ELSEVIER

Chemistry & Biology 8 (2001) 59–69

Chemistry & Biology

www.elsevier.com/locate/chembiol

Research Paper

Novel features in a combined polyketide synthase/non-ribosomal peptide synthetase: the myxalamid biosynthetic gene cluster of the myxobacterium *Stigmatella aurantiaca* Sga15[1]

Barbara Silakowski [a,b], Gabriele Nordsiek [a], Brigitte Kunze [a], Helmut Blöcker [a], Rolf Müller [a,b,*]

[a] *Gesellschaft für Biotechnologische Forschung mbH (GBF), Mascheroderweg 1, 38124 Braunschweig, Germany*
[b] *Institut für Pharmazeutische Biologie, Mendelssohnstr. 1, Technische Universität, 38106 Braunschweig, Germany*

Received 28 August 2000; revisions requested 10 October 2000 revisions received 25 October 2000; accepted 7 November 2000
First published online 20 December 2000

Abstract

Background: Myxobacteria have been well established as a potent source for natural products with biological activity. They produce a considerable variety of compounds which represent typical polyketide structures with incorporated amino acids (e.g. the epothilons, the myxothiazols and the myxalamids). Several of these secondary metabolites are effective inhibitors of the electron transport via the respiratory chain and have been widely used. Molecular cloning and characterization of the genes governing the biosynthesis of these structures is of considerable interest, because such information adds to the limited knowledge as to how polyketide synthases (PKSs) and non-ribosomal peptide synthetases (NRPSs) interact and how they might be manipulated in order to form novel antibiotics.

Results: A DNA region of approximately 50 000 base pairs from *Stigmatella aurantiaca* Sga15 was sequenced and shown by gene disruption to be involved in myxalamid biosynthesis. Sequence analysis reveals that the myxalamids are formed by a combined PKS/NRPS system. The terminal NRPS MxaA extends the assembled polyketide chain of the myxalamids with alanine. MxaA contains an N-terminal domain with homology to NAD binding proteins, which is responsible during the biogenesis for a novel type of reductive chain release giving rise to the 2-aminopropanol moiety of the myxalamids. The last module of the PKS reveals an unprecedented genetic organization; it is encoded on two genes (*mxaB1* and *mxaB2*), subdividing the domains of one module from each other. A sequence comparison of myxobacterial acyl-transferase domains with known systems from streptomycetes and bacilli reveals that consensus sequences proposed to be specific for methylmalonyl-CoA and malonyl-CoA are not always reliable.

Conclusions: The complete biosynthetic gene cluster of the myxalamid-type electron transport inhibitor from *S. aurantiaca* Sga15 has been cloned and analyzed. It represents one of the few examples of combined PKS/NRPS systems, the analysis and manipulation of which has the potential to generate novel hybrid structures via combinatorial biosynthesis (e.g. via module-swapping techniques). Additionally, a new type of reductive release from PKS/NRPS systems is described.

Keywords: Biosynthetic gene cluster; Myxalamid; Myxobacterium; Reductive chain termination; Polyketide synthase/peptide synthetase

1. Introduction

Polyketides and non-ribosomal peptides represent large families of natural products that are assembled using acyl-coenzyme A or amino acid building blocks. Numerous chemical compounds belonging to these groups are widely used as pharmaceuticals, veterinary agents or agrochemicals. Both types are biosynthesized by extremely large polyfunctional enzyme systems, which are organized in modules. Each module is responsible for one discrete chain elongation step and can be subdivided into domains controlling the choice of the extender unit and several types of modifications that can take place on each intermediate. The responsible biosynthetic proteins are known as polyketide synthases (PKSs) and non-ribosomal polypeptide

[1] This article is dedicated to Prof. Dr. E. Leistner on the occasion of his 60th birthday.

* Correspondence: Rolf Müller;
E-mail: rom@gbf.de

1074-5521/01/$ – see front matter
PII: S1074-5521(00)00056-9

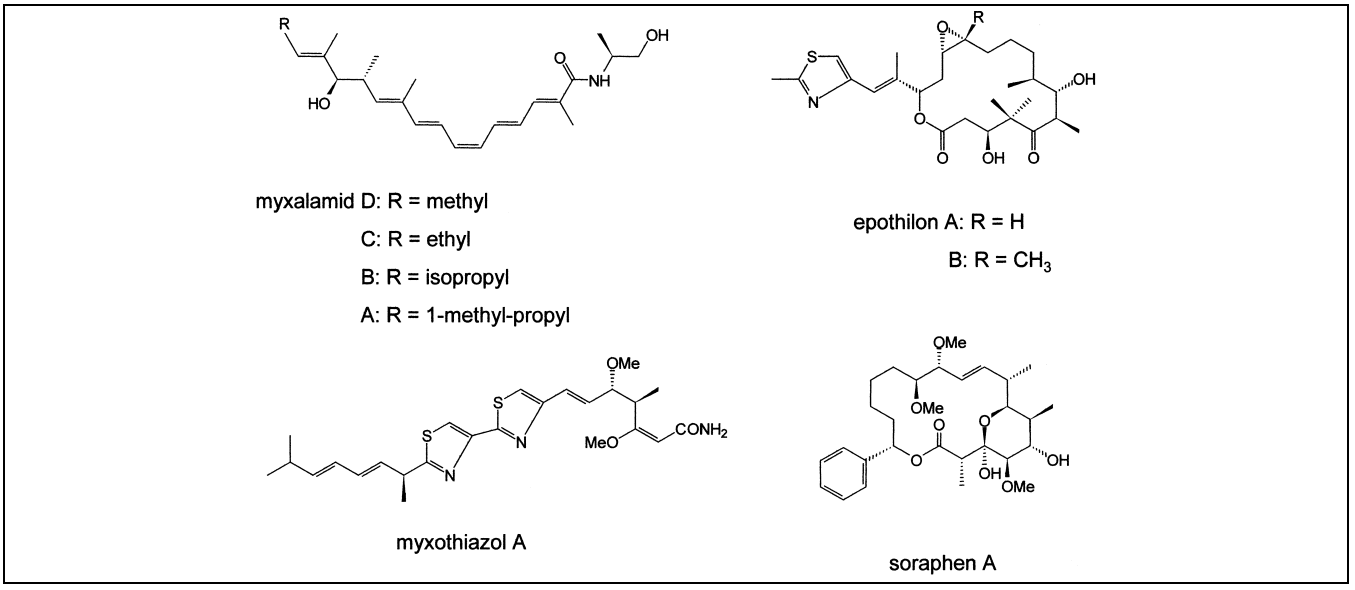

Fig. 1. Structures of natural products isolated from myxobacteria.

synthetases (NRPS) [1]. Until the first hybrid systems were described recently (see below), these biosynthetic machineries have been regarded as unrelated and studied separately.

Myxobacteria are Gram-negative soil bacteria that belong to the δ-group of the *Proteobacteria* [2]. They have been shown to produce an immense variety of secondary metabolites with biological activity (for reviews see [3,4]). Among these are the electron transport inhibitors of the myxothiazol type and the epothilons, which represent promising anticancer agents. These and several other myxobacterial secondary metabolites seem to be biosynthesized by combined PKS/NRPS systems (compare Fig. 1), a few of which have been recently cloned and sequenced [5–7]. From the rapidly growing knowledge of PKS and NRPS systems (for a review see [1]), which is based primarily on research with actinomycetes [8,9], only few hybrid PKS/NRPS systems have been described (rapamycin [10,11], naphthomycin [12], bleomycin [13]), whereas from myxobacteria and other alternative sources of natural products more information in terms of combinations and novel biosynthetic features has been recently gained. Currently four completely sequenced systems have been described, in which modules of PKS and NRPS were even found on single hybrid genes (yersiniabactin of *Yersinia pestis* [14], myxothiazol of *Stigmatella aurantiaca* [5], microcystin of *Microcystis aeruginosa* [15] and mycosubtilin of *Bacillus subtilis* [16]). Combinations of both types of systems have gained a lot of attention due to their potential to generate novel hybrid structures [17].

