Molecular characterization and analysis of the biosynthetic gene cluster for the antitumor antibiotic mitomycin C from *Streptomyces lavendulae* NRRL 2564

Yingqing Mao, Mustafa Varoglu and David H Sherman

Background: The mitomycins are natural products that contain a variety of functional groups, including aminobenzoquinone- and aziridine-ring systems. Mitomycin C (MC) was the first recognized bioreductive alkylating agent, and has been widely used clinically for antitumor therapy. Precursor-feeding studies showed that MC is derived from 3-amino-5-hydroxybenzoic acid (AHBA), D-glucosamine, L-methionine and carbamoyl phosphate. A genetically linked AHBA biosynthetic gene and MC resistance genes were identified previously in the MC producer *Streptomyces lavendulae* NRRL 2564. We set out to identify other genes involved in MC biosynthesis.

Results: A cluster of 47 genes spanning 55 kilobases of *S. lavendulae* DNA governs MC biosynthesis. Fourteen of 22 disruption mutants did not express or overexpressed MC. Seven gene products probably assemble the AHBA intermediate through a variant of the shikimate pathway. The gene encoding the first presumed enzyme in AHBA biosynthesis is not, however, linked within the MC cluster. Candidate genes for mitosane nucleus formation and functionalization were identified. A putative MC translocase was identified that comprises a novel drug-binding and export system, which confers cellular self-protection on *S. lavendulae*. Two regulatory genes were also identified.

Conclusions: The overall architecture of the MC biosynthetic gene cluster in *S. lavendulae* has been determined. Targeted manipulation of a putative MC pathway regulator led to a substantial increase in drug production. The cloned genes should help elucidate the molecular basis for creation of the mitosane ring system, as well efforts to engineer the biosynthesis of novel natural products.

Address: University of Minnesota, Department of Microbiology, Biological Process Technology Institute, 1460 Mayo Memorial Building, Box 196 UFHC, 420 Delaware Street S.E., Minneapolis, MN 55455, USA.

Correspondence: David H. Sherman
E-mail: david-s@biosci.umn.edu

Key words: antibiotic biosynthesis, antibiotic resistance, gene disruption, mitosane, shikimate pathway

Received: **15 January 1999**
Accepted: **5 February 1999**

Published: **23 March 1999**

Chemistry & Biology April 1999, **6**:251–263
http://biomednet.com/elecref/1074552100600251

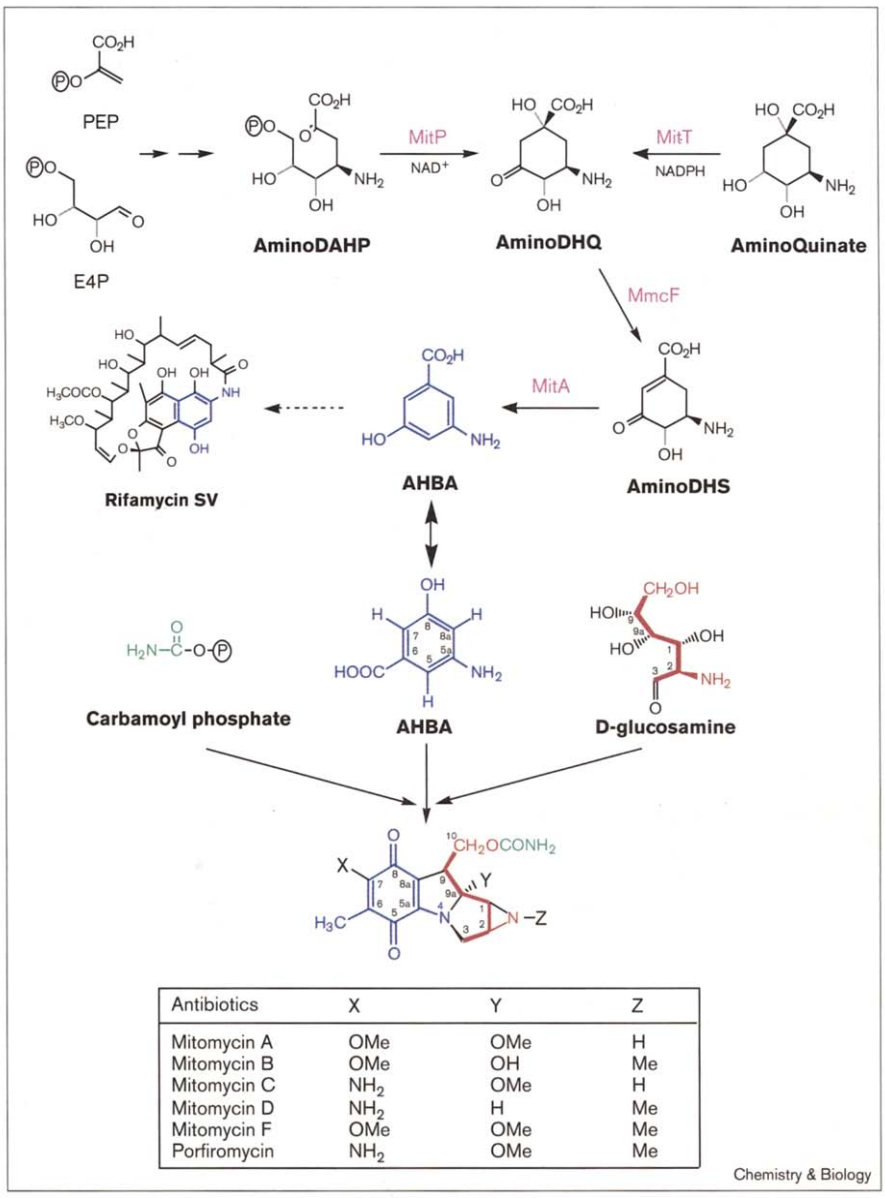
Introduction

Since its discovery and demonstration of anticancer activity in the 1960s, many aspects of the chemistry and biology of mitomycin C (MC) have been investigated. This has provided detailed information on the unprecedented molecular mechanism, unique biological and pharmacological properties, drug resistance and bioactive analogues of MC [1,2]. MC is regarded as the prototype natural-product alkylating agent whose activity is dependent on reductive activation (either chemically, such as low pH, or enzymatically, such as DT-diaphorase or NADH cytochrome c reductase) [3,4]. Activated MC cross-links double-stranded DNA, which in turn induces diverse biological effects including selective inhibition of DNA synthesis, mutagenesis, induction of DNA repair (SOS response), sister-chromatid exchange, signal transduction and induction of apoptosis [5]. Tumor hypoxia and the increased expression of bioreductive enzymes in malignant cells create a selective environment for drug activation and make MC an attractive agent for antitumor therapy [6]. MC remains a clinically important component of combination cancer chemotherapy and radiotherapy of solid tumors [7].

In addition to its biological and pharmacological importance, MC is prominent because its molecular mechanism represents a model for structurally related antitumor antibiotics porfiromycin [8], mitiromycin [9], FR66979 [10], FR900482 [11], FK973 [12], and FK317 [13], as well as structurally unrelated bioreductive agents such as EO9 [14] and tirapazamine [15]. Numerous MC derivatives have been synthesized and tested for enhanced activities, including the recently identified selective protein tyrosine kinase inhibitor, 1a-docosahexaenoyl MC [16,17].

MC is derived, in part, from 3-amino-5-hydroxybenzoic acid (AHBA), a precursor that is also required for biosynthesis of ansamycin antibiotics, including rifamycin [18,19]. Incorporation experiments with radiolabeled precursors have demonstrated that the mitosane core of MC is derived from the junction of AHBA and D-glucosamine [20,21] (Figure 1). The *O*- and *N*- (but not C-) methyl

Figure 1



Antibiotics	X	Y	Z
Mitomycin A	OMe	OMe	H
Mitomycin B	OMe	OH	Me
Mitomycin C	NH_2	OMe	H
Mitomycin D	NH_2	H	Me
Mitomycin F	OMe	OMe	Me
Porfiromycin	NH_2	OMe	Me

Mitomycins and their hypothetical biosynthetic pathway.

groups were shown to be derived from L-methionine, and the C-10 carbamoyl group originates from L-arginine or L-citrulline [22–24].

MC has been an important synthetic target, in addition to being the subject of numerous studies to understand its mechanism of activation and DNA alkylation. Little is known, however, about the details of its convergent assembly from AHBA and D-glucosamine in *Streptomyces lavendulae*. Understanding the derivation of AHBA is important to determine if *de novo* biosynthesis of MC is related to the primary metabolic shikimate pathway, an important route in microorganisms and plants for aromatic amino-acid biosynthesis [25]. Another intriguing question involves the details of cellular self-protection in *S. lavendulae* because the preferred MC alkylation sites in DNA are guanine and cytosine, and MC-induced cell death can result from a single cross-link per genome [26].

To address the above questions, the MC AHBA synthase (*mitA*) and two MC resistance genes (*mcrA* and *mrd*) were cloned from *S. lavendulae*, and *mrd* was found to be closely linked to *mitA* [27–29]. We therefore set out to clone and sequence the large cluster of genes adjoining *mitA* that encode MC structural, regulatory and resistance components. This work provides the first direct genetic information about the biosynthesis of this clinically important antitumor antibiotic, and provides the

foundation for dissecting further the function of individual genes and enzymes involved in assembly of mitosane natural products.

