

Cloning and Characterization of Genes Encoded in dTDP-*D*-mycaminose Biosynthetic Pathway from a Midecamycin-producing Strain, *Streptomyces mycarofaciens*

Lina CONG^{1*} and Wolfgang PIEPERSBERG^{2*}

¹ College of Biology and Food Technology, Dalian Institute of Light Industry, Dalian 116034, China;

² Institute of Chemical Microbiology, Bergische University, D-42097 Wuppertal, Germany

Abstract Two subclusters from *Streptomyces mycarofaciens*, a midecamycin producer, were cloned and partially sequenced. One region was located at the 5' end of the mid polyketide synthase (PKS) genes and contained the genes *midA*, *midB* and *midC*. The other region was at the 3' end of the PKS genes and contained *midK*, *midI* and *midH*. Analysis of the nucleotide sequence revealed that these genes encode dTDP-glucose synthase (*midA*), dTDP-glucose dehydratase (*midB*), aminotransferase (*midC*), methyltransferase (*midK*), glycosyltransferase (*midI*) and an assistant gene (*midH*). All of these genes are involved in the biosynthesis of dTDP-*D*-mycaminose, the first deoxysugar of midecamycin, and in transferring the mycaminose to the midecamycin aglycone in *S. mycarofaciens*. Similar to gene pairs *desVIII/desVII* in *S. venezuelae* and *tylMIII/tylMII* in *S. fradiae*, the product of *midH* probably functions as an auxiliary protein required by the MidI protein for efficient glycosyltransfer in midecamycin biosynthesis.

Key words deoxysugar; dTDP-*D*-mycaminose; midecamycin; *Streptomyces mycarofaciens*

Macrolides constitute a class of antibiotics that contain a macrocyclic lactone ring composed of a polyketide-derived backbone to which one, two or three sugars are commonly attached. They are produced as secondary metabolites by mycelium-forming soil bacteria from the order *Actinomycetales*; the majority are from members of the genera *Streptomyces*, *Micromonospora* and *Saccharopolyspora*. Macrolide antibiotics are widely used as anti-infective, immunosuppressive, insecticidal, and parasiticidal agents in the clinic or for agricultural purposes. The known mechanism of the biological function of the main group of classical macrolides (erythromycin, tylosin and so on) is to bind to the peptidyltransferase center of the 50S subunit of the bacterial ribosome, thereby inhibit-

ing bacterial protein synthesis [3].

Streptomyces mycarofaciens produces midecamycin, a 16-membered macrolide (**Fig. 1**). During the biosynthesis of midecamycin, two deoxyhexose sugars, mycaminose and mycarose, are added to the polyketide lactone ring in sequential order [4]. The genetic organization of macrolide antibiotics has shown a common feature: the genes of polyketide synthase (PKS) involved in the lactone ring formation are clustered in the center of the chromosome with flanking genes of sugar biosynthesis on both sides [5–8]. To date, the PKS gene clusters have been cloned and characterized in the biosynthesis of several macrolide antibiotics [5–10]. However, the genetic analysis of sugar moieties that affect the bioactivity of macrolides has been studied at a much lower level.

Here we report the cloning and characterization of genes encoded in the dTDP-*D*-mycaminose biosynthetic pathway from a midecamycin-producing strain, *S. mycarofaciens*.

Received: November 11, 2006

Accepted: December 25, 2006

This study was supported by a grant from the Fifth Framework Program (FP5) of the European Community for Research, Technological Development and Demonstration Activities

*Corresponding authors:

Lina CONG: Tel, 86-411-86324521; E-mail, congln@dlili.edu.cn

Wolfgang PIEPERSBERG: Tel, 49-202-4392521; E-mail, piepersb@uni-wuppertal.de

DOI: 10.1111/j.1745-7270.2007.00265.x

Quertermous [17].

