Production of the Tubulin Destabilizer Disorazol in *Sorangium cellulosum*: Biosynthetic Machinery and Regulatory Genes

Maren Kopp,[a] Herbert Irschik,[b] Silke Pradella,[b] and Rolf Müller*[a]

Myxobacteria show a high potential for the production of natural compounds that exhibit a wide variety of antibiotic, antifungal, and cytotoxic activities.[1,2] *The genus* Sorangium *is of special biotechnological interest because it produces almost half of the secondary metabolites isolated from these microorganisms. We describe a transposon-mutagenesis approach to identifying the disorazol biosynthetic gene cluster in* Sorangium cellulosum *So ce12, a producer of multiple natural products. In addition to the highly effective disorazol-type tubulin destabilizers,*[3–5] S. cellulosum *So ce12 produces sorangicins, potent eubacterial RNA polymerase inhibitors,*[6] *bactericidal sorangiolides, and the antifungal chivosazoles.*[7,8] *To obtain a transposon library of sufficient size suitable for the identification of the presumed biosynthetic gene clusters, an efficient transformation method was developed. We present here the first electroporation protocol for a strain of the genus* Sorangium. *The transposon library was screened for disorazol-negative mutants. This approach led to the identification of the corresponding trans-acyltransferase core biosynthetic gene cluster together with a region in the chromosome that is likely to be involved in disorazol biosynthesis. A third region in the genome harbors another gene that is presumed to be involved in the regulation of disorazol production. A detailed analysis of the biosynthetic and regulatory genes is presented in this paper.*

Introduction

Myxobacteria offer an alternative and rich source for natural compounds.[2] Most of them can be classified as polyketides, nonribosomally-made peptides, or a combination of both types of biosynthetic pathways. Studies over recent years have revealed the genetic capacity of myxobacteria to be even broader than one could expect from the known number of secondary metabolites.[1,9] Analysis of the myxobacterial genome projects has identified 18 biosynthetic gene clusters in the genome of *Myxococcus xanthus* and numerous biosynthetic gene clusters in the genome of *S. cellulosum* So ce56,[1,10] most of which have not been correlated with known products. Many of the secondary metabolites isolated from myxobacteria exhibit a wide spectrum of biological activities, for example, as antibiotics, antifungal, or cytotoxic substances. Of special interest are the compounds that interact with the cytoskeleton of the cells of higher organisms because they are of potential use for pharmaceutical applications. In myxobacteria five different groups of cytotoxic substances have been identified: chondramides and rhizopodins act on actin filaments,[11–13] whereas tubulysins,[14,15] epothilones, and disorazoles interact with the microtubule network.[16,17] The intensively studied epothilones stabilize microtubule formation and are in phase III clinical trials as anticancer agents. In contrast, the polymerization of tubulin is inhibited by disorazoles and tubulysins,[4,14] which are currently in preclinical studies. The concentration-dependent growth inhibition of L929 mouse fibroblasts by disorazol A1 has an extraordinary IC_{50} of 3 pM.[4]

The highly cytotoxic disorazoles are produced by the strain *S. cellulosum* So ce12.[3,5] The major compound disorazol A1 interferes with microtubule formation, blocks mitosis, and induces apoptosis (Figure 1).[4] As in most studies with natural products, the low yield of the compound causes problems in studying its mode of action and in carrying out further investigations that lead to preclinical development. The production of a complex composition of several derivatives causes additional purification problems. This is especially the case for *S. cellulosum* So ce12, which produces 29 derivatives of disorazoles.[5] Altering production in the natural host or, alternatively, producing the compound in a heterologous host have been explored as possible means for overcoming these problems. Both strategies, however, require the identification and isolation of the genetic information that directs biosynthesis (the biosynthetic gene cluster).