Results

Identification of the mitomycin biosynthetic gene cluster (MC cluster)

The MC cluster was identified by linkage of a cosmid clone containing *mrd* and a gene (*mitA*) that hybridized with the *rifK* gene encoding the rifamycin AHBA synthase [30] from *Amycolatopsis mediterranei*. *mitA* was subsequently shown to be essential for MC biosynthesis because genetic disruption of the chromosomal copy blocked MC production, and could be complemented with exogenous AHBA [29]. Linkage of *mitA* with one of the MC resistance genes (*mrd*) implied that the corresponding biosynthetic genes were adjacent to *mitA*. In this work, cosmid walking was used to obtain overlapping DNA fragments spanning more than 120 kilobases of the *S. lavendulae* chromosome adjacent to *mitA*. Subsequent nucleotide sequence analysis included 55 kb of contiguous DNA, revealing 47 genes involved in MC assembly, regulation and resistance (Tables 1 and 2; Figure 2).

mitT defines the left-hand boundary of the MC cluster

Nucleotide-sequence analysis extended 30 kb downstream of *mitA* and revealed a set of genes corresponding to a type I polyketide synthase (PKS; *orf9*, *orf8*) and a thioesterase (TEII; *orf7*). MC is not derived from the polyketide pathway — an *orf8* disruption mutant showed normal MC production as expected (Table 2). Approximately 20 kb downstream of *mitA*, two genes (*mitT* and *mitS*) encoding a putative aminoquinate dehydrogenase and a kinase, respectively, were located. Both are believed to be involved in AHBA biosynthesis because their equivalents are also present in the rifamycin biosynthetic gene cluster (*rif* cluster) [31]. Whether the six genes between *orf7* and *mitT* are involved in MC biosynthesis remained unclear, however, because the two putative hydroxylases (*orf3*, *orf4*) and the candidate activator gene (*orf1*) could conceivably play a role in MC production. Both *orf3* and *orf4* are predicted to encode cytochrome P450 monooxygenases with Orf4 most similar to OleP and RapN (50% identity, 63% similarity) involved in oleandomycin and rapamycin biosynthesis, respectively [32,33]. Orf3 shows a high degree of similarity to cytochrome P450 105C1 (49% identity, 64% similarity) in *Streptomyces* sp. and cytochrome P450-SU2 in *Streptomyces griseolus* [34,35].

Database analysis revealed that Orf1 belonged to the ActII-ORF4, RedD, DnrI and CcaR family of *Streptomyces* antibiotic pathway specific activators regulating the production of actinorhodin, undecylprodigiosin, daunorubicin and cephamycin, respectively [36–40]. A common feature of this group of activators is that disruption of the corresponding gene abolishes production of the corresponding antibiotic, whereas overexpression results in a several-fold increase in metabolite production. When *orf1* was disrupted, however, the mutant strain showed normal MC production (Table 2). Moreover, the wild-type MC producer containing an additional copy of *orf1* in pKC1139 also had a normal MC production profile (Table 2). Interestingly, *orf4*, one of the cytochrome P450 monooxygenase encoding genes adjacent to *orf1* also showed normal MC production when disrupted (Table 2). *mitT*, therefore, appears to map to the left-hand end of the MC cluster, whereas *orf1–orf9* presumably specify biosynthesis of an unidentified polyketide product.

mmcY defines the right-hand boundary of the MC cluster

Nucleotide-sequence analysis of the MC biosynthetic gene cluster extended 30 kb upstream of *mitA* and several open-reading frames (orfs) corresponding to genes involved in sugar metabolism were identified. They included an acid trehalase (*orf12*), one ABC-type transporter (*orf16*), and four adjacent α-amylases (*orf19*, *orf20*, *orf21* and *orf22*) for starch degradation spanning more than 18 kb (Figure 2). Disruption of four genes (*orf11*, *orf12*, *orf16* and *orf19*) within this region resulted in mutants with wild-type level MC production profiles, indicating that they fall outside of the MC cluster (Table 2). At the beginning of this group of sugar metabolism genes, a gene (*mmcY*) encoding a presumed chitinase is proposed to be the upstream terminus of the MC cluster. This is evident because MC requires D-glucosamine as a biosynthetic precursor, and MmcY shows 75% identity (85% similarity) to the chitinase C gene (*chiC*) product from *S. griseus* that generates *N*-acetylglucosamine from chitin [41]. In addition, mutants with disrupted *orf11* and *orf12* genes had no effect on MC production, whereas disruption of *mmcW* and *mmcX* both affected MC production significantly (Table 2).

MC resistance genes

Antibiotic biosynthetic gene clusters typically include one or more genes for cellular self-protection [42]. Previous work has identified two mitomycin C resistance genes (*mcr* and *mrd*) with *mrd* linked to *mitA* [27–29]. Subsequent analysis showed that Mrd is a resistance protein that binds mitomycin C with 1:1 stoichiometry [28]. However, this resistance mechanism would be extremely inefficient unless the bound drug is transported out of the cell. Indeed, 5 kb upstream of *mrd*, the *mct* gene (a putative MC translocase) encoding a presumed antibiotic transporter was found and shown to be a third MC resistance component [43]. *mct* encodes a 484 amino-acid protein with 14 predicted transmembrane domains. Disruption of *mct* resulted in a mutant *S. lavendulae* strain substantially more sensitive to MC, and coexpression of *mct* with *mrd* in *Escherichia coli* dramatically increased MC resistance levels compared with individual expression of the genes [43]. In contrast, the high-level MC resistance gene (*mcrA*), which encodes an MC oxidase (McrA) capable of re-oxidizing activated MC [44], is not linked with this cluster [27,29]. Interestingly, database

Table 1

Deduced genes and their proposed functions in the MC cluster.

Gene	Amino acids	Accession numbers of homologous proteins	Typical homology (% identity, % similarity)	Proposed function
orf6	414	1020391, 3114701	(54%, 74%)	Dehydrogenase
orf5	176	2496757, 2104395	(38%, 57%)	LinA homolog
orf4	407	99020, 117302	(47%, 66%)	Cytochrome P450 hydroxylase
orf3	410	561882, 987105	(51%, 63%)	Cytochrome P450 hydroxylase
orf2	368	1552858, 1502425	(38%, 54%)	Esterase
orf1	285	118783, 1168271	(39%, 45%)	Transcriptional activator
mitT	270	2792323, 2492956	(56%, 67%)	aminoQuinate Dehydrogenase (RifI homolog)
mitS	315	2792326, 729585	(53%, 63%)	Kinase (RifN homolog)
mitR	514	1170892, 3282517	(26%, 33%)	McrA Homolog
mitQ	164	152404, 2982999	(43%, 50%)	Putative regulator
mitP	343	3056886, 2792321	(70%, 77%)	aminoDHQ Synthase (RifG homolog)
mitO	163	2791588	(32%, 48%)	Unknown
mitN	275	2792343, 2246452	(32%, 48%)	Methyltransferase
mitM	283	2792343, 1001725	(40%, 49%)	Methyltransferase
mitL	520	1502425, 1552858	(28%, 42%)	Esterase
mitK	346	2826429, 2129143	(36%, 51%)	F420-dependent H_4MPT reductase
mitJ	235	2792325, 3056884	(64%, 75%)	Phosphatase (RifM homolog)
mitI	290	–	–	Unknown
mitH	382	2129143, 2826429	(40%, 49%)	F420-dependent H4MPT reductase
mitG	404	3056883	(46%, 61%)	Oxidoreductase (RifL homolog)
mitF	257	1841491, 2506147	(39%, 51%)	Reductase
mitE	707	1040685, 665920	(31%, 52%)	CoA ligase
mitD	383	2648528, 1806159	(22%, 39%)	Unknown
mitC	260	2894171	(54%, 62%)	Unknown
mitB	272	1314568, 1651894	(36%, 43%)	Glycosyltransferase
mitA	388	2147019, 995684	(67%, 76%)	AHBA synthase (RifK homolog)
mmcA	514	–	–	Unknown
mmcB	93	2984024, 113194	(29%, 61%)	Acyl carrier protein
mmcC	470	–		Unknown
mmcD	611	2622915, 3131076	(34%, 53%)	Methyltransferase
mmcE	359	2622160, 2506843	(36%, 56%)	H4MPT:CoM Methyltransferase
mmcF	145	2792346, 1703004	(74%, 81%)	aminoDHQ Dehydratase (RifJ homolog)
mmcG	177	–	–	Unknown
mmcH	254	2105061, 2829569	(36%, 50%)	Unknown
mmcI	264	1568583, 2649391	(55%, 69%)	F420-dependent H4MPT reductase
mmcJ	274	2829504, 2735505	(36%, 47%)	F420-dependent H4MPT reductase
mmcK	460	3218385	(28%, 46%)	Unknown
mmcL	511	2829486, 2228233	(43%, 61%)	Aldehyde dehydrogenase
mmcM	472	1170892, 3282517	(54%, 69%)	McrA homolog
mmcN	395	117302, 2147740	(37%, 56%)	Cytochrome P450 hydroxylase
mmcO	474	–	–	Unknown
mrd	130	1917021	–	Mitomycin resistance determinant (Mrd)
mmcP	443	–	–	Unknown
mmcQ	123	396392, 2851659	(38%, 60%)	Unknown
mmcR	351	1169359, 730913	(40%, 58%)	O-Methyltransferase
mmcS	546	2498662, 3328168	(45%, 58%)	Carbamoyl transferase
mmcT	568	730909, 769829	(38%, 55%)	Hydroxylase
mmcU	160	1652032, 2650349	(52%, 69%)	Sulfate adenylate transferase unit I
mmcV	319	1706266, 882645	(67%, 81%)	Sulfate adenylate transferase unit II
mct	484	282580, 2995318	(52%, 66%)	Mitomycin C translocase (Mct)
mmcW	163	1172058, 127291	(39%, 54%)	Repressor
mmcX	413	–	–	Unknown
mmcY	271	2662299, 116329	(75%, 85%)	Chitinase
orf11	139	2879888	(39%, 58%)	Unknown
orf12	936	1061284, 1352001	(42%, 57%)	Acid trehalase

searches identified two McrA homologs (MitR and MmcM) within the MC cluster, both of which encode putative flavoproteins conserved in the FMN/FAD binding motif. MitR displayed weak similarity to McrA (26% indentity, 33% similarity), whereas MmcM shows end-to-end (54% identity, 69% similarity) alignment with the protein. *mitR* and *mmcM* were genetically disrupted resulting in substantially decreased MC production in the *mitR* mutant strain, in

Table 2

MC production in wild type *S. lavendulae* and gene-disruption mutants.