Myxothiazol [18] and the myxalamids [3,19] represent two electron transport inhibitors, which are produced by several strains of *S. aurantiaca* and both belong to the type of molecules described above. We recently reported the analysis of the gene cluster for myxothiazol [5] and identified a DNA fragment encoding part of a β-ketoacyl-synthase (KS) domain of the myxalamid biosynthetic gene cluster in a general screening approach aimed at the isolation of combined PKS/NRPS gene clusters from myxobacteria [20]. Additionally the myxalamids contain unusual features (Fig. 1): the isoleucine and valine derived starter units 2-methyl-butyryl-CoA and isobutyryl-CoA are used and the polyene-polyketide backbone is linear, terminating in a 2-amino-propanol structure derived from the amino acid alanine (compare [21]).

This communication deals with the characterization of the hybrid PKS/NRPS gene cluster from *S. aurantiaca* Sga15 that is involved in myxalamid biosynthesis. It comprises genes encoding PKS and NRPS, the inactivation of which impairs myxalamid biosynthesis.

2. Results and discussion

2.1. Cloning and identification of the myxalamid gene cluster via gene disruption

A cosmid library of *S. aurantiaca* Sga15 was prepared. Subsequently a variety of cosmids hybridizing to PKS and NRPS probes were isolated and used to amplify specific KS fragments of the PKS genes encoded on the respective DNA locus. One of these fragments, amplified from cosmid C2, was used for a gene inactivation experiment in *S. aurantiaca* Sga15 and resulted in a myxalamid negative phenotype [20]. The sequence of cosmid C2 was determined, revealing the presence of several open reading frames (ORFs) with similarity to PKSs which were designated *mxaB–D* (**myxa**lamid; see Fig. 2). In addition, one further ORF with similarity to NRPSs was detected (*mxaA*). In order to verify the expected involvement of

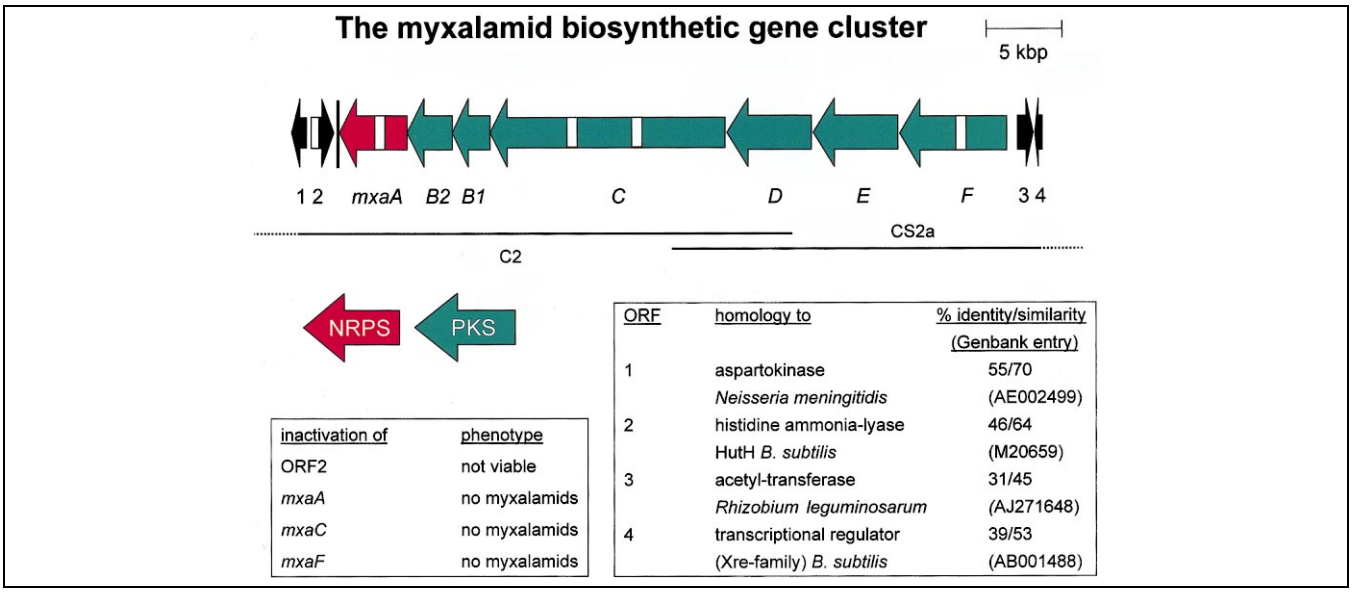

Fig. 2. Myxalamid biosynthetic gene cluster. □ show regions amplified via PCR which were used for the gene inactivation constructions.

mxaA in myxalamid biogenesis, a gene disruption mutant was made by insertion of a PCR generated fragment of *mxaA* into the chromosome of *S. aurantiaca* Sga15, resulting in mutant strain CBS17 (compare Fig. 2). The mutant was verified via Southern analysis (data not shown) and shown to be unable to produce myxalamids (see Fig. 3).

Subsequently a cosmid with minimized overlap was isolated from the library (cosmid CS2a, see Fig. 2) as described in Section 3. The sequence of the insert was analyzed as described and found to overlap with *mxaC*. Three further PKS genes oriented into the same direction as *mxaA–C* were identified and designated *mxaD–F*. Analysis of the modular structure of the myxalamid megasynthetase revealed that *mxaF* most likely encodes the first module of the biosynthetic gene cluster (see below). Thus *mxaF* was subjected to another gene inactivation experiment as described above and the verified mutant was found to be a myxalamid non-producing strain (see Figs. 2 and 3).

2.2. Sequence analysis of the myxalamid biosynthetic genes

The modular organization of PKSs involves activation and condensation of the following carboxylic acid onto the growing chain catalyzed by an acyl-transferase (AT) domain and a KS domain. The resulting β-keto acid may subsequently be processed by β-ketoacyl-reductase (KR) domains, β-hydroxy-acyl-dehydratase (DH) domains and enoyl-reductase (ER) domains (reviewed in [1]). Additional domains for C-methylation [14,22] and O-methylation [5] of intermediates have recently been reported. Sequence motifs typical for these domains in PKSs [10,23] and NRPSs [24,25] were detected in MxaA–F as shown in Table 1 and Fig. 4. The acyl carrier protein (ACP) domains and the peptidyl carrier protein (PCP) domain of MxaA–F contain the Prosite consensus signature of the

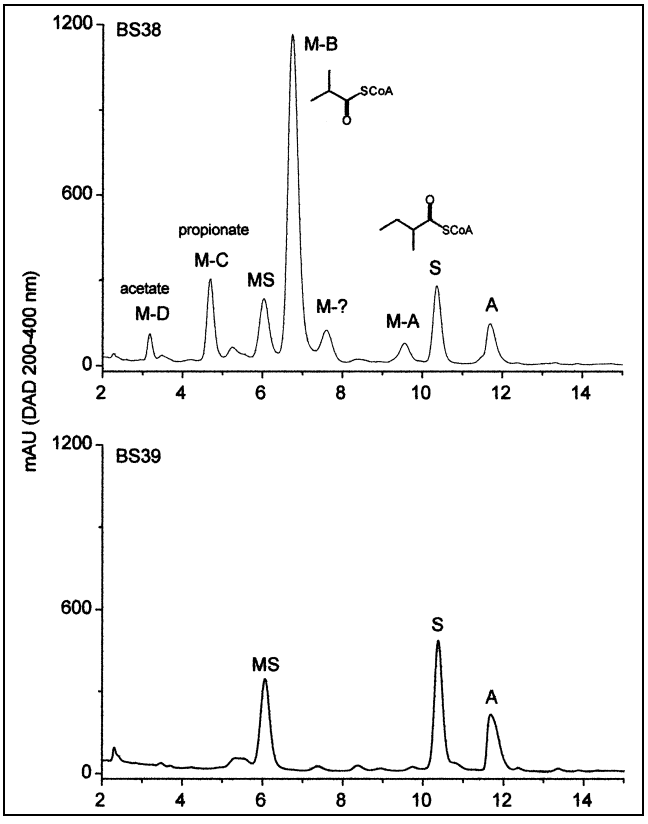

Fig. 3. HPLC analysis of mutants. The analysis of a kanamycin resistant mutant of *S. aurantiaca* Sga15 (strain BS38) grown under the same conditions as the myxalamid mutant BS39 is shown. M-A, myxalamid A; M-B, myxalamid B; M-C, myxalamid C; M-D, myxalamid D; M-?, unknown structural variant of the myxalamids; MS, myxochromid S; S, stigmatellin; A, aurachin A; DAD, diode array detection; mAU, milli-absorption units.