To obtain this information and to subsequently modify the genes, techniques for genetic manipulation have to be available. Combining genetic information and technique would enable the elucidation of disorazol biosynthesis and set the stage for bioengineering and heterologous expression. Nevertheless, for most myxobacterial strains genetic tools are poorly established. In *S. cellulosum* strains So ce90 and So ce26 a triparental mating method has enabled the identification of the soraphen and the epothilone biosynthetic gene clusters.[18–21]

[a] *Dipl.-Pharm. M. Kopp, Prof. Dr. R. Müller*
Saarland University, Pharmaceutical Biotechnology
Im Stadtwald, 66123 Saarbrücken (Germany)
Fax: (+49) 681-302-5473
E-mail: rom@mx.uni-saarland.de

[b] *Dr. H. Irschik, Dr. S. Pradella*
Department of Biology of Natural Products
Department Molecular Biology of Myxobacteria
German Research Centre Biotechnology
Mascheroder Weg 1, 38124 Braunschweig (Germany)

 DOI: 10.1002/cbic.200400459 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

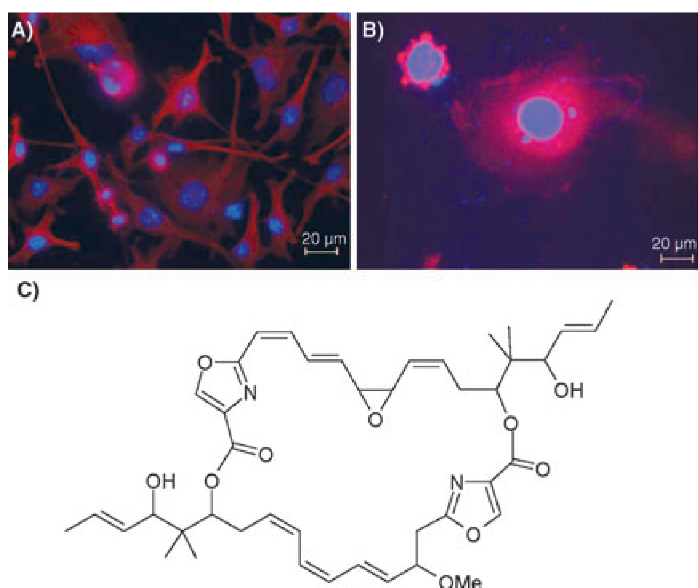

Figure 1. Structure and biological activity of disorazol. Effect of disorazol A1 after 24 h of incubation with interphase microtubules of mouse fibroblasts (L929 cell line). Microtubules were stained with anti-α-tubulin; nuclei were stained with DAPI. A) Control without disorazol A1; B) L929 cells after incubation with disorazol A1 (500 ng ml^{-1}); C) Structure of disorazol A1.

We recently reported a protocol for conjugational gene transfer into *S. cellulosum* So ce56, and a mariner-based transposon mutagenesis system for the disorazol-producing strain *S. cellulosum* So ce12 that is based on biparental mating experiments.[22,23] However, for practical reasons electroporation would be the more convenient and efficient transformation method. Electroporation has been shown to be possible for the transformation of other myxobacteria, such as *Myxococcus xanthus*,[24] *Stigmatella aurantiaca*, and *Angiococcus disciformis*,[25–27] but it has not been successfully used for any strains of the genus *Sorangium*.

Consequently, we established a method for the electroporation of *S. cellulosum* strains which was used to generate a *S. cellulosum* So ce12 transposon library. Following the screening of 1100 mutants we identified four disorazol-negative mutants. After vector recovery from the chromosome, different biosynthetic and regulatory genes of a highly unusual hybrid polyketide synthase (PKS)/nonribosomal peptide synthetase (NRPS) gene cluster were identified in a bacterial artificial chromosome (BAC) library and mapped to three different genomic regions.

Results

Generation of a transposon library from *S. cellulosum* So ce12 by electroporation

The chemical structure of the disorazoles led us to assume that PKS and NRPS are involved in disorazol biosynthesis.[5] Myxobacteria, and especially the genus *Sorangium*, are very prolific producers of secondary metabolites.[1] Since their genetic capacity is even broader than one could expect from the detected metabolites,[1,9] it did not seemed sensible to screen a gene library for the expected hybrid PKS/NRPS biosynthetic gene clusters with a mixed PKS/NRPS probe. As shown recently, the eukaryotic mariner-transposon system is applicable to *S. cellulosum* strains.[23,28] The genome of *S. cellulosum* So ce56 is reported to be around 12 Mbp.[22] Assuming a similar size for *S. cellulosum* So ce12, a transposon mutagenesis approach should identify one disorazol negative strain in approximately 250 mutants. This calculation is based on the number of biosynthetic modules expected to be involved in disorazol biosynthesis and results in a biosynthetic gene cluster approximately 50 kb in size.