Number	Gene disrupted	MC production
0.0	Wild-type control	++
0.1	Additional copy of *orf1* in wild type	++
1	*orf8*	++
2	*orf4*	++
3	*orf1*	++
4	*mitR*	+
5	*mitM*	–
6	*mitI*	–
7	*mitH*	–
8	*mitE*	–
9	*mitB*	–
10	*mitA*	–
11	*mmcA*	–
12	*mmcB*	–
13	*mmcM*	++
14	*mmcP*	–
15	*mmcR*	–
16	*mmcT*	–
17	*mmcW*	++++
18	*mmcX*	++++
19	*orf11*	++
20	*orf12*	++
21	*orf16*	++
22	*orf19*	++

++, normal MC production; –, no MC production; +, reduced MC production; ++++, increased MC production.

contrast to the *mmcM* mutant, which displayed wild-type MC production levels (Table 2).

Regulatory genes

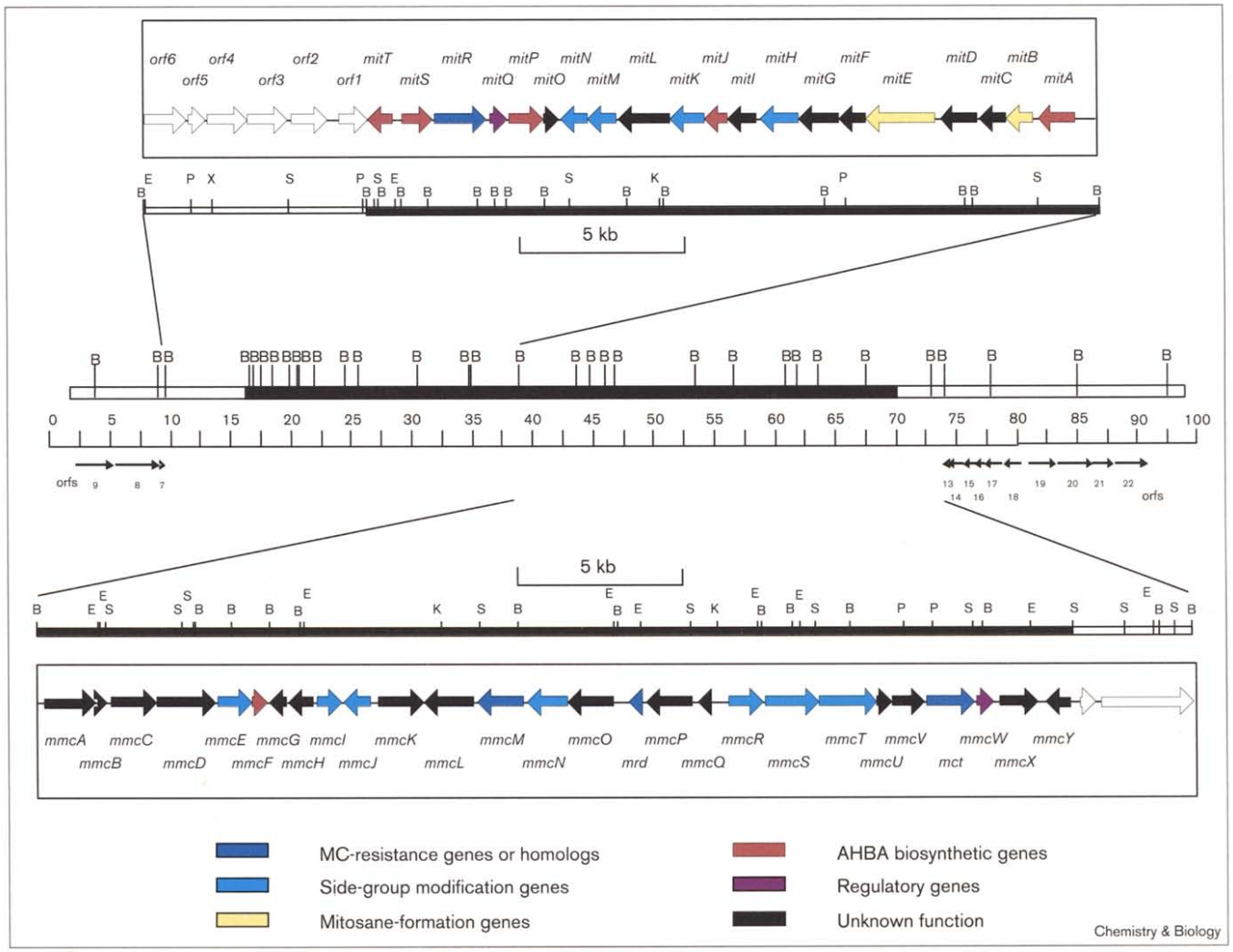
Two genes (*mitQ* and *mmcW*) were identified in the MC cluster and are presumed to be pathway-specific regulators. MitQ belongs to the OmpR–PhoB subfamily of DNA binding regulators in the two-component regulatory system, with the greatest similarity to members of the phosphate assimilation pathway (PhoR–PhoB) [45], ferric enterobactin response pair (PfeR–PfeS) [46], and one histidine protein kinase–response regulator system (HpkA–DrrA) from *Thermotoga maritima* [47]. In contrast to the MitQ group of regulators that typically serve as transcriptional activators [48], MmcW shared high sequence similarity with the MarR group of repressors. The most significant similarity corresponds to EmrR, the negative regulator of the *E. coli* multidrug resistance pump EmrAB [49], and PecS a repressor for pectinase, cellulase and blue pigment production in *Erwinia chrysanthemi* [50]. Significantly, the *mmcW* disruption mutant displayed a several-fold increase in MC production (Table 2).

AHBA biosynthetic genes

Precursor-incorporation studies previously demonstrated that AHBA is an intermediate for both the ansamycin and mitomycin natural products [18,29]. Combining the biochemical, enzymatic and molecular genetic results on the biosynthesis of the ansamycin antibiotic rifamycin, Floss and coworkers [51] have proposed that AHBA is derived from the ammoniated shikimate pathway via phosphoenolpyruvate (PEP) and erythrose 4-phosphate (E4P) by the early incorporation of nitrogen. In the shikimate pathway, PEP and E4P is first converted to 3-deoxy-D-arabino-heptulosonic acid-7-phosphate (DAHP) then stepwise transformed to 3-dehydroquinate (DHQ), 3-dehydroshikimate (DHS) and shikimate, catalyzed by DAHP synthase, DHQ synthase, DHQ dehydratase and shikimate dehydrogenase, respectively [52]. Quinate can also enter the pathway through the action of quinate dehydrogenase to generate DHQ. Evidence to support this new variant of the shikimate pathway includes the following experimental observations. First, all proposed ammoniated shikimate pathway compounds, including PEP, E4P, 3,4-dideoxy-4-amino-D-arabino-heptulosonic acid 7-phosphate (aminoDAHP), 5-deoxy-5-amino-3-dehydroquinic acid (aminoDHQ), and 5-deoxy-5-amino-3-dehydroshikimic acid (aminoDHS), can be readily converted into AHBA by cell-free extracts from the ansamycin producers, whereas none of the early shikimate pathway intermediates, DAHP, DHQ, DHS, quinic acid or shikimic acid, can be incorporated into AHBA under the same conditions [21,51]. Second, the *rif* cluster has been sequenced, and all of the genes encoding early shikimate pathway enzymes are found within the cluster [31]. Finally, the ability of the rifamycin AHBA synthase (RifK) to catalyze dehydration of aminoDHS to AHBA has been demonstrated previously [30]. Our work has also shown that the AHBA synthase gene (*mitA*) in *S. lavendulae* is required for AHBA biosynthesis [29].

A group of AHBA biosynthetic genes similar to those described for *rif* have been identified in the MC cluster. In addition to AHBA synthase, six gene products in the MC cluster showed high sequence similarity (over 43% identity) with their rifamycin AHBA biosynthetic gene homologs. These gene products include aminoDHQ synthase (MitP, RifG equivalent), aminoquinate dehydrogenase (MitT, RifI equivalent), oxidoreductase (MitG, RifL equivalent), phosphatase (MitJ, RifM equivalent), kinase (MitS, RifN equivalent) and aminoDHQ dehydratase (MmcF, RifJ equivalent). In addition to the significant sequence similarity to rifamycin counterparts, all three putative MC shikimate pathway enzymes displayed significant alignment with microbial primary shikimate metabolic enzymes including MitT with the quinate dehydrogenase (AroE) from *Methanococcus jannaschii* (28% identity, 46% similarity) [53], MitP with the DHQ synthase (AroB) from *Mycobacterium tuberculosis* (46% identity, 61% similarity) [54], and MmcF with the DHQ dehydratase from *S. coelicolor* (50% identity, 62% similarity) [55]. Despite extensive sequencing of 15 kb on either side of the mapped right- and left-hand ends of the MC cluster, an aminoDAHP synthase

Figure 2



Organization of the MC biosynthetic gene cluster. The deduced ORFs are drawn to scale, and their corresponding genes are italicized. The filled bars indicate the location of the MC cluster.

Abbreviations of restriction enzymes used: B, *Bam*HI; S, *Sph*I; P, *Pst*I; E, *Eco*RI; X, *Xho*I; K, *Kpn*I.

gene corresponding to RifH (the proposed first enzyme in the *de novo* biosynthesis from PEP and E4P in the *rif* cluster), was not found (Figure 2). Interestingly, a *rifH* homolog has been cloned from *S. lavendulae* genomic DNA through Southern hybridization and shown to be unlinked to the MC cluster (Y.M., unpublished observations).

The existence of the nonshikimate-pathway-related phosphatase/kinase pair in the MC cluster (MitJ/MitS) and the *rif* cluster (RifM/RifN) further support the finding that these two genes are required for AHBA biosynthesis [25]. In addition to the strong homology to RifM, MitJ also showed 56% identity (69% similarity) with ORF8 from the ansamycin antibiotic ansamitocin producer *Actinosynnema pretiosum auranticum*. Other polypeptides with considerable sequence similarity belong to the CBBY family of phosphoglycolate phosphatases in glycolate oxidation [56]. MitS, most similar to RifN (53% identity, 63% similarity), also showed significant similarity with the glucose kinase (involved in glucose repression) from *S. coelicolor* and *Bacillus megaterium* [57,58]. *mitG*, the third nonshikimate-pathway-related AHBA biosynthetic gene in this cluster is also worthy of note because it shows exclusive similarity (46% identity, 61% similarity) with oxidoreductase RifL and its equivalent in *Actinosynnema pretiosum auranticum*. The detailed function of these gene products in AHBA biosynthesis remains to be elucidated.