Table 1
Deduced functions of ORFs in the myxalamid biosynthetic gene cluster

Polypeptide (modules)	Amino acids	GC bias 1st, 2nd, 3rd position	Proposed function
MxaF	2368	71/50/86	PKS domains:
Loading+module 1:			ACP_L, KS, AT_L, AT_1, DH, KR, ACP_2
MxaE	1862	71/51/83	PKS domains:
Module 2:			KS, AT, DH, KR, ACP
MxaD	1840	73/49/83	PKS domains:
Module 3:			KS, AT, DH, KR, ACP
MxaC	3291	71/50/84	PKS domains:
Module 4:			KS, AT, DH, KR, ACP
Module 5:			KS, AT, KR, ACP
Module 6:			KS, AT, DH, KR, ACP
MxaB1	939	68/46/82	PKS domains:
Module 7:			KS, AT
MxaB2	1233	77/53/73	PKS domains:
Module 7:			DH, ER, KR, ACP
MxaA	1515	68/46/80	NRPS domains:
Module 8:			Condensation, adenylation, PCP, reduction

putative binding site for the 4′-phosphopantetheine (Ppant) co-factor (Prosite signature number PS00012, R2082 and L2104). The codon bias (see Table 1) of the genes reported is in accordance with other genes from myxobacteria [26]. The last module of the PKS part of the cluster is encoded on two genes, which to date is without precedent in other PKS systems. Initially we thought that this finding was due to a sequencing error and re-evaluated the DNA region, in which a frame shift would result in the module being encoded on one single ORF.

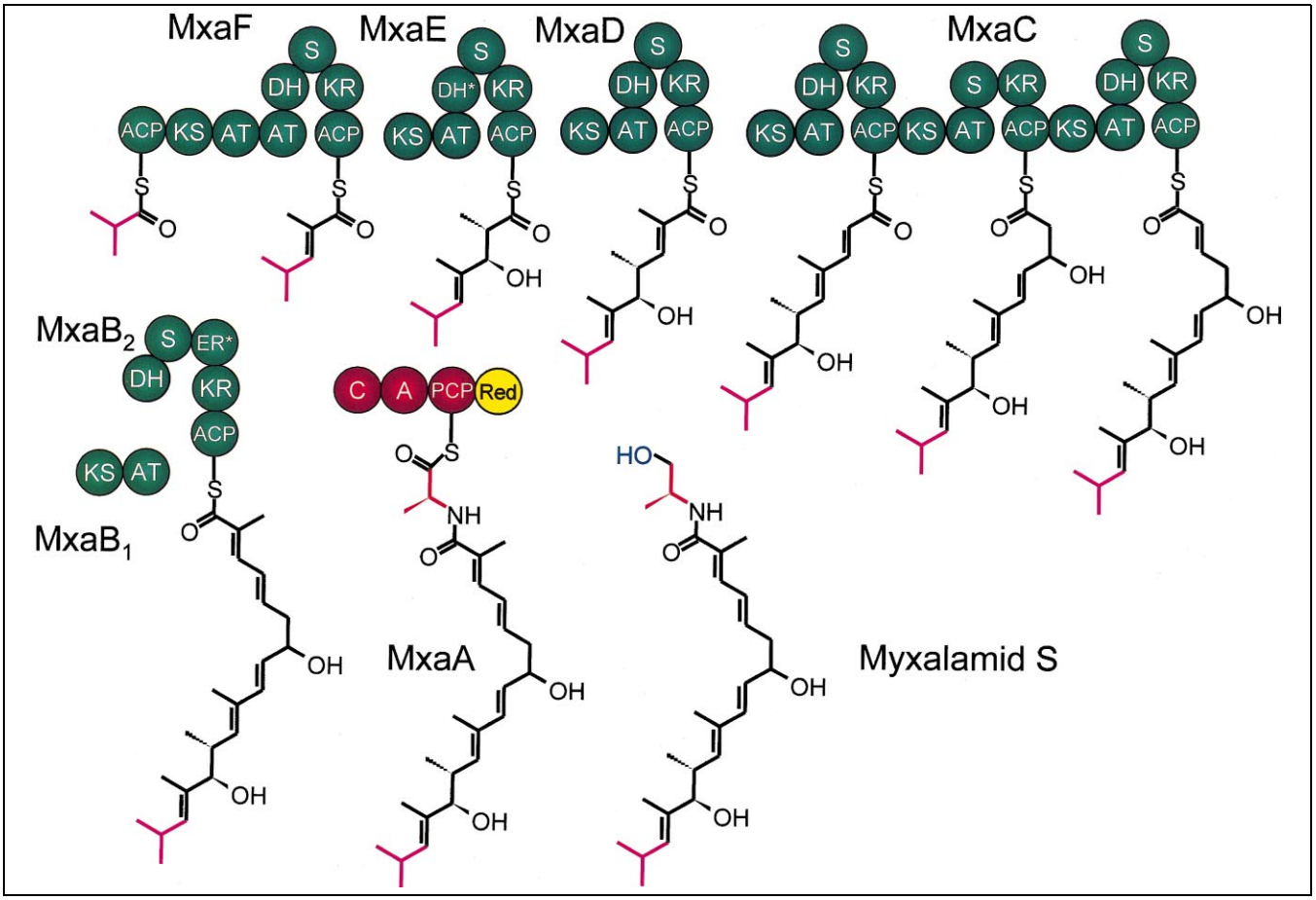

Fig. 4. Model for myxalamid biosynthesis. PKS domains shown in green, NRPS domains shown in red. The unusual terminal reductase domain (Red) of MxaA is indicated in yellow. '*' indicates a presumably inactive domain. The biosynthesis of myxalamid S is shown (see text). S, spacer region.

This region (1320 bp from the beginning of the AT domain in $mxaB_2$ to the end of the DH domain in $mxaB_1$) was covered by sequencing reactions from cosmid C4 at least four times on each DNA strand indicating no ambiguous region. In order to verify the sequence directly from the chromosome, the region was amplified from *S. aurantiaca* Sga15 via PCR using a proofreading polymerase and re-sequenced from four independently generated PCR products. No differences to the original sequence were found. The overall G+C content of the sequenced *mxa* region spanning approximately 55 kbp is 67%.