DNA sequencing and analysis

DNA sequencing was carried out with cosmid by standard shotgun cloning to obtain at least 4-fold coverage. Primer walking was used to close the gaps. The sequence was assembled using the Sequencher software package (Gene Codes, Ann Arbor, USA) and analyzed with MacVector (Accelrys, San Diego, USA) and the BLAST server of the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/BLAST>).

Nucleotide sequence accession number

The nucleotide sequence reported here has been submitted to GenBank with accession numbers DQ672716, DQ672717, DQ672718, DQ672719, and DQ672720.

Results

Identification of pathway-specific gene probes

To identify the specific genes in midecamycin biosynthesis, the DNA fragments *midB'* (303 bp) and *midI'* (420 bp) were amplified by PCR from the genomic DNA of *S. mycarofaciens*. Sequencing analysis indicated that *MidB'* encodes dTDP-glucose 4,6-dehydratase and *MidI'* encodes a glycosyltransferase. The evidence is that the product of *midB'* in a 101 amino acid overlap is 76% iden-

tical to TylAII in *S. fradiae* and 70% identical to StrE in *S. griseus*. Both proteins TylAII and StrE are the putative dTDP-glucose 4,6-dehydratase [15,18]. The comparison of the *midI'* product to other known glycosyltransferases revealed that *MidI'* in a 142 amino acid overlap is 56% identical to TylMII and 53% identical to EryCIII in *Saccharopolyspora erythraea* [12,16]. Analysis of several previously identified macrolide deoxysugar clusters suggested that these two genes are separately located on both sides of the PKS gene cluster [9,15–18]. Therefore, the obtained gene fragments of *midB'* and *midI'* were used as pathway-specific probes to screen the genomic DNA library of *S. mycarofaciens*.

Cloning and sequencing of the partial *mid* biosynthetic gene cluster

Initial colony hybridization analysis with the *midB'* probe against the genomic DNA library of *S. mycarofaciens* revealed two positive cosmids, cosM1 (30.2 kb) and cosM2 (22.6 kb) [Fig. 2(A)]. Both cosmids are overlapped 12.4 kb, and sequencing showed that the extending region of cosM2 contains the 5' end sequence of the *mid* PKS genes. By chromosomal walking using a probe of a 2.3 kb DNA fragment close to the end of cosM1 to screen the library, a positive cosmid cosM4 (29.3 kb) was obtained that extended away from cosM1 in a region of approximately 18 kb. Using the *midI'* probe to screen the genomic library, a positive cosmid cosM3 (26.3 kb) was obtained. One end of cosM3 contains the 3' end sequence