Only about 300 mutants could be generated per conjugation experiment when using the method of transposon mutagensis previously described for *S. cellulosum* So ce12[23]. Several conjugations did not lead to any disorazol nonproducing mutant and the efficiency of the method could not be significantly increased.

Despite serious efforts in the past, no protocol for electroporation was available for any *Sorangium* species, but investigations carried out with *S. cellulosum* So ce56 (O. Perlova, K. Gerth, and R.M., unpublished results) led us to re-evaluate the electroporation conditions used for several strains of the genus *Sorangium*. Conditions for successful electrotransformation of *S. cellulosum* So ce12 were identified by varying parameters, such as the resistance and voltage. Using the mariner transposon, pMiniHimarHyg, we were able to integrate the DNA into the *S. cellulosum* So ce12 chromosome after electroporation for the first time. pMiniHimarHyg harbors the transposable element of pMycoMarHyg but it lacks the genes for conjugational DNA transfer. The reduced size of the electroporated plasmid was expected to increase the transformation efficiency.

Experience from previous conjugation experiments showed a reduced sensitivity of *S. cellulosum* So ce12 to the selection antibiotic hygromycin B when plating high quantities of cells.[23] Consequently, the cells were cultured in hygromycin B-containing medium prior to plating. Furthermore, the maximum amount of cells that can be plated on one agar plate had to be determined. A bioassay with the yeast, *Rhodotorula glutinis*, was used for the screening of *S. cellulosum* So ce12 transposon mutants for disorazol production. In addition to the identification of production-negative mutants we proved that the transposon randomly integrates into the chromosome. Figure 2 shows the Southern blot of randomly chosen *S. cellulosum* So ce12 transposon mutants probed with the digoxigenin-labeled hygromycin resistance cassette. Electroporation efficiency was determined to be $6–8\times10^{-7}$, based on the number of cells used per experiment. Approximately 1100 colonies were screened for disorazol production. Four disorazol nonproducing mutants were identified (strains So12_EX_2793, So12_EX_13-3, So12_EX_13-21, So12_EXI_IE-2) in the bioassay and the results were confirmed by using DAD-HPLC (Figure 3) and HPLC-MS (data not shown). Recovery of the transposon together with the adjacent sequence from these four mutants resulted in the plasmids pTn-Rec_2793, pTn-Rec_13-3, pTn-Rec_13-21, and pTn-Rec_IE-2. These plasmids harbor the hygromy-

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

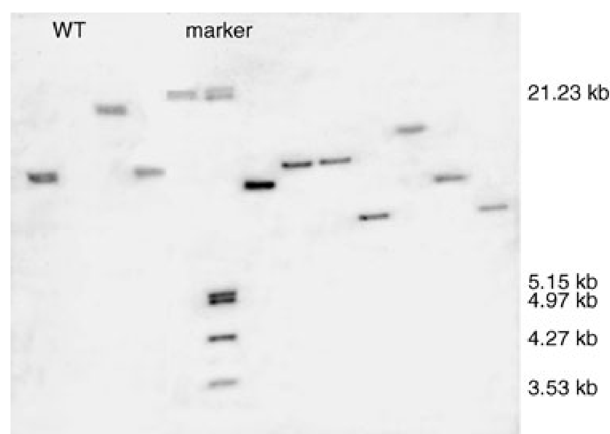

Figure 2. Southern blot of *S. cellulosum* S ce12 transposon mutants. Chromosomal DNA was digested with *Mlu*I and probed with the digoxigenin labeled hygromycin resistance cassette. WT: chromosomal DNA from wild type *S. cellulosum* So ce12; marker: digoxigenin-labeled DNA molecular weight marker.