2.3. Model for the myxalamid biosynthetic pathway

The main component produced by *S. aurantiaca* Sga15, myxalamid B (compare Fig. 4), was shown to block the respiratory chain at the site of complex I in beef heart submitochondrial particles, i.e. NADH:ubiquinone oxidoreductase [19]. Here it is demonstrated, that the biosynthetic machinery for the myxalamids follows a multi-step process, resembling the biochemistry performed by type I PKS and NRPS systems. Alternative starters can be used to initiate the biosynthesis of the polyketide backbone finally resulting in myxalamids A–D (compare Figs. 1 and 4). The modular structure of type I PKSs usually starts with an AT or a CoA-ligase domain responsible for the recognition (and, in the case of CoA-ligases, for activation) of the starter molecule followed by transfer of the activated substrate to the first ACP domain (compare the biosynthetic gene clusters of erythromycin [23], rapamycin [10,27] and rifamycin [28]). In the case of *mxaF*, the modular organization looks different: the protein starts with an ACP_L, followed by KS, two ATs (AT_L and AT_1), DH, KR followed by the second ACP_2. The unusual type of AT arrangement in MxaF has only been reported twice in other proteins, which catalyze the first steps of polyketide biogenesis in two other myxobacterial compounds: soraphen (U.S. patent 5693774) from *Sorangium cellulosum* and myxothiazol [5] from *S. aurantiaca* DW4/3-1. In these, as in myxalamid, unusual starter molecules are used: activated forms of benzoic acid derived from phenylalanine in soraphen [29] and activated forms of 3-methylbutyrate derived from leucine in myxothiazol (see Fig. 1). We suppose that the first AT of MxaF (AT_L) is responsible for the recognition of the different starter units used for myxalamid biosynthesis (2-methylbutyrate in myxalamid A, isobutyrate in myxalamid B, propionate in myxalamid C and acetate in myxalamid D), whereas the second AT (AT_L) loads methylmalonate, which is subsequently used for the first chain extension step (compare Figs. 4 and 5). Whether the arrangement of the ATs is of consequence for the specificity of the unusual starter molecules and the three-dimensional structure of the PKS, remains to be determined. A comparison of the ATs of the myxalamid biosynthetic proteins with the domains of other known myxobacterial gene clusters (epothilon, myxothiazol, soraphen) is given in Fig. 5, indicating that the definition of core regions, presumed to be responsible for substrate specificity, is not always reliable. There are several ATs that do not fit into the consensus sequences published (e.g. [30,31]), the most striking examples being the malonate specific AT1 of MxaC and the methylmalonyl specific AT2 of the myxothiazol biosynthetic protein MtaB (compare Fig. 5). The difficulty in assigning specificities for ATs can be exemplified by a comparison of the AT domains of the epothilon biosynthetic genes, which were characterized by sequence comparison. AT2 of EpoC was assigned malonate specificity to via sequence comparison [6], and it was speculated that the methyl group at the corresponding position, C12 in the epothilon B structure (compare Fig. 1), is derived from an *S*-adenosyl-methionine dependent methylation. Nevertheless, it is clear from feeding experiments that the occurrence of the C12 methyl group in epothilon B (which does not occur in epothilon A; compare Fig. 1) is derived from alternative incorporation of malonate into epothilon A or methylmalonate into epothilon B (Gerth, K., Reichenbach, H. and Höfle, G., unpublished). This clearly shows that AT2 of EpoC must be able to accept both extender units. If one compares the homology of the ATs, one additionally finds that AT2 of EpoC (accepting methyl-malonate and malonate) and AT1 of EpoC (processing only malonate) are almost identical sequences, with just seven amino acids (aa) being exchanged in the 320 aa sequence, four of which are replaced by similar aa (Fig. 5). It has recently been proposed that an N-terminal 35 aa segment, encoded by an *Nhe*I-*Sty*I fragment of the DEBS AT2 domain or its counterparts in other AT domains, plays an important role in controlling the substrate specificity [32]. Nevertheless, there is only one threonine–alanine exchange in this aa stretch in the epothilon ATs mentioned. One highly conserved region within ATs is the GHSxGE motif, in which the x position is occupied by an isoleucine in AT1 and by an alanine in AT2 of EpoC. Whether any of these slight exchanges is responsible for the difference in substrate recognition remains to be determined, e.g. via site directed mutagenesis. Fig. 5B shows that almost all of the aa believed to specify malonate or methylmalonate recognition can be exchanged by their counterparts. In light of these exceptions to the 'specificity rules' it seems hardly possible to define new consensus sequences via sequence comparison alone. Further studies are necessary, ideally involving the crystallography of AT modules with their respective substrates. As in peptide synthetase systems [33], the KS and also possibly the ACP domains of polyketide synthases could play an additional role in substrate specificity.

Subsequent action of KR and DH domains during myxalamid biosynthesis can be assumed to result in the first double bond of the myxalamids (Fig. 4). Large spacer regions (S) more than 300 aa in size are located between the AT and the KR domains of each module. These have been described for almost all PKS and fatty acid biosyn-

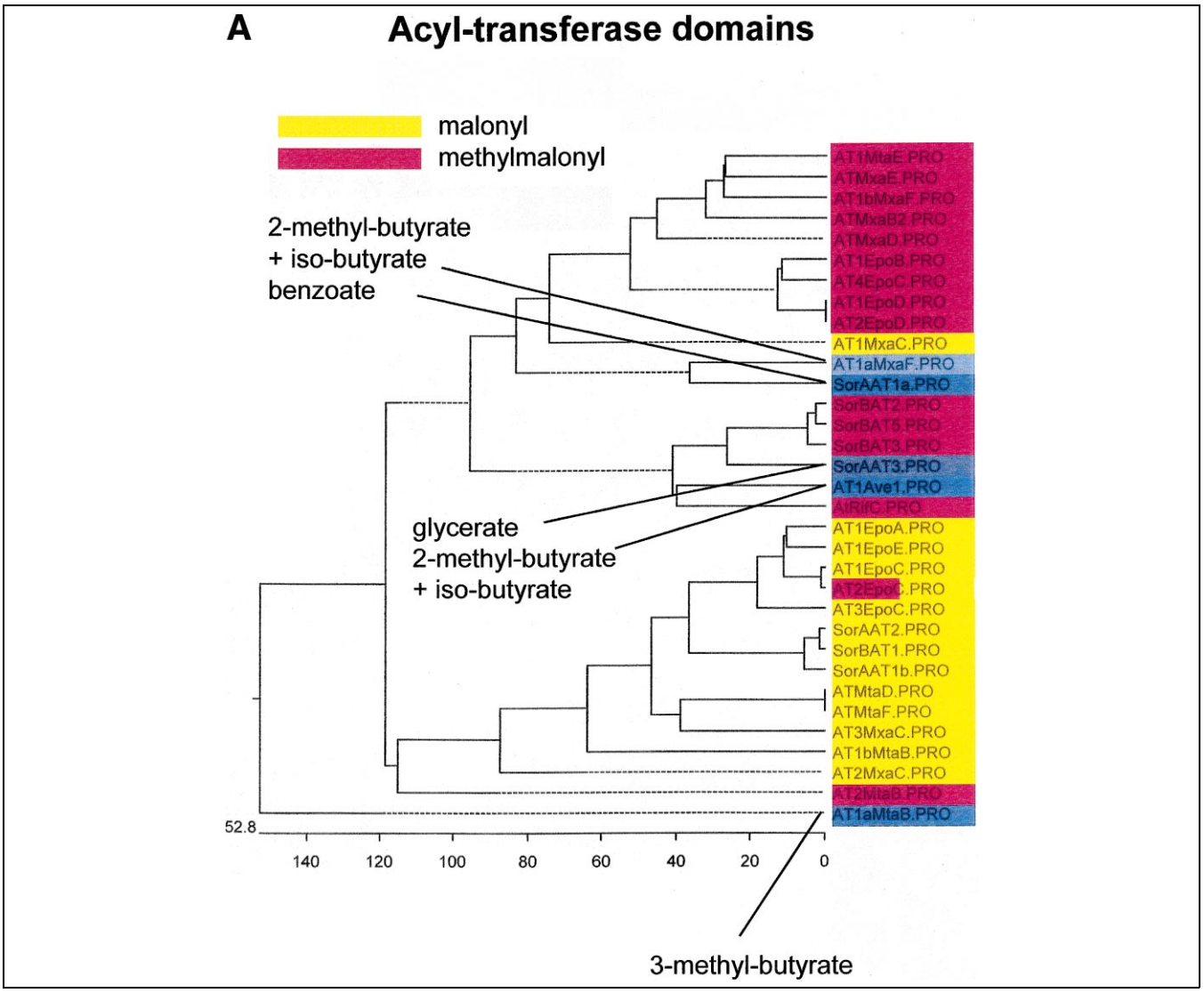

Fig. 5. Acyl-transferase domain alignment. A: Dendrogram of AT sequences from the epothilon (Epo; ref. see text), soraphen (Sor; ref. see text), myxothiazol (Mta; ref. see text), myxalamid (Mxa), avermectin (Ave; [31]) and rifamycin (Rif; [28]) biosynthetic proteins. Malonyl (m) and methylmalonyl (mm) specific ATs are shown in yellow and red, respectively. The ATs in blue are those with a different substrate specificity, which is referred to in the dendrogram. The length of each pair of branches represents the distance between sequence pairs, while the units at the bottom of the tree indicate the number of substitution events. A dotted line on a phenogram indicates a negative branch length. B: Alignment of the four AT domains discussed in the text. Boxes indicate the regions believed to determine substrate specificity [31], with the consensus sequence for each region given below the alignment. Amino acids in bold are those that are believed to specify malonate or methyl-malonate binding. Amino acids shown in red in the alignments do not fit into this hypothesis. Amino acids in green indicate the only differences between AT1 and AT2 of EpoC. '*' indicate identical, ':' similar aa.