Mitosane-formation genes

Precursor-incorporation studies established that the mitosane core is assembled from the condensation of AHBA and D-glucosamine [20,21,24]. Although no specific

gene products can be assigned for forming the three bonds bridging AHBA and D-glucosamine, two genes downstream of *mitA* (*mitB* and *mitE*) probably encode enzymes that mediate one of these reactions. MitB shows local sequence similarity with a group of glycosyltransferases involved in glycopeptide antibiotic and polysaccharide biosynthesis, the typical function of which is to attach an activated sugar residue to a core compound [29,59]. MitE has weak similarity (22% identity and 45% similarity) to the two cloned 4-hydroxybenzoate-CoA ligases from *Rhodopseudomonas palustris* involved in the anaerobic degradation of aromatic compounds [60]. MitE also showed similarity to a group of long-chain fatty acid CoA ligases, as well as to the *O*-succinylbenzoic acid CoA synthetase in Vitamin K2 biosynthesis [61]. *mitB* and *mitE* disruption mutants both have an MC-deficient phenotype (Table 2).

Based on reasonable biosynthetic principles, the condensation of AHBA with D-glucosamine can be initiated in two different ways — either initial formation of a C_{8a}–C_9 bond by an acylation or alkylation reaction, or formation of a Schiff base between the AHBA nitrogen and D-glucosamine C1 aldehyde, followed by the ring closure at C_{8a}–C_9. MitR, one of the two McrA homologs might be involved in one of the ring-closure reactions. Interestingly, MitR showed high sequence homology to the plant berberine bridge enzyme (BBE; 30% identity, 37% similarity) in benzophenanthridine alkaloid formation, where it catalyzes an unusual C–C bond formation of the berberine bridgehead carbon of (*S*)-scoulerine from the *N*-methyl carbon of (*S*)-reticuline [62]. Using a mechanism similar to BBE, it is possible that MitB is involved in C_{8a}–C_9 bond formation. The decreased MC production in the *mitR* disruption mutant might be due to the existence of isoenzymes (e.g. MmcM) that could catalyze the reaction in the absence of functional MitR.

Genes for tailoring of the mitosane core

Complete assembly of MC requires functionalization of several sites on the core mitosane ring system. First, complete reduction of the carbonyl group at C-6 must occur. Second, hydroxylation at C-5 and C9a must proceed followed by methylation at C-9a. Third, amination at C-7 must occur presumably through initial hydroxylation followed by transamination. Fourth, oxidation of the hydroxyl groups at C-5 and C-8 to form the benzoquinone are required. Fifth, intramolecular amination of C-1 by N-1a to form the aziridine ring must be completed, and finally, carbamoylation at C-10 completes assembly of the molecule. Several putative enzymes found in this cluster are candidates for these modifications and are discussed below.

Methylation

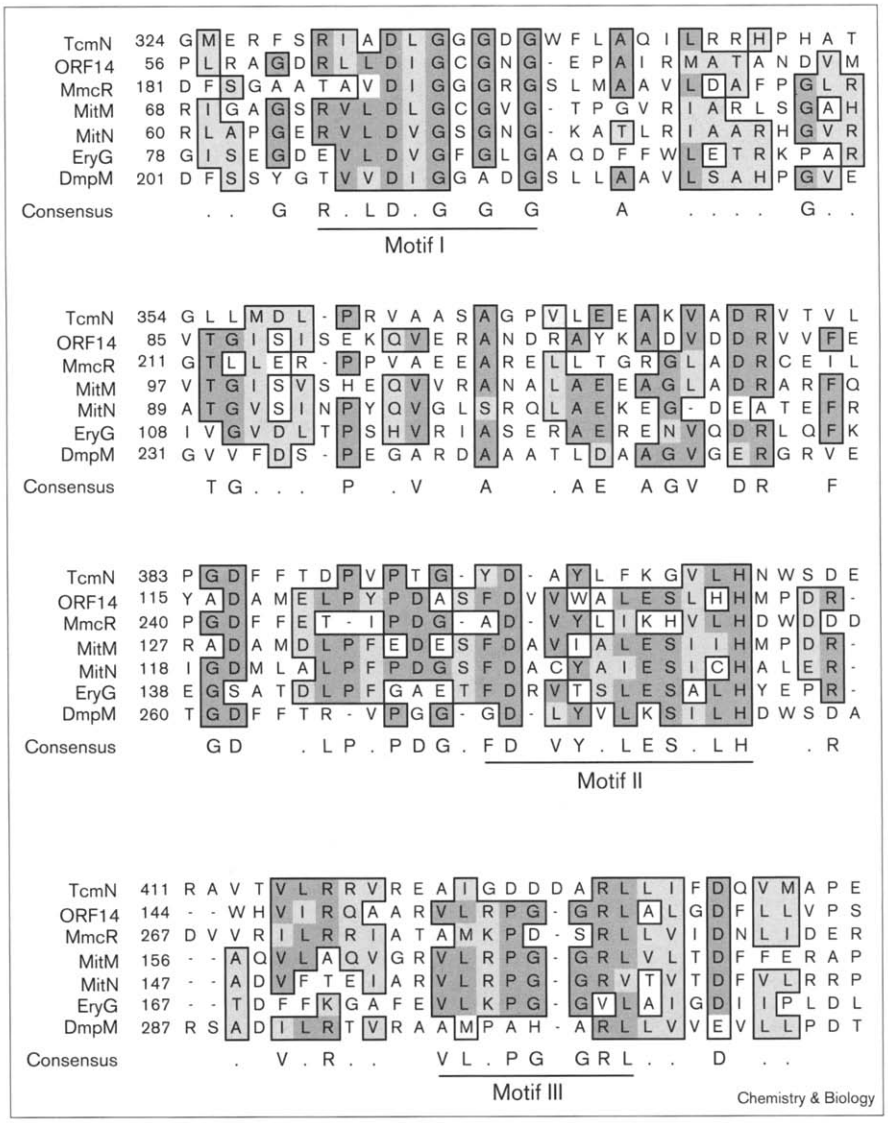
In contrast to MC, which has an *O*-methyl group at C-9a, mitomycin A and mitomycin B also contain a C-7 *O*-methyl group, and mitomycin B, mitomycin D and porfiromycin have an *N*-methyl group on the aziridine ring (Figure 1). Radiolabeled precursor incorporation studies showed that all of the *O*- and *N*-methyl (but not the C-methyl) groups in the mitomycin molecules are derived from L-methionine [22]. Typically, the methyl donor for most C1 reactions is *S*-adenosyl-L-methionine (SAM), which can be formed through activation of L-methionine by ATP. Three SAM-dependent methyltransferase genes were identified in this cluster (encoding MitM, MitN and MmcR), all of which have three conserved *S*-adenosylmethionine- or *S*-adenosylhomocysteine-binding motifs [63] (Figure 3). Interestingly, database searches of MitM and MitN (probably responsible for the MC C-9a sidechain methylation) revealed a group of plant δ-(24)-sterol C-methyltransferases, but have a closer phylogenetic relationship with the rifamycin *O*-methyltransferase (ORF14) and erythromycin *O*-methyltransferase (EryG) [31,64] (Figure 4). In contrast, protein database searches revealed that MmcR is most related to other *Streptomyces* antibiotic biosynthetic *O*-methyltransferases with greatest similarity to *O*-demethylpuromycin *O*-methyltransferase (44% identity, 60% similarity) from *S. anulatus* and carminomycin 4-*O*-methyltransferase from *S. peucetius* [65,66]. MmcR might be involved in the *O*-methylation of the phenol ring of MC before oxidation to the quinone. Both *mmcR* and *mitM* were shown to be essential for MC biosynthesis because disruption of each one completely abolished MC production (Table 2).

A SAM-independent methyltransferase, MmcD, was also identified in the MC cluster. MmcD shares strong sequence homology with the magnesium-protoporphyrin IX monomethyl ester oxidative cyclase (34% identity, 53% similarity) from *Methanobacterium thermoautotrophicum* (GenBank accession #2622915), as well as the phosphonoacetaldehyde methyltransferase from *Streptomyces wedmorensis* [67], the P-methyltransferase from *Streptomyces hygroscopicus* [68] and the fortimicin KL methyltransferase from *Micromonospora olivasterospora* [69]. Instead of SAM, this group of methyltransferases uses methylcobalamine or a structurally related protoporphyrin as the direct methyl donor. Although the greatest number of matches were made to protoporphyrin methyltransferases, this enzyme is expected to have another function in the MC biosynthetic pathway as all the *O*- and *N*-methyl groups of MC have been shown to be derived from SAM-dependent methyltransferases.