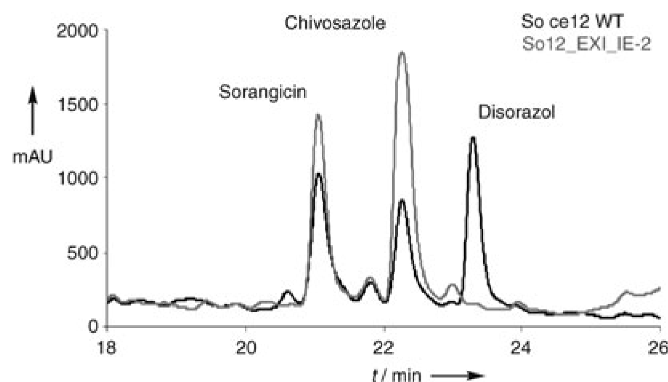

Figure 3. HPLC chromatogram of culture extracts of *S. cellulosum* So ce12 wild type (So12_WT) and transposon mutant So12_EXI_IE-2. Disorazol was identified by comparison to authentic reference standards (retention time, UV, mass spec). Disorazol masses and their absence in the disorazol negative mutants was also verified by HPLC-MS analysis (data not shown).

cin resistance gene and the λpir-dependent origin of replication oriR6K together with parts of *S. cellulosum* So ce12 chromosomal DNA that originally flanked the transposition site. The sequences of these chromosomal regions identified the inactivated genes. Analysis of the sequences revealed all four transformants to be unique and blast searches identified two of the predicted proteins as fragments of a PKS and a NRPS (Table 1).

In order to sequence the corresponding gene cluster a BAC library comprising 1920 BACs (each containing *S. cellulosum* So ce12 chromosomal DNA in *E. coli*) was constructed. Statistical analysis of the BAC clones revealed that the library represents at least 96 Mbp of cloned chromosomal DNA. Assuming that the genome size of *S. cellulosum* So ce12 approximately corresponds to that of So ce56,[22] the BAC library represents at least an eight-fold coverage of the *S. cellulosum* So ce12 chromosome. The PKS and the NRPS fragments obtained from the transposon recovery were hybridized with the BAC library. Eight BAC clones that hybridized with both probes could be isolated. After determination of the end sequences, one of the BACs was completely sequenced; 86.06 kb of the approximately 140 kb insert was sequenced with $>99.99\,\%$ accuracy. This region shows an average GC content of 72.17 %. During the preparation of this manuscript, parts of the sequence reported here were also published in the patent WO04053065A2.

Four genes with homologies to PKS and NRPS encoding genes were designated *disA–D* (see Table 2 and Figure 4). Typical PKS and NRPS domain sequences could be identified by comparison with databases.[29–31] The core biosynthetic gene cluster starts with *disA,* which is preceded by a putative ribosomal binding site (RBS; GGAAA) and is located 11 bp upstream of the designated start codon (GTG); *disB* presumably starts with an ATG and has a putative RBS (GGGG) which appears to be located 7 bp upstream of this start codon. The last gene of the putative transcriptional unit, *disC*, encodes a mixed PKS/NRPS. The gene most likely starts with an ATG that is preceded by a putative RBS (GAGGA) located 8 bp upstream. 36 bp downstream of the first putative start codon of *disC* an alternative start codon (TTG) can be found. A probable transcriptional terminator is located downstream of this gene. The inverted repeats are found at positions 63671–63698 and 63704–63731 of the sequence. This finding indicates independent transcription of the downstream genes which include *disD*.

Downstream of *orf9*, encoding a hypothetical protein (Table 2), *disD*, was identified and a putative ribosomal binding site (GAGGA) is located 7 bp upstream of the designated ATG start codon.

This gene shows significant similarities to the bifunctional proteins LnmG, from the leinamycin biosynthetic gene cluster, and MmpIII, from the mupirocin biosynthetic gene cluster.[32,33] DisD exhibits two discrete domains (Table 2 and Figure 4); the putative acyltransferase (AT) domain (InterPro: IPR001227) at

Table 1. Recovered plasmids and the proposed function of the proteins encoded on the inactivated genes.

Plasmid	Proposed Function of the Similar Protein	Source	Identity/Similarity
pTn-Rec_2793	BarG (NRPS) barbamide biosynthetic gene cluster	*Lyngbya majuscula*	39 %/57 %
pTn-Rec_13-3	5′ to transposition site: no prediction		
	3′ to transposition site: carbamoyltransferase BlmD	*Rhodopirellula baltica* SH1	28 %/45 %
pTn-Rec_13-21	LnmJ (PKS) Leinamycin biosynthetic gene cluster	*Streptomyces atroolivaceus*	29 %/40 %
pTn-Rec_IE-2	β-lactamase	*Oceanobacillus iheyensis*	38 %/53 %
	putative esterase	*Rhodopirellula baltica* SH 1	30 %/48 %

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 2. Proteins encoded within the sequenced region including the disorazol biosynthetic gene cluster and their putative function.