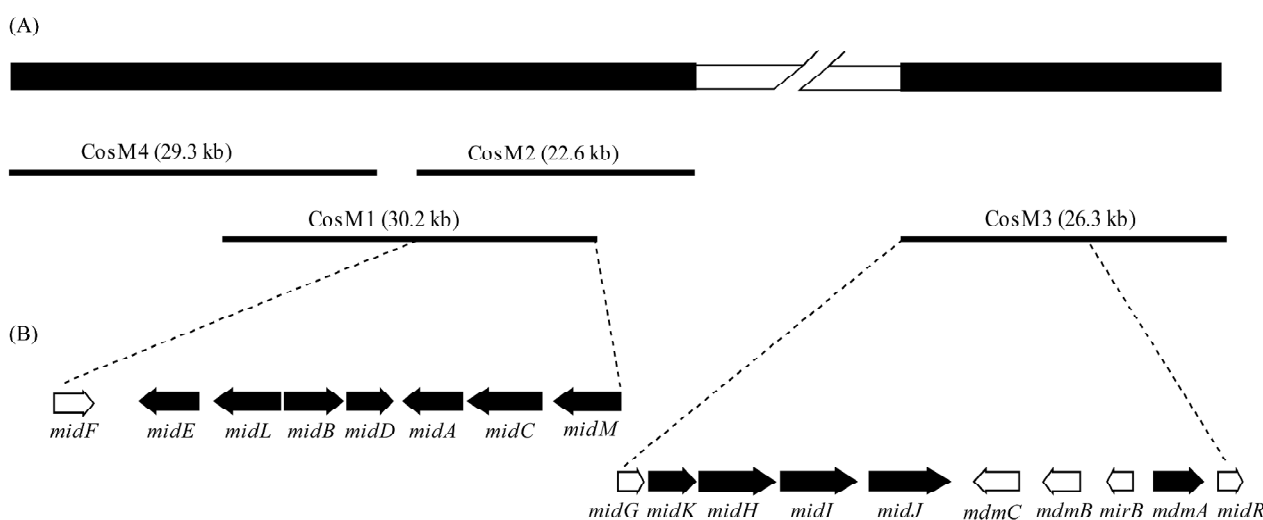


Fig. 2 Organization of the partial gene cluster in midecamycin biosynthesis

(A) Schematic representation of cosmid clone location. The black bars represent the region of *Streptomyces mycarofaciens* chromosomal DNA found in existing cosmids. The white bar indicates the unresearched region of the chromosome with a break shown as a double slant. (B) Genes identified within the analyzed sequence (cosM1 and cosM3). The arrowheads indicate transcription direction. Black arrow bars represent open reading frame genes and white arrow bars indicate incomplete genes.

of the *mid* PKS genes. Although finding the *mid* PKS genes was not the target of this study, it could facilitate the identification of the overall organization and orientation of the gene cluster and the genes encoding 6-deoxysugar biosynthetic enzymes. Therefore, cosmids cosM1, cosM2 and cosM4 should be upstream of the *mid* PKS genes, whereas cosmid cosM3 should be located downstream [Fig. 2(A)].

Extensive sequencing was done from cosmids cosM1 and cosM3. Computer-assisted analysis of the sequenced regions led to identification of the genes shown in Fig. 2(B) and listed in Table 1. They include genes for mycaminoses biosynthesis and attachment as well as genes for PKS, resistance, regulation and chain modification.

Analysis of genes involved in mycaminoses biosynthesis and attachment

The *midA*, *midB* and *midC* genes were found in a subcluster upstream of the *mid* PKS genes. The deduced product of *midA* showed significant sequence identity to putative dTDP-glucose synthases from several streptomycetes [9,19,20]. The complete *midB* gene includes the DNA fragment of the *midB'* probe. The deduced product

of *midB* is highly similar in sequence to dTDP-glucose-4,6-dehydratases from several deduced gene products in macrolide antibiotics [18,19,21]. The deduced product of *midC* shows significant sequence similarity with putative aminotransferases from various deduced gene products in different antibiotic biosynthetic pathways [18,20,22]. The BLASTP analysis of the MidA, MidB and MidC proteins is shown in Table 2.

The *midK*, *midI* and *midH* genes were found in a subcluster downstream of the *mid* PKS genes. The deduced product of *midK* shows significant sequence similarity to a family of enzymes proposed to function as S-adenosylmethionine-dependent methyltransferases [16,20,23]. The complete *midI* gene includes the DNA fragment of the *midI'* probe. The deduced product of *midI* shows convincing end-to-end sequence similarity to the known glycosyltransferases of several macrolide antibiotics [9,16,20]. In database searches, the deduced product of *midH* shows sequence similarity to *TylMIII* in *S. fradiae*, *DnrQ*

Table 1 Deduced functions of the putative genes identified in the midcamycin biosynthetic gene cluster

Gene designation	aa (n)	Proposed function
<i>midA</i>	303	dTDP-glucose synthase
<i>midB</i>	326	dTDP-glucose-4,6-dehydratase
<i>midC</i>	398	Aminotransferase
<i>midK</i>	249	Methyltransferase
<i>midH</i>	414	Auxiliary protein
<i>midI</i>	421	Glycosyltransferase
<i>midM</i>	372	Cytochrome P-450
<i>midD</i>	264	Thioesterase
<i>midL</i>	372	dTDP-4-keto-6-deoxyglucose 2,3-dehydratase
<i>midJ</i>	448	Crotonyl CoA reductase
<i>midG</i> (incomplete)	133	Midcamycin PKS
<i>midF</i> (incomplete)	228	Regulation
<i>midR</i> (incomplete)	140	Regulation
<i>mdmA</i>	271	Resistance
<i>mirB</i> (incomplete)	135	Resistance
<i>midE</i>	338	4 ^{''} -O-propionyl transferase
<i>mdmB</i> (incomplete)	213	3-O-acyltransferase
<i>mdmC</i> (incomplete)	252	O-methyltransferase

aa, amino acid.