C-6 carbonyl reduction

The C-6 methyl group was shown previously to be derived from the reduction of the carboxylic acid of AHBA, because [carboxy-^{13}C] AHBA can be efficiently, and specifically, incorporated into the C-6 methyl group of porfiromycin [20]. In the MC cluster, four F420-dependent tetrahydromethanopterin (H_4MPT) reductase genes (encoding MitH, MitK, MmcI and MmcJ) and one H_4MPT:CoM

Figure 3



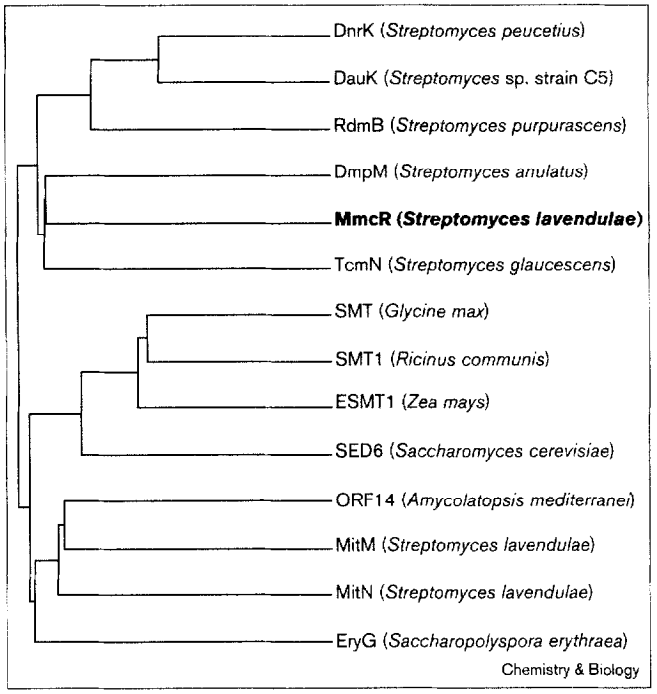
The three SAM-dependent methyltransferase conserved motifs can be found in MitM, MitN and MmcR. DmpM [65], TcmN [100], ORF14 [31] and EryG [101] are *O*-methyltransferases for puromycin, tetracenomycin C, rifamycin and erythromycin biosynthesis, respectively.

methyltransferase gene (encoding MmcE) are candidates for the C-6 carbonyl reduction. In the methanogenesis pathway of *Methanobacterium thermoautotrophicum*, two cofactor F420-dependent H_4MPT reductases, and one cofactor CoM dependent methyltransferase are required in the seven-step reduction from CO_2 to CH_4. Steps 4 to 6 from CH-H_4MPT to CH_2-H_4MPT, and CH_3-H_4MPT to CH_3-CoM are catalyzed by N^5, N^{10}-methylene-H_4MPT dehydrogenase, N^5, N^{10}-methylene-H_4MPT reductase and N^5-methyl-H_4MPT:CoM methyltransferase, respectively [70,71]. All four enzymes (MitH, MitK, MmcI, MmcJ) in this cluster showed local sequence similarity with the cloned F420-dependent H_4MPT reductase (42% identity, 62% similarity in several 50 amino-acid fragments) [72,73]. One of these genes, *mitH* was disrupted, and the mutant strain had an MC-deficient phenotype (Table 2). MmcE is notable because the deduced protein sequence contains two domains showing significant alignment (33% identity, 56% similarity) to the amino terminus of H_4MPT:CoM methyltransferase from *Methanobacterium thermoautotrophicum* [74], and the remaining carboxyl terminus is related to fatty acid biosynthetic acyl carrier proteins (ACPs) [75,76]. The potential function of this ACP-like domain in MC biosynthesis remains unknown, as does the role of a distinct gene (*mmcB*) encoding a putative ACP identified just upstream of *mmcE*. Significantly, the disruption of *mmcB* resulted in total abrogation of MC production (Table 2).

Hydroxylation

The two putative hydroxylases (encoded by *mmcN*, *mmcT*) identified in the MC cluster are candidates for catalyzing hydroxylation at the C-5, C-7, and C-9a positions on the

Figure 4



Sequence similarity of MitM, MitN and MmcR with other O-methyltransferases (DmpM [65], TcmN [100], ORF14 [31], EryG [101], RdmB [102], DnrK [66], and DauK [103]) and C-methyltransferases (SMT [64], ESMT1 [104], SMT1 [105], and SED6 [106]). The dendrogram was constructed with the program PILEUP [95].

mitosane system. MmcN belongs to the cytochrome P450 family of monooxygenases, with greatest homology (37% identity, 56% similarity) to the two herbicide-inducible cytochrome P450s (P450-SU1 and P450-SU2) from *S. griseolus*, as well as to RapJ and RapN in the rapamycin biosynthetic gene cluster from *S. hygroscopicus* [33,35]. MmcT showed highest similarity to the tetracenomycin C hydroxylase (TcmG) in *Streptomyces glaucescens* (38% identity, 55% similarity), with lower but significant sequence similarity to a group of phenol or hydroxybenzoate hydroxylases [77]. Genetic disruption of *mmcT* completely blocked MC biosynthesis (Table 2).

Carbamoylation

The carbamoyl group of MC is derived intact from L-citrulline or L-arginine with carbamoyl phosphate as the incorporated precursor [21]. In eubacteria, carbamoyl phosphate can be generated from L-glutamine, HCO_3^-, and ATP by the enzyme carbamoyl phosphate synthetase, which is indispensable for pyrimidine biosynthesis. One candidate carbamoyl transferase gene (*mmcS*) was identified directly upstream of *mmcT*. MmcS belongs to the NodU/CmcH family of *O*-carbamoylation enzymes, with the greatest similarity (35% identity, 44% similarity) to NolO from *Rhizobium* sp. [78]. Other members with significant alignment in this family include NolO from *Bradyrhizobium japonicum* [79] and NodU from *Rhizobium* sp. for 6-*O*-carbamoylation of Nod factors [80], and CmcH from *Nocardia lactamdurans* and *S. clavuligerus* for 3′-hydroxymethylcephem *O*-carbamoylation in cephamycin biosynthesis [81].

Discussion

Bridging primary and secondary metabolism

The shikimate pathway is an essential metabolic route in microorganisms and plants for aromatic amino-acid biosynthesis. Genes that encode the early shikimate pathway enzymes from various organisms have been well studied and are often dispersed along the chromosome as revealed by sequencing of complete microbial genomes [53,54,82]. The finding that the ansamycin and mitomycin natural products are derived in part from an ammoniated shikimate pathway whose genes are clustered on the bacterial chromosome is a significant difference to the primary metabolic network, and could suggest an important evolutionary bridge leading to secondary metabolism. The lack of incorporation of early shikimate pathway intermediates into mitomycin and ansamycin metabolites indicated the existence and ultimate substrate specificity of the alternate ammoniated shikimate pathway enzymes. The conversion of aminoDAHP and aminoshikimic acid by the corresponding primary shikimate pathway enzymes to aminoDHQ and aminoDHS, respectively [51], however, suggested that substrate specificity in the primary metabolic shikimate pathway is mainly determined by the initial reaction step. This notion is further supported by the disruption results for *rifG* and *rifI* mutants showing only a minimal affect on rifamycin production [25].

In addition to the absence of an aminoDAHP synthase gene, the organization of the AHBA biosynthetic genes in the MC cluster is quite different compared to the *rif* cluster. In *rif* (with the exception of *rifJ*), all AHBA biosynthetic genes are found within a defined subcluster that are organized into a single apparent operon. In contrast, almost all of the *mit*/*mmc*-encoded AHBA genes are scattered within the 55 kb MC cluster. Thus, as opposed to the multifunctional polyketide gene clusters whose linearity of architecture reflects a precise pattern of biosynthesis, the MC cluster is biochemically less transparent based on a similar primary analysis. Nevertheless, with a complete set of MC biosynthetic genes in hand it will be possible to begin a thorough dissection of the mode of assembly and function of each enzyme. In addition, the MC cluster provides a good model for analyzing genetic evolution both vertically, from the primary metabolic shikimate pathway to the secondary shikimate pathway related route, and horizontally by comparing different groups of secondary metabolic biosynthetic clusters.

The MC biosynthetic network

In a typical liquid culture of *S. lavendulae*, MC production is initiated 24 h after inoculating the seed culture, reaches

maximum production in 2 days, and maintains drug synthesis during stationary phase for another 2 days. Compared with high-level MC resistance of the wild-type *S. lavendulae* (>150 μg/ml), MC production is relatively low (<5 μg/ml MC). The significant gap between the self-resistance and production levels makes it possible to improve drug production through genetic engineering. In this work, disruption of either the candidate repressor gene (*mmcW*) or the downstream *mmcX* (encoding an unknown membrane protein) in the MC cluster resulted in a several-fold increase in MC production. The existence of a repressor gene(s) is not uncommon in *Streptomyces* antibiotic biosynthetic gene clusters. Previous examples include *mmyR* from the methylenomycin cluster [83], *actII–orfI* in the actinorhodin cluster [84], *jadR(2)* in jadomycin biosynthesis [85] and *dnrO* in the daunorubicin cluster [86]. Disruption of *jadR* [20] and *mmyR* also resulted in increased levels of the corresponding antibiotics [83,85].

In order to avoid autotoxicity, drug-producing microorganisms must evolve self-protection systems. Currently, three types of self-protection mechanisms have been identified in *S. lavendulae* for MC resistance, including MC binding (Mrd), efflux (Mct) and reversing MC reductive activation (McrA). In principle, resistance genes must be expressed before drug formation. In this respect, it is interesting to note the linkage of the MC resistance genes with putative regulatory genes. Expression of the high-level resistance gene *mcrA* has been demonstrated to be regulated by the downstream gene *mcrB*, which is presumably cotranscribed with *mcrA* [27]. Though the function of the McrA homolog MitR in the MC cluster remains unknown, *mitR* is also followed by a cotranscribed regulatory gene (*mitQ*), and the putative MC translocase gene, *mct*, is followed by the repressor gene, *mmcW*. Genetic linkage of membrane transporter/resistance and repressor genes have been described in a number of cases, including *tetA/tetR* in tetracycline resistance [87], *tcmA/tcmR* in tetracenomycin C resistance [88], *actII–orf2/actII–orf1* in actinorhodin resistance [84], and the *qacA/qacR* pair for multidrug resistance in *S. aureus* [89]. The specific regulatory role that MmcW plays in MC resistance and biosynthesis remains to be elucidated.

Significance

Although mitomycin C (MC) was first isolated more than 40 years ago and has been used in anticancer chemotherapy since the 1960s, the mechanistic details and order of its biosynthesis has remained unclear. This work represents a complete characterization of the gene cluster in *Streptomyces lavendulae* responsible for MC biosynthesis. Our results are clearly consistent with precursor incorporation studies gathered in the 1970s, showing that MC is biosynthetically derived from D-glucosamine, L-methionine, carbamoyl phosphate and AHBA, and also support the use of the variant *de novo* shikimate pathway leading to AHBA [21,51]. Many, if not all of the genes responsible for the formation of the mitosane and aziridine rings are located within the boundary of the 55 kb MC cluster. These genes are of special interest because their functions are poorly understood and they might be useful as probes for identifying related natural-product biosynthetic genes from other microorganisms and plants.