Protein (Gene)	Size (Da/bp)	NRPS and PKS Part of the Gene Cluster – Proposed Function (Protein Domains with their Position in the Sequence)
DisA (*disA*)	647772/18036	PKS Domains: KS1 (3–428), DH1 (953–1144), KR1(1528–1779), ACP1 (1821–1889), KS2 (1971–2395), KR2 (2856–3105), MT2 (3225–3463), ACP2 (3537–3606), ACP2b (3672–3741), KS3 (3779–4201), KR3 (4642–4898), ACP3 (4918–4987), KS4 (5059–5490), DH4 (5649–5878)
DisB (*disB*)	672408/18771	PKS Domains: KR4 (238–492), ACP4 (547–615), KS5 (676–1114), DH5 (1274–1476), KR5 (1836–2093), ACP5 (2108–2176), KS6 (2255–2686), DH6 (2944–3149), KR6 (3490–3738), ACP6 (3776–3824), KS7 (3876–4304), DH7 (4472–4679), KR7 (5049–5302),ACP7 (5316–5398), KS8 (5500–5926), ACP8 (6123–6192)
DisC (*disC*)	409960/11379	NRPS Domains: HC1a (58–506), HC1b (532–955), A1 (1035–1551), PCP1 (1580–1647), OX (1649–1836), PKS Domains: KS9 (1882–2309), ACP9 (2542–2609), KS10 (2668–3098), ACP10 (3399–3468), TE (3521–3701)
DisD (*disD*)	90953/2526	PKS-Domains: AT (1–280), OR (393–839)

ORFs Encoded Upstream and Downsteam of *disA–disD*

Gene	Size (Da/bp)	Orientation (strand)	Proposed Function of the Similar Protein	Similarity to Source	Similarity/Identitiy	Acc. No. of the Similar Protein
orf1	49316/1374	–	*blr*4832 Sugar (and other) transporter	*Bradyrhizobium japonicum* USDA 110	49%/67%,	NC_004463.1
orf2	51696/1449	–	probable two-component response regulator, signal receiver domain	*Pseudomonas aeruginosa* PAO1	49%/66%	NC_002516.1
orf3	45545/1293	+	hypothetical protein	*Leptospira interrogans serovar Lai str.* 56601	27%/40%	NC_004342.1
orf4	56119/1641	–	no prediction			
orf5	48994/1371	–	probable two-component response regulator, signal receiver domain	*Pseudomonas aeruginosa* PAO1	51%/69%	NC_002516.1
orf6	105961/3021	–	sensory box histidine kinase	*Pseudomonas putida* KT2440	39%/55%	NC_002947.3
orf7	34954/975	–	phosphotransferase	*Escherichia coli*	29%/40%	Q47395
orf8	37435/1053	+	putative serine/threonine protein kinase	*Streptomyces avermitilis* MA-4680	33%/48%	NC_003155.2
disA–C		+				
orf9	30717/822	+	no functional prediction			
disD		+				
orf10	23476/642	–	phosphotransferase	*Bacillus subtilis subsp.* subtilis str. 168	38%/56%	NC_000964.2
orf11	46773/1287	+	putative sugar transporter	*Streptomyces avermitilis* MA-4680	27%/41%	NC_003155.2
orf12	32992/912	+	ABC membrane transporter homologue	*Brevibacterium fuscum var. dextranlyticum*	36%/53%	Q93RD7
orf13	31993/882	+	ABC membrane transporter homologue	*Brevibacterium fuscum var. dextranlyticum*	51%/72%	Q93RD6
orf14	86590/2355	+	putative sugar hydrolase	*Streptomyces coelicolor* A3(2)	61%/72%	NP_733521
orf15	105005/2892	+	putative sugar hydrolase	*Streptomyces coelicolor* A3(2)	46%/59%	NP_629813
orf16	121293/3273	+	serine-threonine protein kinase	*Mycobacterium leprae* TN	36%/53%	NP_301681
orf17	23384/642	–	no prediction			
orf18	35402/999	–	no prediction			
orf19	25075/657	–	no prediction			

the N terminus of the protein shows high similarity to known AT domains and the conserved AT core motif (GxSxG) can be found.[34] Analysis of the amino acids responsible for substrate specificity suggests that malonyl-CoA is the preferred substrate.[35] The C terminus sequence of DisD is similar to proteins of the oxidoreductase superfamily.