Table 2 BLASTP analysis of gene products involved in mycaminoses biosynthesis and attachment from *Streptomyces mycarofaciens*

Gene product	Best BLASTP matches (%identity/protein/organism)	Accession No.
MidA	70/ChmAI/ <i>S. bikiniensis</i>	AY509120
	66/AveBIII/ <i>S. avermitilis</i>	NC_003155
	60/DesIII/ <i>S. venezuelae</i>	AF079762
MidB	72/AprE/ <i>S. tenebrarius</i>	AAG18457
	64/TylAII/ <i>S. fradiae</i>	U08223
MidC	64/AveBII/ <i>S. avermitilis</i>	NC_003155
	58/OleN2/ <i>S. antibioticus</i>	AF055579
	56/TylB/ <i>S. fradiae</i>	U08223
MidK	54/DesV/ <i>S. venezuelae</i>	AF079762
	57/TylMI/ <i>S. fradiae</i>	X81885
	54/OleM1/ <i>S. antibioticus</i>	AJ002638
MidI	53/DesVI/ <i>S. venezuelae</i>	AF079762
	59/ChmCIII/ <i>S. bikiniensis</i>	AY509120
	57/TylMII/ <i>S. fradiae</i>	X81885
MidH	54/DesVII/ <i>S. venezuelae</i>	AF079762
	35/TylMIII/ <i>S. fradiae</i>	X81885
	32/DnrQ/ <i>S. peucetius</i>	L47164
	30/DesVIII/ <i>S. venezuelae</i>	AF079762

MidA, dTDP-glucose synthase; MidB, dTDP-glucose-4,6-dehydratase; MidC, aminotransferase; MidK, methyltransferase; MidI, glycosyltransferase; MidH, auxiliary protein.

in *S. peuceitius*, and *DesVIII* in *S. venezuelae* [16,20,24]. **Table 2** shows the results of BLASTP analysis of the MidK, MidI and MidH proteins.

Discussion

Several studies have shown that various 6-deoxyhexoses, present in a range of antibiotic molecules, are made from *D*-glucose-1-phosphate by way of dTDP-glucose and dTDP-4-keto-6-deoxy-glucose before the pathway diverges [25,26]. Based on sequence analysis and comparison, the gene *midA* as dTDP-glucose synthase and *midB* as dTDP-glucose-4,6-dehydratase are supposed to be responsible for the early steps in the midecamycin biosynthetic gene cluster. In particular, a novel gene order of *midA* and *midB* was found in this study whereby both genes are convergent and separated by *midD*. In most other antibiotics *midA* and *midB* are usually co-directional neighbors [18,20].

The MidC protein deduced as an aminotransferase might catalyze the conversion of dTDP-3-keto-6-deoxyglucose to dTDP-3-amino-6-deoxyglucose during dTDP-*D*-mycaminose biosynthesis. The aminotransferase enzyme is thought to be dependent on pyridoxal phosphate as a co-factor [27]. The sequence similarities between MidC and other macrolide aminotransferases are most apparent in a region of the protein that contains the conserved lysine

residue, which is supposed to be the attachment site for pyridoxal phosphate (data not shown). The MidK protein displayed as a methyltransferase might act on the amino group of dTDP-3-amino-6-deoxyglucose during mycaminose biosynthesis, perhaps by introducing two methyl groups at that site. The sequence alignment of the MidK protein with other macrolide methyltransferases shows that they possess three of the consensus sequence motifs typical of methyltransferases that use *S*-adenosylmethionine as a co-substrate [28] (data not shown).