thetic systems [1,34] but their function is not clear. These S regions can also be found in all myxobacterial PKS systems sequenced so far and show an identity of up to 35% on the aa level. In addition, they share up to 20% identical aa with ER domains from different sources. The S region present in module 1 even harbors an aa stretch (Lx**L**xxx**S**GVGxx**L**xxx**L**) which resembles the LxHxxxGGVGxxAxxxA motif important for functionality of ER domains [35]. Nevertheless, there is no indication that these S regions may function as ER domains, because myxalamid analogs without the corresponding double bond have not been isolated (Höfle, G.; personal communication). MxaE–C contain the following five PKS modules necessary for chain extension (compare Fig. 4). In addition, *mxaE* encodes a DH domain, which is not needed for the assembly of the product. Although the typical core region for DH domains (LxxHxxxGxxxxP) can be found, one has to assume that this domain is inactive. Interestingly, module 5 in MxaC lacks the DH domain necessary to create myxalamids A–D. Myxalamid S (compare Fig. 4) has also been isolated from *S. aurantiaca* Sga15 (Höfle, G., personal communication) representing the main compound (myxalamid B) derivative, in which water elimination does not occur in the corresponding position. Myxalamid S might be produced primarily and water elimination may be performed by an enzyme

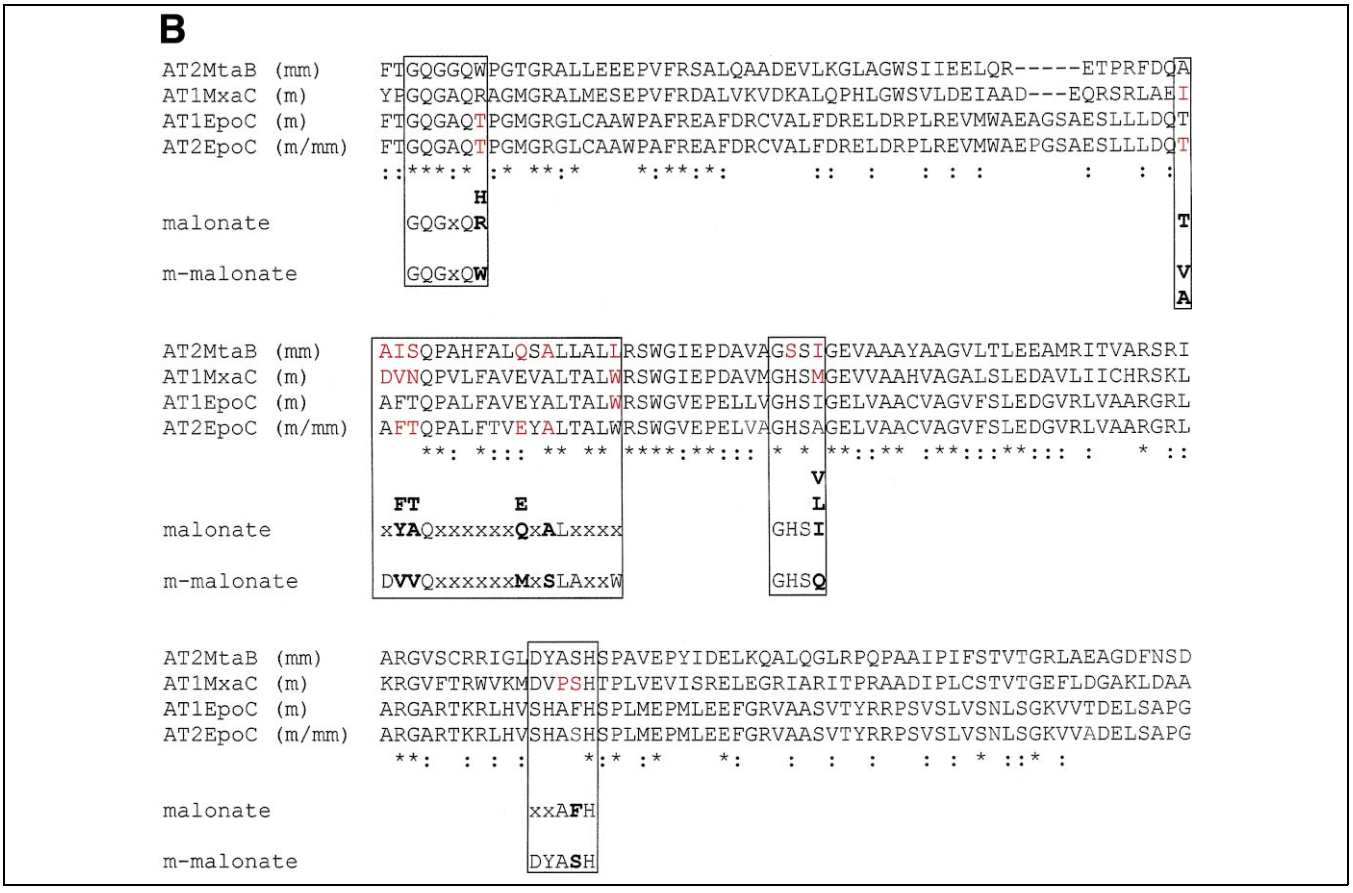

Fig. 5 (*continued*).

not encoded in the gene cluster, or by another DH domain of the myxalamid megasynthetase (possibly the one encoded in *mxaE*). It may even occur non-enzymatically giving rise to the conjugated pentadiene in the myxalamids A–D.

The growing chain is then extended by the $MxaB_{1-2}$ module, that is encoded on two separate proteins. One can envisage here how domain fusions take place via mutation, giving rise to type I systems. One single nucleotide insertion could combine the complete module on one ORF. $MxaB_{1-2}$ must interact during biogenesis, because the ACP domain necessary for a minimal module (KS-AT-ACP) is missing on MxaB1 and needs to be complemented by the ACP on MxaB2 (Fig. 4). Relatively hydrophilic amino-terminal linker regions resembling the interpolypeptide linkers defined recently [36] can be found in the sequences separating the non-covalently connected modules. These include the assumed interpolypeptide linker between MxaB2 and MxaB1. MxaB2 carries an inactive ER domain (**W**x**L**xxx**S**G**L**GxxAxxxA instead of the conserved LxHxxxGGVGxxAxxxA, see above).

Finally, the assembled polyketide is extended via MxaA, the terminal NRPS present in the system. The adenylation domain of MxaA carries all the core regions for A domains in NRPS [24] and most likely it is this domain which activates alanine, because feeding experiments have been performed using *Myxococcus stipitatus* Mxs40, which produces phenalamids. These natural products are myxalamid derivatives with a starter unit derived from phenyl acetate. In these experiments ^{13}C-labelled alanine was shown to be incorporated into the 2-amino-propanol moiety of the phenalamids [21], indicating that the same biochemistry is performed during myxalamid biosynthesis. Although structure-based models for the amino acid recognition by NRPSs exist, specificity pocket structures for alanine show weak homology [37,38]. During the review process of this manuscript a publication occurred, in which Shen and coworkers describe the bleomycin biosynthetic gene cluster. They redefine the consensus sequence of the alanine binding pocket [39], the updated version of which perfectly matches the sequence found in MxaA. The condensation domain of MxaA harbors all the hallmarks of a typical NRPS C domain [24] (C1: 232–262; C2: 370–396; C3: 619–648; C4: 844–864; C5: 1075–1107; C6: 1183–1203; C7: 1231–1254) with some amino acid changes, in comparison to the proposed core consensus sequences. Whether these are responsible for the substrate specificity of the myxalamid polyketide intermediate, that is extended by MxaA, remains to be determined. *mxaA* might be a good candidate gene to study PKS–NRPS

Reductase domains

	core R1	core R2	R3	core R4	core R5	core R6
	VLLTGATGYLG	VxxxVRA	GDL	VYPYxxLRxPNVxxT	GYxxSKWxxE	RPG
MxaA	**I**LLTGATGFLG	**I**YCLVR**S**	**G**D**I**	**L**YPYES**M**RA**A**NVLGT	GYAQSKWVAE	RP**E**
MxcG	VLLTGATGF**V**G	VVCLVRA	**A**DL	V**RE**YGSL**Q**A**T**NVRGT	GYQQSKWAAE	RPG
SafA	VLLTGATGYLG	VYCLVRA	GDL	VYPYSALRGPNVHGT	GYTGSKWVAE	n.D.
HetM	VFLTG**G**TGFLG	VYCLVRA	GDL	V**F**PYSAL**KA**ANVLGT	GYSQTKWAAE	n.D.
Lys2	VFVTGVTGFLG	VFAHVRA	GDL	VYPYAKLRDPNVIST	GYGQSKWAAE	n.D.