The cloned genes presented here provide a key starting point for future molecular genetic and biochemical studies on MC biosynthesis and natural-product assembly. The advantage of having this information has already been demonstrated through genetic disruption of the candidate repressor gene (*mmcW*) that provided a several-fold increase in MC production. In addition, expression and disruption of selected genes should be useful for engineering the biosynthesis of clinically valuable mitomycin analogues, as well as more complex hybrid natural-product systems. Finally, the MC resistance and regulatory genes identified in this cluster will provide important insight into the MC biosynthetic and regulatory network in *S. lavendulae*.

Materials and methods

Bacterial strains and cloning vectors

S. lavendulae NRRL2564 was used as the source strain for cosmid library construction and the creation of gene-disruption mutants. *E. coli* DH5α was used as the host strain for constructing the library and subsequent DNA manipulation. *E. coli* strain S17-1 [90] served as the conjugative host for introducing foreign DNA into *S. lavendulae*. The cosmid library was constructed with the *E. coli–Streptomyces* shuttle vector pNJ1 [91], and pUC119 was routinely used as a vector for subcloning and sequencing. The conjugative *E. coli–Streptomyces* shuttle vector pKC1139 [92] was used for gene disruption in *S. lavendulae*.

DNA manipulation

Standard *in vitro* techniques were used for DNA manipulation [93]. *S. lavendulae* genomic DNA was harvested using standard procedures [94]. A library of size-fractionated genomic DNA in pNJ1 [91] was screened with the rifamycin AHBA synthase (*rifK*) gene probe from *Amycolatopsis mediterranei* [30]. Through subsequent cosmid walking, a contiguous 120 kb region of *S. lavendulae* chromosomal DNA containing the putative MC biosynthetic genes was mapped. M13 forward and reverse primers were used for sequencing (Gibco BRL). To accomplish this, individual fragments of less than 5 kb were subcloned into pUC119 and serial deletion subclones were generated using the exonuclease III Erase-a Base System (Promega).

DNA sequencing and analysis

Automatic DNA sequencing was carried out with the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems), and analyzed on an Applied Biosystems model 377 DNA Sequencer at the University of Minnesota Advanced Genetic Analysis Center. Both DNA strands were sequenced redundantly a minimum of three times. Sequence compilation was performed with MacVector (Oxford Molecular Group) and GeneWorks (Oxford Molecular Group) software, and sequence homology analysis was accomplished using BLAST [95] and GCG programs [96].

Disruption mutant construction

A 1.4 kb *Apa*L1–*Hind*III fragment from pFD666 [97] containing the *aphII* gene for kanamycin resistance was routinely used as the selection marker for the creation of gene-disruption constructs. The target

genes were subcloned into pUC119, cut at a unique internal restriction site, blunt-ended, and ligated with the end-blunted selection marker. The inserts were then transferred from pUC119 to pKC1139, and conjugated into wild-type *S. lavendulae* [29]. Transconjugants were selected on AS1 plates [98], overlaid with apramycin, kanamycin and nalidixic acid followed by propagation on R5T plates at 37°C for several generations. Disruption mutants were selected based on the phenotype changing from apramycin- and kanamycin-resistant to apramycin-sensitive and kanamycin-resistant. Replacement of the chromosomal copy of the target gene with the disrupted plasmid-borne copy was confirmed by Southern blot hybridization.

Mitomycin C analysis

MC production was evaluated using 3-day cultures in Nishikohri media [99]. The culture broth was extracted twice with equal volumes of ethyl acetate. After removing the chemical solvent by vacuum, the crude broth extract was dissolved in 50% methanol and 50% 50 mM pH 7.2 Tris buffer and monitored by HPLC (C_{18} reverse phase column) at 363 nm. A continuous methanol gradient from 20% to 60% in methanol/50 mM pH 7.2 Tris buffer system over 24 min was employed to resolve MC from other crude extract components. A 90% $CHCl_3$/10% MeOH solvent system was used to resolve and detect MC on TLC plates.

Accession numbers

The entire sequence reported here, together with the two previously submitted loci [29,43], have been deposited in the GenBank database under the accession number AF127374.

Acknowledgements

This work was supported in part by the Biological Process Technology Institute Seed Grant Program, University of Minnesota (to D.H.S.). M.V. was the recipient of a postdoctoral fellowship from National Cancer Institute Training Grant CA09138.

References

1. Hata, T., *et al.* & Hoshi, T. (1956). Mitomycin, a new antibiotic from *Streptomyces. J. Antibiot. Ser. A* **9**, 141-146.
2. Verweij, J. (1997). Mitomycins. *Cancer Chemother. Biol. Response Modif.* **17**, 46-58.
3. Boyer, M.J. (1997). Bioreductive agents: a clinical update. *Oncol. Res.* **9**, 391-395.
4. Cummings, J., Spanswick, V.J., Tomasz, M. & Smyth, J.F. (1998). Enzymology of mitomycin C metabolic activation in tumour tissue – implications for enzyme-directed bioreductive drug development. *Biochem. Pharmacol.* **56**, 405-414.
5. Tomasz, M. & Palom, Y. (1997). The mitomycin bioreductive antitumor agents: cross-linking and alkylation of DNA as the molecular basis of their activity. *Pharmacol. Therap.* **76**, 73-87.
6. Spanswick, V.J., Cummings, J. & Smyth, J.F. (1998). Current issues in the enzymology of mitomycin C metabolic activation. *Gen. Pharmacol.* **31**, 539-544.
7. Henderson, I.C. (1993). Recent advances in the usage of mitomycin. Proceedings of a symposium. Hawaii, March 21-24, 1991. *Oncology* **1**, 1-83.
8. Pan, S.S. & Iracki, T. (1988). Metabolites and DNA adduct formation from flavoenzyme-activated porfiromycin. *Mol. Pharmacol.* **34**, 223-228.
9. Wakaki, K., *et al.* & Fujimoto, Y. (1958). Isolation of new fractions of antitumor mitomycins. *Antibiot. Chemother.* **8**, 228-240.
10. Paz, M.M. & Hopkins, P.B. (1997). DNA-DNA interstrand cross-linking by FR66979 - intermediates in the activation cascade. *J. Am. Chem. Soc.* **119**, 5999-6005.
11. Williams, R.M., Rajski, S.R. & Rollins, S.B. (1997). FR900482, a close cousin of mitomycin C that exploits mitosene-based DNA cross-linking. *Chem. Biol.* **4**, 127-137.
12. Hirai, O., *et al.*, & Mine, Y. (1994). Microbial mutagenicity and *in vitro* chromosome aberration induction by FK973, a new antitumor agent. *Mutation Res.* **324**, 43-50.
13. Naoe, Y., *et al.*, & Shimomura, K. (1998). FK317 – a novel substituted dihydrobenzoxazine with potent antitumor activity which does not induce vascular leak syndrome. *Cancer Chemother. Pharmacol.* **42**, 31-36.
14. Smitskamp-Wilms, E., Hendriks, H.R. & Peters, G.J. (1996). Development, pharmacology, role of DT-diaphorase and prospects of the indoloquinone EO9. *Gen. Pharmacol.* **27**, 421-429.
15. Evans, J.W., Yudoh, K., Delahoussaye, Y.M. & Brown, J.M. (1998). Tirapazamine is metabolized to its DNA-damaging radical by intranuclear enzymes. *Cancer Res.* **58**, 2098-2101.
16. Kasai, M. & Arai, H. (1995). Novel mitomycin derivatives. *Exp. Opin. Ther. Patents* **5**, 757-770.
17. Shikano, M., *et al.*, & Uehara, Y. (1998). 1a-docosahexaenoyl mitomycin C: a novel inhibitor of protein tyrosine kinase. *Biochem. Biophys. Res. Commun.* **248**, 858-863.
18. Becker, A.M., Herlt, A.J., Hilton, G.L., Kibby, J.J. & Rickards, R.W. (1983). 3-Amino-5-hydroxybenzoic acid in antibiotic biosynthesis. VI. Directed biosynthesis studies with ansamycin antibiotics. *J. Antibiot.* **36**, 1323-1328.
19. Kibby, J.J. & Rickards, R.W. (1981). The identification of 3-amino-5-hydroxybenzoic acid as a new natural aromatic amino acid. *J. Antibiot.* **34**, 605-607.
20. Anderson, M.G., Kibby, J.J., Rickards, R.W. & Rothschild, J.M. (1980). Biosynthesis of the mitomycin antibiotics from 3-amino-5-hydroxybenzoic acid. *J. Chem. Soc. Chem. Comm.* 1277-1278.
21. Hornemann, U. (1981). Biosynthesis of the mitomycins. In *Biosynthesis.* (Corcoran, J. W., ed.), pp. 295-312, The Chemical Society, London.
22. Bezanson, G.S. & Vining, L.C. (1971). Studies on the biosynthesis of mitomycin C by *Streptomyces verticillatus. Can. J. Biochem.* **49**, 911-918.
23. Hornemann, U. & Eggert, J.H. (1975). Utilization of the intact carbamoyl group of L-(NH_2CO-^{13}C,^{15}N) citrulline in mitomycin biosynthesis by *Streptomyces verticillatus. J. Antibiot.* **28**, 841-843.
24. Hornemann, U., Kehrer, J.P., Nunez, C.S. & Ranieri, R.L. (1974). D-glucosamine and L-citrulline, precursors in mitomycin biosynthesis by *Streptomyces verticillatus. J. Am. Chem. Soc.* **96**, 320-322.
25. Floss, H.G. (1997). Natural products derived from unusual variants of the shikimate pathway. *Nat. Prod. Rep.* **14**, 433-452.
26. Tomasz, M. (1995). Mitomycin C: small, fast and deadly (but very selective). *Chem. Biol.* **2**, 575-579.
27. August, P.R., Flickinger, M.C. & Sherman, D.H. (1994). Cloning and analysis of a locus (*mcr*) involved in mitomycin C resistance in *Streptomyces lavendulae. J. Bacteriol.* **176**, 4448-4454.
28. Sheldon, P.J., Johnson, D.A., August, P.J., Liu, H.-w. & Sherman, D.H. (1997). Characterization of a mitomycin-binding drug resistance mechanism from the producing organism, *Streptomyces lavendulae. J. Bacteriol.* **179**, 1796-1804.
29. Mao, Y.Q., Varoglu, M. & Sherman, D.H. (1999). Genetic localization and molecular characterization of two key genes (*mitAB*) required for biosynthesis of the antitumor antibiotic mitomycin C. *J. Bacteriol.* **181**, 2199-2208.
30. Kim, C.G., Yu, T. W., Fryhle, C.B., Handa, S. & Floss, H.G. (1998). 3-Amino-5-hydroxybenzoic acid synthase, the terminal enzyme in the formation of the precursor of mC_7N units in rifamycin and related antibiotics. *J. Biol. Chem.* **273**, 6030-6040.
31. August, P.R., *et al.*, & Floss, H.G. (1998). Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the *rif* biosynthetic gene cluster of *Amycolatopsis mediterranei* S699. *Chem. Biol.* **5**, 69-79.
32. Rodriguez, A.M., Olano, C., Mendez, C., Hutchinson, C.R. & Salas, J.A. (1995). A cytochrome P450-like gene possibly involved in oleandomycin biosynthesis by *Streptomyces antibioticus. FEMS Microbiol. Lett.* **127**, 117-120.
33. Schwecke, T., *et al.*, & Leadlay, P.F. (1995). The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. *Proc. Natl. Acad. Sci. USA* **92**, 7839-7843.
34. Horii, M., Ishizaki, T., Paik, S.Y., Manome, T. & Murooka, Y. (1990). An operon containing the genes for cholesterol oxidase and a cytochrome P-450-like protein from a *Streptomyces sp. J. Bacteriol.* **172**, 3644-3653.
35. Omer, C.A., *et al.* & O'Keefe, D.P. (1990). Genes for two herbicide-inducible cytochromes P-450 from *Streptomyces griseolus. J. Bacteriol.* **172**, 3335-3345.
36. Fernandez-Moreno, M.A., Caballero, J.L., Hopwood, D.A. & Malpartida, F. (1991). The *act* cluster contains regulatory and antibiotic export genes, direct targets for translational control by the *bldA* tRNA gene of *Streptomyces. Cell* **66**, 769-780.
37. Perez-Llarena, F.J., Liras, P., Rodriguez-Garcia, A. & Martin, J.F. (1997). A regulatory gene (*ccaR*) required for cephamycin and clavulanic acid production in *Streptomyces clavuligerus*: amplification results in overproduction of both beta-lactam compounds. *J. Bacteriol.* **179**, 2053-2059.

