Tuikkanen, J., Jussila, S., Cong, L. & Mäntsälä, P. (1996). A gene cluster involved in nogalamycin biosynthesis from *Streptomyces nogalater*: sequence analysis and complementation of early-block mutants in the anthracycline pathway. *Mol Gen Genet* **251**, 113–120.

Yu, T.-W., Bibb, M. J., Revill, W. P. & Hopwood, D. A. (1994). Cloning, sequencing, and analysis of the griseusin polyketide synthase gene cluster from *Streptomyces griseus*. *J Bacteriol* **176**, 2827–2634.