	Core R7
	LExxVGFLxxP
MxaA	**SD**SE**L**G**D**LLM**F**
MxcG	**YC**VG**Q**G**L**LQRP
SafA	**YL**VDIGFLQRP
HetM	**YL**IQ**T**GFLTL**A**
Lys2	n.D.

Fig. 6. Alignment of the core regions of reductase domains. Aligned are the core regions of the domains of the saframycin synthetase (SafA), of the probable polyketide synthase of *Anabaena* spec. (HetM; [49]), MxaA, MxcG [42] and of the α-aminoadipic acid reductase of *Saccharomyces cerevisiae* [41]. Differences to the core region [24] are indicated in bold. n.D. means that we were unable to detect the region in the sequence.

interaction, because alternative polyketide intermediates might be extended with alanine and reductively set free from the PCP using *mxaA* in combinatorial approaches.

The chain termination process catalyzed by MxaA is another interesting feature of the enzyme. The terminal reductase domain contains an NAD(P)H binding site similar to the one found in the Lys2 protein from *Saccharomyces cerevisiae* (compare Fig. 6), which was assumed to be characteristic for a family of aminoacyladenylate forming enzymes different from NRPS [40]. Biochemical characterization of the *lys2* gene product demonstrated that the protein is responsible for the reduction of α-aminoadipate to α-aminoadipate semialdehyde, which is followed by a transamination giving rise to lysine [41] (compare Fig. 7). We have cloned a second NRPS (MxcF) harboring such a reductase domain from the myxochelin biosynthetic operon of *S. aurantiaca* Sga15 [42]. Here it is assumed that the reductase domain sets free the aldehyde and might also be responsible for further reduction to the alcohol structure found in myxochelin A (compare Fig. 7). One further NRPS system is known to harbor an analogous domain; this is the saframycin megasynthetase from *Myxococcus xanthus* [43]. It is speculated that this enzyme reductively releases a polypeptide intermediate, setting free an aldehyde, that forms Schiff bases with free amino groups of the molecule, finally resulting in the formation of saframycin. During myxalamid biogenesis the reductase domain most likely performs the reductive release of the

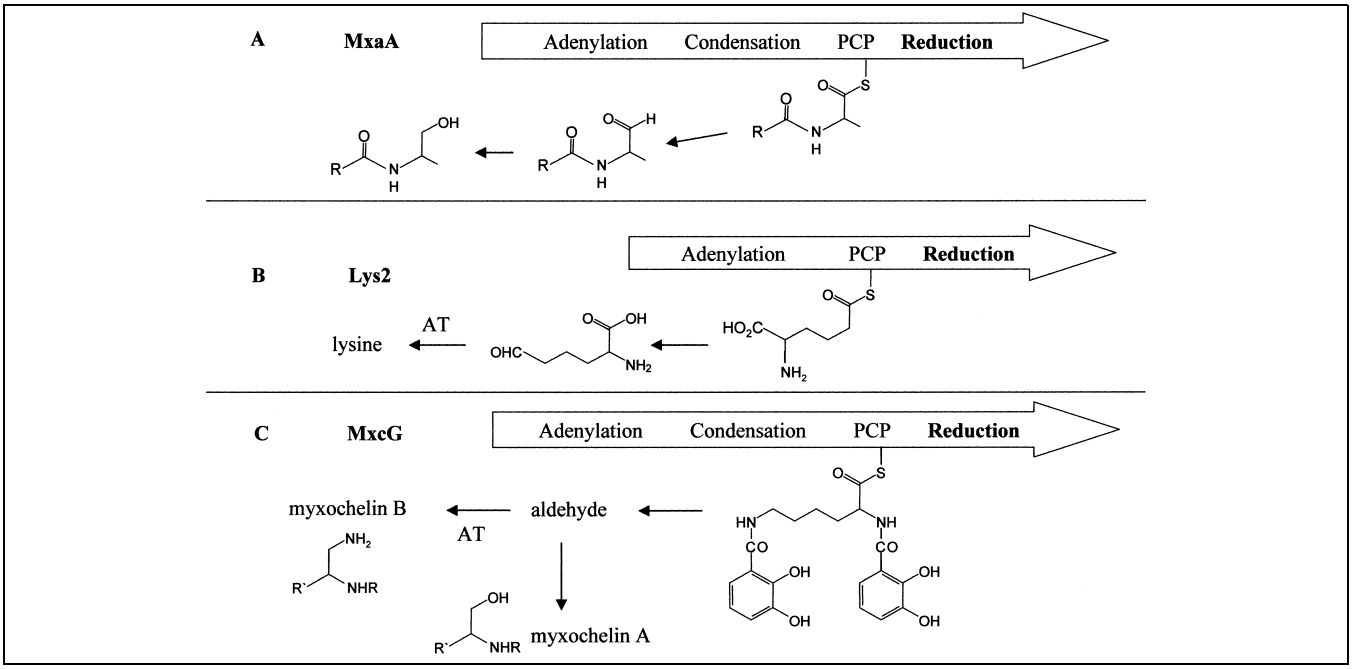

Fig. 7. Proposed function of the reductase domains in MxaA (A), Lys2 (B) and MxcG (C). AT, amino-transferase.

hybrid polyketide/peptide, giving rise to the linear aldehyde and also further reduces the aldehyde to the terminal 2-amino-propanol structure present in the natural product (compare Fig. 7). We sequenced downstream and upstream of the hypothetical ends of the myxalamid gene cluster (approximately 15 kbp in each direction) and did not detect any further gene with homology to oxidoreductases that could be responsible for the reduction of the aldehyde to the alcohol. Nevertheless, it cannot be excluded that the released aldehyde is further processed by an enzyme not encoded in the cluster. Interestingly, there is also one example for a reductase domain being present at the end of an uncharacterized PKS system (HetM of *Anabaena* spec.), where most likely the ACP bound thioester of the polyketide product is reductively released.

2.4. *Defining the border regions of the myxalamid biosynthetic gene cluster*

Analysing the sequence downstream of *mxaA* a putative stem loop structure was found, indicating transcriptional termination behind the gene (Fig. 2). Two further ORFs (ORF1 and ORF2) were identified, encoding genes with similarity to aspartokinases and histidine ammonia lyases, respectively (compare Fig. 2). ORF2 was subjected to gene inactivation experiments using pCBS21 but it was not possible to create mutants of the gene. This finding indicates that the gene encoding ORF2 is essential for growth and is not related to myxalamid biosynthesis.

Upstream of *mxaF* a region of 264 bp can be found which does not seem to encode any proteins. The next putative ORF identified (ORF3, compare Fig. 2) shows similarity to acetyl-transferases. The gene is divergently transcribed from *mxaF*, indicating that a tandem promotor is located within the 264 bp gap. Behind this gene ORF4 similar to several transcriptional repressors can be found. The encoding gene is divergently transcribed to ORF3. No role for any of the detected proteins in myxalamid biogenesis is obvious. Whether the transcriptional repressor has any relevance for the biosynthesis of the secondary metabolite remains to be elucidated.

3. Significance

Polyketide synthases (PKSs) and non-ribosomal peptide synthetases (NRPS) are large multifunctional and modular enzyme systems which are responsible for the biosynthesis of an immense variety of natural products with biological activity. Understanding the genetic basis and the biochemistry of these systems facilitates genetic engineering aimed at the rational design of novel structures with improved activity. Recently biosynthetic pathways using natural hybrids of PKS and NRPS systems have gained a lot of attention, because they enable the analysis of the co-operation of both enzyme types. The myxalamid biosynthetic gene cluster of *S. aurantiaca* Sga15 represents one of the few combined systems reported and analyzed to date. The gene cluster reported is not only a promising candidate to study further the interaction of PKS and NRPS, but also reveals several novel biosynthetic features.

4. Materials and methods

4.1. *Bacterial strains and culture conditions*

Escherichia coli strains and *S. aurantiaca* Sga15 and its descendants were cultured as described previously [44].