38. Takano, E., Gramajo, H.C., Strauch, E., Andres, N., White, J. & Bibb, M.J. (1992). Transcriptional regulation of the *redD* transcriptional activator gene accounts for growth-phase-dependent production of the antibiotic undecylprodigiosin in *Streptomyces coelicolor* A3(2). *Mol. Microbiol.* **6**, 2797-2804.
39. Tang, L., Grimm, A., Zhang, Y.X. & Hutchinson, C.R. (1996). Purification and characterization of the DNA-binding protein DnrI, a transcriptional factor of daunorubicin biosynthesis in *Streptomyces peucetius*. *Mol. Microbiol* **22**, 801-813.
40. Wietzorrek, A. & Bibb, M. (1997). A novel family of proteins that regulates antibiotic production in streptomycetes appears to contain an OmpR-like DNA-binding fold. *Mol. Microbiol* **25**, 1181-1184.
41. Ohno, T., *et al.* & Watanabe, T. (1996). A modular family 19 chitinase found in the prokaryotic organism *Streptomyces griseus* HUT 6037. *J. Bacteriol.* **178**, 5065-5070.
42. Seno, E.T. & Baltz, R.H. (1989). *Structural Organization and Regulation of Antibiotic Biosynthesis and Resistance Genes in Actinomycetes*. CRC Press, Boca Raton, FL.
43. Sheldon, P.J., Mao, Y.Q., He, M. & Sherman, D.H. (1999). Mitomycin resistance in *Streptomyces lavendulae* includes a novel drug-binding protein-dependent export system. *J. Bacteriol.*, in press.
44. Johnson, D.A., August, P.R., Shackleton, C., Liu, H.-w. & Sherman, D.H. (1997). Microbial resistance to mitomycins involves a redox relay mechanism. *J. Am. Chem. Soc.* **119**, 2576-2577.
45. Makino, K., Shinagawa, H., Amemura, M. & Nakata, A. (1986). Nucleotide sequence of the *phoB* gene, the positive regulatory gene for the phosphate regulon of *Escherichia coli* K-12. *J. Mol. Biol.* **190**, 37-44.
46. Dean, C.R., Neshat, S. & Poole, K. (1996). PfeR, an enterobactin-responsive activator of ferric enterobactin receptor gene expression in *Pseudomonas aeruginosa*. *J. Bacteriol.* **178**, 5361-5369.
47. Lee, P.J. & Stock, A.M. (1996). Characterization of the genes and proteins of a two-component system from the hyperthermophilic bacterium *Thermotoga maritima*. *J. Bacteriol.* **178**, 5579-5585.
48. Mizuno, T. & Tanaka, I. (1997). Structure of the DNA-binding domain of the OmpR family of response regulators. *Mol. Microbiol.* **24**, 665-667.
49. Lomovskaya, O., Lewis, K. & Matin, A. (1995). EmrR is a negative regulator of the *Escherichia coli* multidrug resistance pump EmrAB. *J. Bacteriol.* **177**, 2328-2334.
50. Praillet, T., Nasser, W., Robert-Baudouy, J. & Reverchon, S. (1996). Purification and functional characterization of PecS, a regulator of virulence-factor synthesis in *Erwinia chrysanthemi*. *Mol. Microbiol.* **20**, 391-402.
51. Kim, C.G., *et al.*, & Floss, H.G. (1996). Biosynthesis of 3-amino-5-hydroxybenzoic acid, the precursor of mC_7N units in ansamycin antibiotics. *J. Am. Chem. Soc.* **118**, 7486-7491.
52. Dewick, P.M. (1998). The biosynthesis of shikimate metabolites. *Nat. Prod. Rep.* **15**, 17-58.
53. Bult, C.J., *et al.*, & Venter, J.C. (1996). Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **273**, 1058-1073.
54. Cole, S.T., *et al.*, & Barrell, B.G. (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537-544.
55. White, P.J., Young, J., Hunter, I.S., Nimmo, H.G. & Coggins, J.R. (1990). The purification and characterization of 3-dehydroquinase from *Streptomyces coelicolor*. *Biochem. J.* **265**, 735-738.
56. Schaferjohann, J., Yoo, J.G., Kusian, B. & Bowien, B. (1993). The *cbb* operons of the facultative chemoautotroph *Alcaligenes eutrophus* encode phosphoglycolate phosphatase. *J. Bacteriol.* **175**, 7329-7340.
57. Angell S., Schwarz E., & Bibb, M.J. (1992). The glucose kinase gene of *Streptomyces coelicolor* A3(2): its nucleotide sequence, transcriptional analysis and role in glucose repression. *Mol. Microbiol.* **6**, 2833-2844.
58. Spath, C., Kraus, A. & Hillen, W. (1997). Contribution of glucose kinase to glucose repression of xylose utilization in *Bacillus megaterium*. *J. Bacteriol.* **179**, 7603-7605.
59. Yamazaki, M., Thorne, L., Mikolajczak, M., Armentrout, R.W. & Pollock, T.J. (1996). Linkage of genes essential for synthesis of a polysaccharide capsule in *Sphingomonas strain* S88. *J. Bacteriol.* **178**, 2676-2687.
60. Gibson, J., Dispensa, M., Fogg, G.C., Evans, D.T. & Harwood, C.S. (1994). 4-Hydroxybenzoate-coenzyme A ligase from *Rhodopseudomonas palustris*: purification, gene sequence, and role in anaerobic degradation. *J. Bacteriol.* **176**, 634-641.
61. Kwon, O., Bhattacharyya, D.K. & Meganathan, R. (1996). Menaquinone (vitamin K_2) biosynthesis: overexpression, purification, and properties of 0-succinylbenzoyl-coenzyme A synthetase from *Escherichia coli*. *J. Bacteriol.* **178**, 6778-6781.
62. Dittrich, H. & Kutchan, T.M. (1991). Molecular cloning, expression, and induction of berberine bridge enzyme, an enzyme essential to the formation of benzophenanthridine alkaloids in the response of plants to pathogenic attack. *Proc. Natl Acad. Sci. USA* **88**, 9969-9973.
63. Kagan, R.M. & Clarke, S. (1994). Widespread occurrence of three sequence motifs in diverse S-adenosylmethionine-dependent methyltransferases suggests a common structure for these enzymes. *Arch. Biochem. Biophy.* **310**, 417-427.
64. Shi, J., Gonzales, R.A. & Bhattacharyya, M.K. (1996). Identification and characterization of an S-adenosyl-L-methionine: Δ24-sterol-C-methyltransferase cDNA from soybean. *J. Biol. Chem.* **271**, 9384-9389.
65. Lacalle, R.A., Ruiz, D. & Jimenez, A. (1991). Molecular analysis of the *dmpM* gene encoding an O-demethyl puromycin O-methyltransferase from *Streptomyces alboniger*. *Gene* **109**, 55-61.
66. Madduri, K., Torti, F., Colombo, A.L. & Hutchinson, C.R. (1993). Cloning and sequencing of a gene encoding carminomycin 4-O-methyltransferase from *Streptomyces peucetius* and its expression in *Escherichia coli*. *J. Bacteriol.* **175**, 3900-3904.
67. Hidaka, T., Goda, M., Kuzuyama, T., Takei, N., Hidaka, M. & Seto, H. (1995). Cloning and nucleotide sequence of fosfomycin biosynthetic genes of *Streptomyces wedmorensis*. *Mol. Gen. Genet.* **249**, 274-280.
68. Hidaka, T., Hidaka, M., Kuzuyama, T. & Seto, H. (1995). Sequence of a P-methyltransferase-encoding gene isolated from a bialaphos-producing *Streptomyces hygroscopicus*. *Gene* **158**, 149-150.
69. Kuzuyama, T., Seki, T., Dairi, T., Hidaka, T. & Seto, H. (1995). Nucleotide sequence of fortimicin KL1 methyltransferase gene isolated from *Micromonospora olivasterospora*, and comparison of its deduced amino-acid sequence with those of methyltransferases involved in the biosynthesis of bialaphos and fosfomycin. *J. Antibiot.* **48**, 1191-1193.
70. Deppenmeier, U., Muller, V. & Gottschalk, G. (1996). Pathways of energy conservation in methanogenic archaea. *Arch. Microbiol.* **165**, 149-163.
71. Thauer, R.K., Hedderich, R. & Fischer, R. (1993). *Reactions and enzymes involved in methanogenesis from CO_2 and H_2*. Chapman and Hall, New York, NY.
72. Nolling, J., Pihl, T.D. & Reeve, J.N. (1995). Cloning, sequencing, and growth phase-dependent transcription of the coenzyme F_{420}-dependent N^5,N^{10}-methylenetetrahydromethanopterin reductase-encoding genes from *Methanobacterium thermoautotrophicum* ΔH and *Methanopyrus kandleri*. *J. Bacteriol.* **177**, 7238-7244.
73. Vaupel, M. & Thauer, R.K. (1995). Coenzyme F_{420}-dependent N^5,N^{10}-methylenetetrahydromethanopterin reductase (Mer) from *Methanobacterium thermoautotrophicum* strain Marburg. Cloning, sequencing, transcriptional analysis, and functional expression in *Escherichia coli* of the *mer* gene. *Eur. J. Biochem.* **231**, 773-778.
74. Stupperich, E., Juza, A., Hoppert, M. & Mayer, F. (1993). Cloning, sequencing and immunological characterization of the corrinoid-containing subunit of the N^5-methyltetrahydromethanopterin: coenzyme-M methyltransferase from *Methanobacterium thermoautotrophicum*. *Eur. J. Biochem.* **217**, 115-121.
75. Morbidoni, H.R., de Mendoza, D. & Cronan, J., Jr. (1996). *Bacillus subtilis* acyl carrier protein is encoded in a cluster of lipid biosynthesis genes. *J. Bacteriol.* **178**, 4794-4800.
76. Platt, M.W., Miller, K.J., Lane, W.S. & Kennedy, E.P. (1990). Isolation and characterization of the constitutive acyl carrier protein from *Rhizobium meliloti*. *J. Bacteriol.* **172**, 5440-5444.
77. Decker, H., Motamedi, H. & Hutchinson, C.R. (1993). Nucleotide sequences and heterologous expression of *tcmG* and *tcmP*, biosynthetic genes for tetracenomycin C in *Streptomyces glaucescens*. *J. Bacteriol.* **175**, 3876-3886.
78. Jabbouri, S., *et al.*, & Broughton, W.J. (1998). *nolO* and *noeI* (HsnIII) of *Rhizobium sp.* NGR234 are involved in 3-O-carbamoylation and 2-O-methylation of Nod factors. *J. Biol. Chem.* **273**, 12047-12055.
79. Luka, S., Sanjuan, J., Carlson, R.W. & Stacey, G. (1993). *nolMNO* genes of *Bradyrhizobium japonicum* are co-transcribed with *nodYABCSUIJ*, and *nolO* is involved in the synthesis of the lipo-oligosaccharide nodulation signals. *J. Biol. Chem.* **268**, 27053-27059.
80. Jabbouri, S., *et al.* & Broughton, W.J. (1995). Involvement of *nodS* in N-methylation and *nodU* in 6-O-carbamoylation of *Rhizobium sp.* NGR234 *nod* factors. *J. Biol. Chem.* **270**, 22968-22973.
81. Coque, J.J., Perez-Llarena, F.J., Enguita, F.J., Fuente, J.L., Martin, J.F. & Liras, P. (1995). Characterization of the *cmcH* genes of *Nocardia lactamdurans* and *Streptomyces clavuligerus* encoding a functional 3'-hydroxymethylcephem O-carbamoyltransferase for cephamycin biosynthesis. *Gene* **162**, 21-27.