Analysis of the domain organization of *disA–C* (see Figure 4 and Table 2) identified ten PKS and one NRPS module without internal AT domains. The latter are known to be necessary for the function of a minimal PKS module. The acyl carrier protein (ACP) domains (except for ACP2b) and the peptidyl carrier protein (PCP) domain contain the consensus motif of the putative binding site for the 4′-phosphopantetheine (Ppant) cofactor.[31,36]

The first seven PKS modules followed by an NRPS module appear to be sufficient for the incorporation of the seven acetate units and one serine. In silico analysis of all ketosynthase (KS) domains showed that KS8, KS9, and KS10 lack one of the highly conserved histidine residues of the KS core regions, as shown in Figure 5.[29] Module 2 of *disA* encodes a methyltransferase (MT) domain which has the conserved *S*-adenosylmethionine (SAM) binding motifs described in the literature.[37,38] The NRPS module encodes tandem heterocyclization (HC) domains. Both HC domains include the known NRPS core motifs.[31]

Analysis of the adenylation domain in the NRPS module of DisC identified serine as the most likely candidate for activation (amino-acid residues: DVWHFSLV) according to the nonribosomal code.[39,40]

An oxidation (OX) domain is found downstream of the PCP domain of DisC. This domain shows both core motifs defined by Du et al.[41]

The genes adjacent to the 5′- and 3′-end of *disA–D* were analyzed for potential involvement in disorazol biosynthesis. Table 2 shows similar proteins from the database as determined with blastp.

In addition to the two transpositions into PKS- and NRPS-encoding genes, which were used for the identification of the core disorazol biosynthetic gene cluster, two further knockouts (strains So12_EX_13-3 and So12_EXI_IE-2) could be gener-