4.2. *DNA manipulations, analysis, sequencing and PCR*

Chromosomal DNA of *S. aurantiaca* strains was prepared as described [45]. PCR was carried out using HotStarTaq Polymerase (Qiagen) according to the manufacturer's protocol. DMSO was added to a final concentration of 5%. The conditions using the Eppendorf gradient mastercycler were as follows: 15 min at 95°C for activation of the polymerase, denaturation for 30 s at 95°C, annealing for 30 s at 57–60°C and extension for 45 s at 72°C; 30 cycles and a final extension for 10 min at 72°C. PCR products were purified with the high pure PCR product purification kit (Roche Molecular Biochemicals). Screening of the cosmid library of *S. aurantiaca* Sga15 [20] was performed as described [5]. The following primer pair was used to identify cosmid CS2a: CS2A1: 5′-CCCGAATGGCCGCGTCACACC-3′ and CS2A2: 5′-TCGGCATCAGCGGAACCAAC-3′. Clones giving amplification products of the expected size (465 bp) were counterselected with primer pair CS2A3: 5′-CCCGCTGATGCCGAAGGAAC-3′ and CS2A4: 5′-CGTGCGTGGCGAGGGCTGTGG-3′ to identify a cosmid with minimal overlap.

Sequencing of cosmids CS2 and CS2a was performed by a shotgun approach as follows: Sheared fragments of the two cosmids were subcloned separately into pTZ18R. At least 500 clones were selected from each cosmid library, plasmid DNA prepared (Millipore) and sequenced using DYEnamic™ ET terminator cycle sequencing premix kit (Amersham Pharmacia Biotech) and UPO/RPO primer (MWG-BioTech). The gels were run on ABI-377 sequencers and data were assembled and edited using the XGAP program [46]. In order to verify that *mxaB* is encoded by two genes, Pfu polymerase (Stratagene) was used to amplify a 1320 bp PCR product from chromosomal DNA of *S. aurantiaca* Sga15 using primers C24 (5′-CGGCTCGCCAGTTCCTCG-TAATG-3′) and C26 (5′-TGGGGCACAGCATGGGGGA-AGT-3′). Four independently amplified PCR products were sequenced and found to contain a DNA fragment identical to the analogous fragment from cosmid C2. Southern analysis of DNA was performed using the standard protocol for homologous probes of the DIG high prime DNA labelling and detection starter kit II (Roche Molecular Biochemicals). All other DNA manipulations were performed according to standard protocols [47]. Amino acid and DNA alignments were done using the programs

Clustal W [48] and the Lasergene software package (DNASTAR Inc.).

4.3. Construction and analysis of the S. aurantiaca Sga15 mutants

Two knock-out mutants unable to produce myxalamid have been described [20]. These were generated using KS fragments amplified by degenerate primers from cosmids hybridizing with PKS probes. These mutants (C2, C14) carry the integrated plasmid in the region of KS2 and KS3 of *mxaC*. Using the primers C3: 5′-AGGCTTGGCGGGCATGCTTGAGTA-3′ and C4: 5′-TCCGGAACGAGGCGCTGCTGTGTA-3′ and cosmid CS2 as template DNA, a fragment of *mxaA* (534 bp in size) was amplified and cloned into pCR2.1-TOPO using the TOPO TA cloning kit (Invitrogen). The resulting plasmid pCBS17 was electroporated after denaturation of the DNA into *S. aurantiaca* Sga15 [5] resulting in mutant CBS17. All other *S. aurantiaca* Sga15 mutants were constructed using the same approach. The *mxaF* mutant CBS39 was obtained after electroporation with plasmid pCBS39, which is a pCR2.1-TOPO derivative harboring a 590 bp fragment, amplified by primers RM174: 5′-GCCCGCACACGATCGTCAAGCACT-3′ and RM175: 5′-CTCGATCGCCGCCCGGAAGACAGG-3′ from cosmid CS2a.

We tried to obtain another set of mutants starting from a PCR product (1534 bp) of the complete histidine ammonium lyase-like gene generated using primers C15: 5′-CTCATATGTCCCGCCCCCGCTTGAACA-3′ and C16: 5′-GGCTGCAGCTGTCCGCGCGTCGCGGT-3′ cloned into the vector resulting in pCBS20. This plasmid was digested with *Stu*I/*Xho*I, sticky ends were blunted with Klenow enzyme and the fragment religated giving pCBS21. pCBS21 contains the 5′-end of the gene (760 bp) and was used for electroporation three times, but no transformants could be generated.

4.4. Production and analysis of myxalamids and further secondary metabolites from S. aurantiaca Sga15 and its descendants

For antibiotic production the mutants were cultivated in Zein liquid medium containing Zein 0.8%, peptone from casein tryptically digested 0.1%, $MgSO_4 \cdot 7H_2O$, 0.1%, kanamycin 60 μg/ml, HEPES buffer 50 mM, pH 7.2, and 1% of the adsorber resin AD-16 (Rohm and Haas). Batch cultures of 100 ml in 250 ml Erlenmeyer flasks were incubated at 30°C on a gyratory shaker at 160 rpm for 4–5 days.

After cultivation of the mutants, both the cell mass and the adsorber resin contained secondary metabolites and were extracted together twice with acetone. The spectrum of secondary metabolites produced by the different mutants was determined in aliquots of concentrated acetone extracts by diode array detected HPLC analysis using a Hewlett Packard series 1100 instrument. Chromatographic conditions were as follows: Column ET 125×2 mm and precolumn, Nucleosil120-5-C_{18}; solvent used: 0.2% aqueous acetic acid (A) and acetonitrile (B). For the first 5 min a mixture of solvent A and B (1:1) was used. Subsequently (5–15 min), the percentage of solvent B was increased in a gradient up to 80%; flow rate 0.5 ml/min; detection 200–400 nm.

The assignment of the different myxalamid structures was based on diode array coupled mass spectrometry which was also used to determine whether the mutants created were indeed myxalamid non-producers.

Nucleotide accession number: The nucleotide sequence reported here has been submitted to the EMBL database under accession number AF319998.

Acknowledgements

We thank Prof. Dr. H. Reichenbach and Prof. Dr. G. Höfle for their support at the GBF. B.S. and R.M. should like to thank the Deutsche Forschungsgemeinschaft for financial support via grant Mu1254/3-1. We acknowledge the skilfull technical assistance of M. Weilharter and the helpful comments of J. Scriven, S. Beyer and the reviewers regarding this manuscript.

References

[1] D.E. Cane, Polyketide and nonribosomal polypeptide biosynthesis, Chem. Rev. 97 (1997) 2463–2706.
[2] H. Reichenbach, Biology of the myxobacteria: ecology and taxonomy, in: M. Dworkin, D. Kaiser (Eds.), Myxobacteria II, American Society for Microbiology, Washington, DC, 1993, pp. 13–62.
[3] H. Reichenbach, G. Höfle, Production of bioactive secondary metabolites, in: M. Dworkin, D. Kaiser (Eds.), Myxobacteria II, American Society for Microbiology, Washington, DC, 1993, pp. 347–397.
[4] H. Reichenbach, G. Höfle, Myxobacteria as producers of secondary metabolites, in: S. Grabley, R. Thieriecke (Eds.), Drug Discovery from Nature, Springer Verlag, Berlin, 1999, pp. 149–179.
[5] B. Silakowski, R. Müller et al., New lessons for combinatorial biosynthesis from Myxobacteria: The myxothiazol biosynthetic gene cluster of *Stigmatella aurantiaca* DW4/3-1, J. Biol. Chem. 274 (1999) 37391–37399.
[6] I. Molnar, J.M. Ligon et al., The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from *Sorangium cellulosum* So ce90, Chem. Biol. 7 (2000) 97–109.
[7] L. Tang, B. Julien et al., Cloning and heterologous expression of the epothilon gene cluster, Science 287 (2000) 640–642.
[8] D. Hopwood, Genetic contributions to understanding polyketide synthases, Chem. Rev. 97 (1997) 2465–2497.
[9] D.A. Hopwood, D.H. Sherman, Molecular genetics of polyketides and its comparison to fatty acid biosynthesis, Annu. Rev. Genet. 24 (1990) 37–66.
[10] T. Schwecke, P.F. Leadley et al., The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin, Proc. Natl. Acad. Sci. USA 92 (1995) 7839–7843.
[11] I. Molnar, P.F. Leadley et al., Organisation of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of genes flanking the polyketide synthase, Gene 169 (1996) 1–7.
[12] S. Chen, E. Leistner et al., Biosynthesis of ansatrienin (mycotrienin) and naphtomycin: identification and analysis of two separate biosynthetic gene clusters in *Streptomyces collinus* TÜ 1892, Eur. J. Biochem. 261 (1999) 98–107.
[13] B. Shen, L. Du, C. Sanchez, M.T. Chen, D.J. Edwards, Bleomycin biosynthesis in *Streptomyces verticillus* ATCC15003: A model of hybrid peptide and polyketide biosynthesis, Bioorg. Chem. 27 (1999) 155–191.
[14] A.M. Gehring, R.D. Perry et al., Iron acquisition in plague: modular