82. Blattner, F.R., *et al.*, & Shao, Y. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453-1474.
83. Chater, K.F. & Bruton, C.J. (1985). Resistance, regulatory and production genes for the antibiotic methylenomycin are clustered. *EMBO J.* **4**, 1893-1897.
84. Caballero, J.L., Malpartida, F. & Hopwood, D.A. (1991). Transcriptional organization and regulation of an antibiotic export complex in the producing *Streptomyces* culture. *Mol. Gen. Genet.* **228**, 372-380.
85. Yang, K., Han, L. & Vining, L.C. (1995). Regulation of jadomycin B production in *Streptomyces venezuelae* ISP5230: involvement of a repressor gene, *jadR2*. *J. Bacteriol.* **177**, 6111-6117.
86. Otten, S.L., Ferguson, J. & Hutchinson, C.R. (1995). Regulation of daunorubicin production in *Streptomyces peucetius* by the *dnrR2* locus. *J. Bacteriol.* **177**, 1216-1224.
87. Guilfoile, P.G. & Hutchinson, C.R. (1992). The *Streptomyces glaucescens* TcmR protein represses transcription of the divergently oriented *tcmR* and *tcmA* genes by binding to an intergenic operator region. *J. Bacteriol.* **174**, 3659-3666.
88. Guilfoile, P.G. & Hutchinson, C.R. (1992). Sequence and transcriptional analysis of the *Streptomyces glaucescens tcmAR* tetracenomycin C resistance and repressor gene loci. *J. Bacteriol.* **174**, 3651-3658.
89. Grkovic, S., Brown, M.H., Roberts, N.J., Paulsen, I.T. & Skurray, R.A. (1998). QacR is a repressor protein that regulates expression of the *Staphylococcus aureus* multidrug efflux pump QacA. *J. Biol. Chem.* **273**, 18665-18673.
90. Mazodier, P., Petter, R. & Thompson, C. (1989). Intergeneric conjugation between *Escherichia coli* and *Streptomyces* species. *J. Bacteriol.* **171**, 3583-3585.
91. Tuan, J.S., Weber, J.M., Staver, M.J., Leung, J.O., Donadio, S. & Katz, L. (1990). Cloning of the genes involved in erythromycin biosynthesis from *Saccharopolyspora erythraea* using a novel Actinomycete–*Escherichia coli* cosmid. *Gene* **90**, 21-29.
92. Bierman, M., Logan, R., O'Brien, K., Seno, E.T., Rao, R.N. & Schoner, B.E. (1992). Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene* **116**, 43-49.
93. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual.* (2nd edn), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
94. Hopwood, D.A., *et al.*, & Schrempf, H.S. (1985). *Genetic Manipulation of Streptomyces: A Laboratory Manual.* John Innes Institute, Norwich, United Kingdom.
95. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410.
96. Devereux, J., Haeberli, P. & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**, 387-395.
97. Denis, F. & Brzezinski, R. (1992). A versatile shuttle cosmid vector for use in *Escherichia coli* and actinomycetes. *Gene* **111**, 115-118.
98. Baltz, R.H. (1980). Genetic recombination in *Streptomyces fradiae* by protoplast fusion and cell regeneration. *Dev. Ind. Microbiol.* **21**, 43-54.
99. Nishikohri, K. & Fukui, S. (1975). Biosynthesis of mitomycin in *Streptomyces caespitosus*. Relationship of intracellular vitamin B_{12} level to mitomycin synthesis and enzymatic methylation of a demethylated derivative of mitomycin. *Eur. J. Appl. Microbiol.* **2**, 129-145.
100. Summers, R.G., Wendt-Pienkowski, E., Motamedi, H. & Hutchinson, C.R. (1992). Nucleotide sequence of the *tcmII-tcmIV* region of the tetracenomycin C biosynthetic gene cluster of *Streptomyces glaucescens* and evidence that the *tcmN* gene encodes a multifunctional cyclase-dehydratase-O-methyl transferase. *J. Bacteriol.* **174**, 1810-1820.
101. Haydock, S.F., Dowson, J.A., Dhillon, N., Roberts, G.A., Cortes, J. & Leadlay, P.F. (1991). Cloning and sequence analysis of genes involved in erythromycin biosynthesis in *Saccharopolyspora erythraea*: sequence similarities between EryG and a family of S-adenosylmethionine-dependent methyltransferases. *Mol. Gen. Genet.* **230**, 120-128.
102. Niemi, J. & Mantsala, P. (1995). Nucleotide sequences and expression of genes from *Streptomyces purpurascens* that cause the production of new anthracyclines in *Streptomyces galilaeus*. *J. Bacteriol.* **177**, 2942-2945.
103. Dickens, M.L., Ye, J. & Strohl, W.R. (1995). Analysis of clustered genes encoding both early and late steps in daunomycin biosynthesis by *Streptomyces sp.* strain C5. *J. Bacteriol.* **177**, 536-543.
104. Grebenok, R.J., Galbraith, D.W. & Penna, D.D. (1997). Characterization of *Zea mays* endosperm C-24 sterol methyltransferase: one of two types of sterol methyltransferases in higher plants. *Plant Mol. Biol.* **34**, 891-896.
105. Bouvier-Nave, P., Husselstein, T., Desprez, T. & Benveniste, P. (1997). Identification of cDNAs encoding sterol methyl-transferases involved in the second methylation step of plant sterol biosynthesis. *Eur. J. Biochem.* **246**, 518-529.
106. Hardwick, K.G. & Pelham, H.R. (1994). SED6 is identical to ERG6, and encodes a putative methyltransferase required for ergosterol synthesis. *Yeast* **10**, 265-269.

Because *Chemistry & Biology* operates a 'Continuous Publication System' for Research Papers, this paper has been published via the internet before being printed. The paper can be accessed from http://biomednet.com/cbiology/cmb – for further information, see the explanation on the contents pages.