Based on sequence analysis and comparison, the MidI protein as a glycosyltransferase is believed to be responsible for attachment of mycaminose to midecamycin lactone. A sequence alignment of MidI with other known glycosyltransferases shows that all these proteins retain a characteristic motif, P-NVR-VDFVPL-ALLP-C---VHHGG-GT--TA--HG-P, present in UDP-glycosyltransferases [29] (data not shown). As the gene order of *midH* directly upstream from *midI* is the same as the gene pairs *tylMIII/tylMII* in *S. fradiae* and *desVIII/desVII* in *S. venezuelae* [1,2], and the sequence of MidH showed similarity with the TylMIII and DesVIII, the product of *midH* probably functions as an auxiliary protein required by the MidI protein for efficient glycosyltransfer in midecamycin biosynthesis.

As shown in **Fig. 3**, a biosynthetic route from glucose-1-phosphate to dTDP-*D*-mycaminose, and the sugar at-

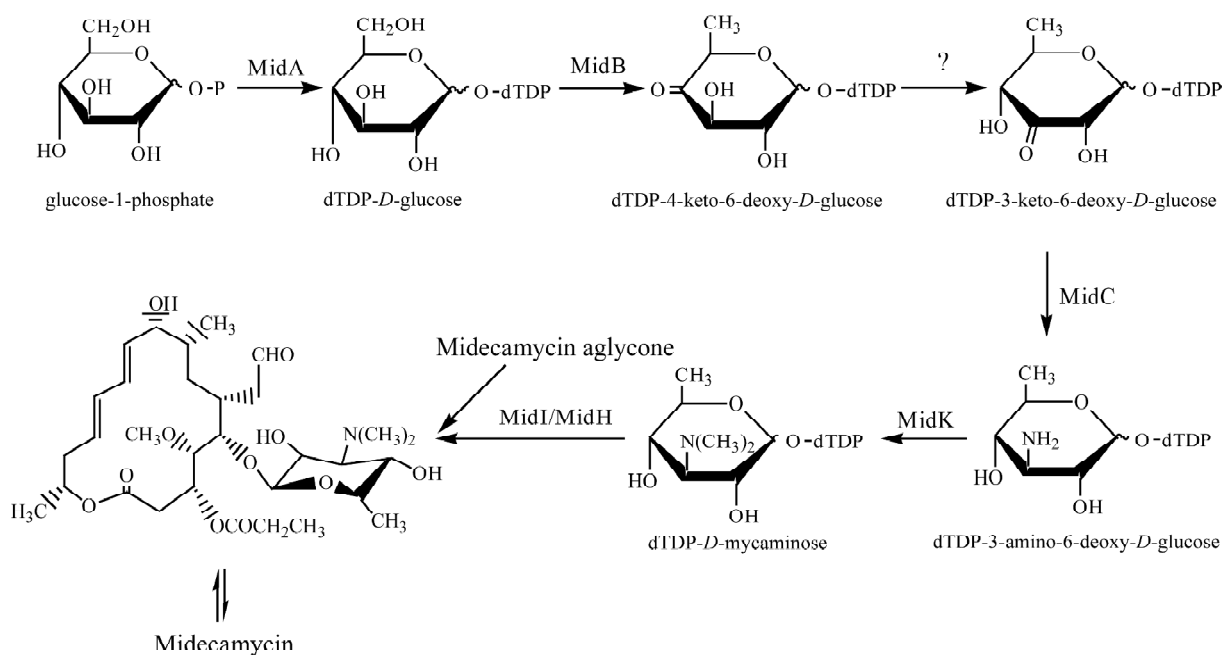


Fig. 3 Proposed biosynthetic route to dTDP-*D*-mycaminose in *Streptomyces mycarofaciens* and related *mid*-encoding proteins