logic in enzymatic biogenesis of yersiniabactin by *Yersinia pestis*, Chem. Biol. 5 (1998) 573–586.
[15] T. Nishizawa, M. Shirai et al., Polyketide synthase gene coupled to the peptide synthetase module involved in the biosynthesis of the cyclic heptapeptide microcystin, J. Biochem. 127 (2000) 779–789.
[16] E.H. Duitman, J. Vater et al., The mycosubtilin synthetase of *Bacillus subtilis* ATCC6633: a multifunctional hybrid between a peptide synthetase, an amino transferase, and a fatty acid synthase, Proc. Natl. Acad. Sci. USA 96 (1999) 13294–13299.
[17] D. Cane, C. Walsh, The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases, Chem. Biol. 6 (1999) R319–R325.
[18] K. Gerth, H. Irschik, H. Reichenbach, W. Trowitzsch, Myxothiazol, an antibiotic from *Myxococcus fulvus* (myxobacterales). I. Cultivation, isolation, physico-chemical and biological properties, J. Antibiot. 33 (1980) 1474–1479.
[19] K. Gerth, G. Thierbach et al., The myxalamids, new antibiotics from *Myxococcus xanthus* (myxobacterales). I. Production, physico-chemical and biological properties, and mechanism of action, J. Antibiot. 36 (1983) 1150–1156.
[20] S. Beyer, B. Kunze, B. Silakowski, R. Müller, Metabolic diversity in myxobacteria - identification of the myxalamid and the stigmatellin biosynthetic gene cluster of *Stigmatella aurantiaca* Sga15 and a combined polyketide–(poly)peptide gene cluster from the epothilon producing strain *Sorangium cellulosum* So ce90, Biochim. Biophys. Acta 1445 (1999) 185–195.
[21] W. Trowitzsch, E. Forche, V. Wray, H. Reichenbach, E. Jurkiewicz, G. Hunsmann, G. Höfle, Phenalamide, neue HIV-Inhibitoren aus *Myxococcus stipitatus* Mx s40, Liebigs Ann. Chem. 16 (1992) 659–664.
[22] J. Kennedy, K. Auclair, S.G. Kendrew, C. Park, J.C. Vederas, C.R. Hutchinson, Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis, Science 284 (1999) 1368–1372.
[23] S. Donadino, M.J. Staver, J.B. McAlpine, S.J. Swanson, L. Katz, Modular organization of genes required for complex polyketide biosynthesis, Science 252 (1991) 675–679.
[24] D. Konz, M.A. Marahiel, How do peptide synthetases generate structural diversity?, Chem. Biol. 6 (1999) R39–R48.
[25] H. Kleinkauf, H. von Döhren, Products of secondary metabolism, in: H. Kleinkauf, H. von Döhren (Eds.), Biotechnology, Vol. 2, Verlag Chemie, Weinheim, 1997, pp. 277–322.
[26] L. Shimkets, The myxobacterial genome, in: M. Dworkin, D. Kaiser (Eds.), Myxobacteria II, American Society for Microbiology, Washington, DC, 1993, pp. 85–108.
[27] I. Molnar, P.F. Leadlay et al., Organisation of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of genes flanking the polyketide synthase, Gene 169 (1996) 1–7.
[28] P.R. August, H.G. Floss et al., Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the *rif* biosynthetic gene cluster of *Amycolatopsis mediterranei* S699, Chem. Biol. 5 (1998) 69–79.
[29] G. Höfle, H. Reichenbach, The biosynthetic potential of the myxobacteria, in: W. Kuhn, H.P. Fiedler (Eds.), Sekundärmetabolismus bei Mikroorganismen, Attempo Verlag, Tübingen, 1995, pp. 61–78.
[30] S. Haydock, P.F. Leadley et al., Divergent structural motifs correlated with the substrate specificity of (methyl)malonyl-CoA:acylcarrier protein transacylase domains in the modular polyketide synthases, FEBS Lett. 374 (1995) 246–248.
[31] H. Ikeda, T. Nonomiya, M. Usami, T. Ohta, S. Omura, Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*, Proc. Natl. Acad. Sci. USA 96 (1999) 9509–9514.
[32] J. Lau, H. Fu, D.E. Cane, C. Khosla, Dissecting the role of acyltransferase domains of modular polyketide synthases in the choice and stereochemical fate of extender units, Biochemistry 38 (1999) 1643–1651.
[33] P. Belshaw, C. Walsh, T. Stachelhaus, Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis, Science 284 (1999) 486–489.
[34] S. Smith, The animal fatty acid synthase: one gene, one polypeptide, seven enzymes, FASEB J. 8 (1994) 1248–1259.
[35] S.J. Kakavas, L. Katz, D. Stassi, Identification and characterization of the niddamycin polyketide synthase genes from *Streptomyces caelestis*, J. Bacteriol. 179 (1997) 7515–7522.
[36] R.S. Gokhale, S.Y. Tsuji, D.E. Cane, C. Khosla, Dissecting and exploiting intermodular communication in polyketide synthases, Science 284 (1999) 482–485.
[37] G. Challis, J. Ravel, C. Townsend, Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains, Chem. Biol. 7 (2000) 211–224.
[38] T. Stachelhaus, H.D. Mootz, M.A. Marahiel, The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases, Chem. Biol. 6 (1999) 493–505.
[39] L. Du, C. Sanchez, M.T. Chen, D. Edwards, B. Shen, The biosynthetic gene cluster for the antitumor drug bleomycin from *Streptomyces verticillus* ATCC15003 supporting functional interactions between nonribosomal peptide synthetases and a polyketide synthase, Chem. Biol. 7 (2000) 623–642.
[40] J. Casqueiro, S. Gutierrez, O. Banuelos, F. Fierro, J. Veölasco, J. Martin, Characterization of the *lys2* gene of *Penicillium chrysogenum* encoding α-aminoadipic acid reductase, Mol. Gen. Genet. 259 (1998) 549–556.
[41] D. Ehmann, A. Gehring, C. Walsh, Lysine biosynthesis in *Saccharomyces cerevisiae*: Mechanism of α-aminoadipate reductase (Lys2) involves posttranslational phosphopantetheinylation by Lys5, Biochemistry 38 (1999) 6171–6177.
[42] B. Silakowski, B. Kunze, G. Nordsiek, H. Blöcker, G. Höfle, R. Müller, The myxochelin iron transport regulon of the myxobacterium *Stigmatella aurantiaca* Sga15, Eur. J. Biochem. 267 (2000) 6476–6485.
[43] A. Pospiech, J. Bietenhadeer, T. Schupp, Two multifunctional peptide synthetases and an O-methyltransferase are involved in the biosynthesis of the DNA-binding antibiotic and antitumour agent saframycin Mx1 from *Myxococcus xanthus*, Microbiology 142 (1996) 741–746.
[44] B. Silakowski, H. Ehret, H.U. Schairer, *fbfB*, a gene encoding a putative galactose oxidase, is involved in *Stigmatella aurantiaca* fruiting body formation, J. Bacteriol. 180 (1998) 1241–1247.
[45] B. Neumann, A. Pospiech, H.U. Schairer, Rapid isolation of genomic DNA from Gram negative bacteria, Trends Genet. 8 (1992) 332–333.
[46] J. Bonfield, K. Smith, R. Staden, A new DNA sequence assembly program, Nucleic Acids Res. 23 (1995) 4992–4999.
[47] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.
[48] J.D. Thompson, D.G. Higgins, T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Res. 22 (1994) 4673–4680.
[49] T.A. Black, C.P. Wolk, Analysis of a Het-mutation in *Anabaena* sp. strain PCC 7120 implicates a secondary metabolite in the regulation of heterocyst spacing, J. Bacteriol. 176 (1994) 2282–2292.