tachment to midecamycin lactone in *S. mycarofaciens* is proposed. In this pathway, a gene encoding 3,4-isomerase, which is responsible for converting dTDP-4-keto-6-deoxy-*D*-glucose to dTDP-3-keto-6-deoxy-*D*-glucose, is still missing. We have since extended another 20 kb of sequencing region on both sides of CosM1 and CosM3 to identify the missing gene and genes involved in mycarose biosynthesis, the second deoxysugar of midecamycin. Based on the present study, it is concluded that the proposed biosynthetic pathway is very similar to the parallel genes identified in the tylosin biosynthetic gene cluster of *S. fradiae*.

Acknowledgement

We thank Prof. John SMITH for help in preparing the manuscript.

References

- Borisova SA, Zhao L, Melancon III CE, Kao CL, Liu HW. Characterization of the glycosyltransferase activity of desVII: Analysis of and implications for the biosynthesis of macrolide antibiotics. *J Am Chem Soc* 2004, 126: 6534–6535
- Melancon CE 3rd, Takahashi H, Liu HW. Characterization of *tylM3/tylM2* and *mydC/mycB* pairs required for efficient glycosyltransfer in macrolide antibiotic biosynthesis. *J Am Chem Soc* 2004, 126: 16726–16727
- Gale EF, Cundliffe E, Reynolds PE, Richmond MH, Waring MJ. *The Molecular Basis of Antibiotic Action*. 2nd edn. London: Wiley 1981
- Neu HC. *In vitro* activity of midecamycin, a new macrolide antibiotic. *Antimicrob Agents Chemother* 1983, 24: 443–444
- Reeves AR, English RS, Lampel JS, Post DA, Vanden Boom TJ. Transcriptional organization of the erythromycin biosynthetic gene cluster of *Saccharopolyspora erythraea*. *J Bacteriol* 1999, 181: 7098–7106
- Cundliffe E, Bate N, Butler A, Fish S, Gandechea A, Merson-Davies L. The tylosin-biosynthetic genes of *Streptomyces fradiae*. *Antonie Van Leeuwenhoek* 2001, 79: 229–234
- Campelo AB, Gil JA. The candicidin gene cluster from *Streptomyces griseus* IMRU 3570. *Microbiology* 2002, 148: 51–59
- Haydock SF, Appleyard AN, Mironenko T, Lester J, Scott N, Leadlay PF. Organization of the biosynthetic gene cluster for the macrolide concanamycin A in *Streptomyces neyagawaensis* ATCC 27449. *Microbiology* 2005, 151: 3161–3169
- Ward SL, Hu Z, Schirmer A, Reid R, Revill WP, Reeves CD, Petrakovsky OV *et al.* Chalcomycin biosynthesis gene cluster from *Streptomyces bikiniensis*: Novel features of an unusual ketolide produced through expression of the *chm* polyketide synthase in *Streptomyces fradiae*. *Antimicrob Agents Chemother* 2004, 48: 4703–4712
- Anzai Y, Saito N, Tanaka M, Kinoshita K, Koyama Y, Kato F. Organization of the biosynthetic gene cluster for the polyketide macrolide mycinamicin in *Micromonospora griseorubida*. *FEMS Microbiol Lett* 2003, 218: 135–141
- Sambrook J, Russel DW. *Molecular Cloning: A Laboratory Manual*, 3rd edn. New York: Cold Spring Harbor Laboratory Press 2001
- Salah-Bey K, Doumith M, Michel JM, Haydock S, Cortes J, Leadlay PF, Raynal MC. Targeted gene inactivation for the elucidation of deoxysugar biosynthesis in the erythromycin producer *Saccharopolyspora erythraea*. *Mol Gen Genet* 1998, 257: 542–553
- Kakinuma S, Takada Y, Ikeda H, Tanaka H, Ōmura S, Hopwood DA. Cloning of large DNA fragments, which hybridize with actinorhodin biosynthesis genes, from kalafungin and nanaomycin A methyl ester producers and identification of genes for kalafungin biosynthesis of the kalafungin producer. *J Antibiotics* 1991, 44: 995–1005
- Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA. *Practical Streptomyces Genetics*. Norwich: The John Innes Foundation 2000
- Pissowotzki K, Mansouri K, Piepersberg W. Genetics of streptomycin production in *Streptomyces griseus*: Molecular structure and putative function of genes strELMB2N. *Mol Gen Genet* 1991, 231: 113–123
- Gandechea AR, Large S, Cundliffe E. Analysis of four tylosin biosynthetic genes from the *tylLM* region of the *Streptomyces fradiae* genome. *Gene* 1997, 184: 197–203
- Weis JH, Quertermous T. Construction of recombinant DNA libraries. In: Ausubel FM, Brent R, Kingston RE, Moore DD, Smith JA, Seidman JG, Struhl K eds. *Current Protocols in Molecular Biology*. 4th edn. New York: John Wiley and Sons 1987
- Merson-Davies LA, Cundliffe E. Analysis of five tylosin biosynthetic genes from the *tylBA* region of the *Streptomyces fradiae* genome. *Mol Microbiol* 1994, 13: 349–355
- Ikeda H, Ishikawa J, Hanamoto A, Shinose M, Kikuchi H, Shiba T, Sakaki Y *et al.* Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat Biotechnol* 2003, 21: 526–531
- Xue Y, Zhao L, Liu HW, Sherman DH. A gene cluster for macrolide antibiotic biosynthesis in *Streptomyces venezuelae*: Architecture of metabolic diversity. *Proc Natl Acad Sci USA* 1998, 95: 12111–12116
- Li T, Shang G, Xia H, Wang Y. Cloning of the sugar related biosynthesis gene cluster from *Streptomyces tenebrarius* H6. *Sheng Wu Gong Cheng Xue Bao* 2001, 17: 329–331
- Quiros LM, Aguirrezabalaga I, Olano C, Mendez C, Salas JA. Two glycosyltransferases and a glycosidase are involved in oleandomycin modification during its biosynthesis by *Streptomyces antibioticus*. *Mol Microbiol* 1998, 28: 1177–1185
- Olano C, Rodriguez A, Michel J, Mendez C, Raynal M, Salas J. Analysis of a *Streptomyces antibioticus* chromosomal region involved in oleandomycin biosynthesis, which encodes two glycosyltransferases responsible for glycosylation of the macrolactone ring. *Mol Gen Genet* 1998, 259: 299–308
- Otten SL, Liu X, Ferguson F, Hutchinson CR. Cloning and characterization of the *Streptomyces peucetius dnrQS* genes encoding a daunosamine biosynthesis enzyme and a glycosyl transferase involved in daunorubicin biosynthesis. *J Bacteriol* 1995, 177: 6688–6729
- Piepersberg W. Pathway engineering in secondary metabolite-producing actinomycetes. *Crit Rev Biotechnol* 1994, 14: 251–285
- Salas JA, Mendez C. Biosynthesis pathways for deoxysugars in antibiotic-producing actinomycetes: Isolation, characterization and generation of novel glycosylated derivatives. *J Mol Microbiol Biotechnol* 2005, 9: 77–85
- Pascarella S, Bossa F. Similarity between pyridoxal/pyridoxamine phosphate dependent enzymes involved in dideoxy and deoxyaminosugar biosynthesis and other pyridoxal phosphate enzymes. *Protein Sci* 1994, 3: 701–705

- 28 Kagan RM, Clarke S. Widespread occurrence of three sequence motifs in diverse S-adenosylmethionine-dependent methyltransferases suggests a common structure for these enzymes. *Arch Biochem Biophys* 1994, 310: 417–427
- 29 Jenkins G, Cundliffe E. Cloning and characterization of two genes from *Streptomyces lividans* that confer inducible resistance to lincomycin and macrolide antibiotics. *Gene* 1991, 108: 55–62

Edited by
Stephen F. HAYDOCK