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

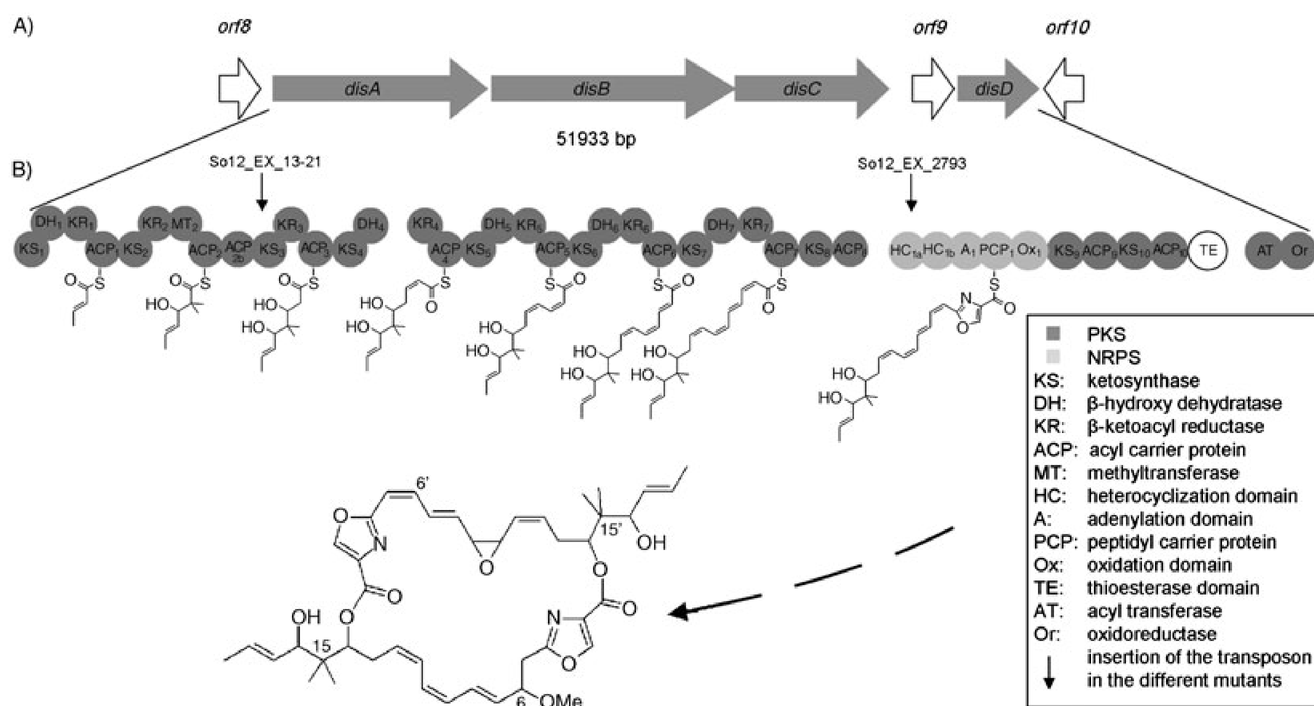

Figure 4. Model of disorazol biosynthesis. A) The organization of the disorazol biosynthetic gene cluster including the adjacent genes *orf8* and *orf10*. The orientation of all the genes at the 5′ end of *disA* and 3′ end of *disD* are indicated in Table 2. B) Modular organization of the genes *disA–D*, together with the intermediates, according to our proposal for disorazol biosynthesis. The insertions of the transposon in the different mutants are marked with ↓.

cons.	DxxCSSxL	H	H
KS1	IAIDAACASSLVAL	YVEAHGTGTSL	SNIGHLEAAA
KS2	LAIDTACSSSLVAV	YIECHGTGTAL	SNLGHLEAAA
KS3	VPVDTACSSSLMAI	YVETHGTGTKL	TGIGHTLAAS
KS4	EPINTACSSALIAV	YIETHGTGTAL	TNVGHLEAAA
KS5	IASDTACSSSLVAI	YVEAHGTGTEL	SNVGHLEGAA
KS6	LTVSTACSSSLLAV	YVEAHGTGTAL	SNIGHLESAA
KS7	VALDTMCSSSLTAI	YVEAHGTGTSL	SNVGHLESAA
KS8	MAVDTTCASSLTAI	YVEA**A**ANGAPL	SNIGHLEGAS
KS9	VAIDTSCSSGLTAL	YVEC**A**ATGSGI	PNIGHLESAS
KS10	LAVDTACASSLTAL	AYIE**V**AANGSS	SNIGHPEAAS

Figure 5. Analysis of the KS domain core-regions. The alignment shows the regions of KS domain core-motifs of KS1–10 of DisA–C. Highly conserved amino acids are shaded in grey, deviations from the core motif are marked in bold.

ated. PCR analysis of all *dis*-PKS/NRPS positive BACs described above with primers that bind to the sequences obtained from the recovered plasmids pTn-Rec_13-3 and pTn-Rec_IE-2, revealed that the inactivated genes are not located on the sequenced BAC and therefore are not part of the core biosynthetic gene cluster. Due to the similarity of the gene product mutated by the transposon in strain So12_EXI_IE-2 to esterases, the complete recovered plasmid was also sequenced. This provided information about all the genes located next to the esterase gene. Figure 6 and Table 3 show the analysis of the plasmid and the similarities of the encoded proteins to others found in the database.

Similarly, the genomic region corresponding to the fourth gene inactivation was cloned, and sequences derived from the region adjacent to the transposon did not reveal genes that could be directly involved in disorazol biosynthesis (see Table 1). Further sequencing identified a gene at the 3′ end of the cloned region that encodes a putative carbamoyltransferase.

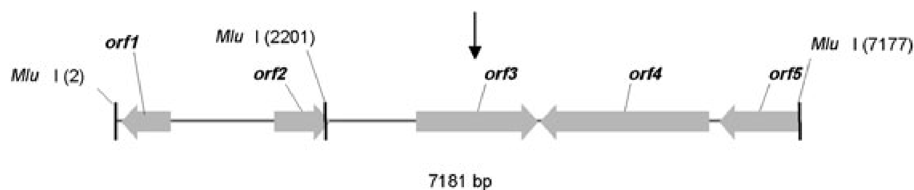

Figure 6. Organization of the genes encoded adjacent to the transposition site of mutant So12_EXI_IE-3 that was cloned into the recovered plasmid pTn-Rec_IE-2. The insertion site of the transposon is marked with ↓. *Mlu*I restriction sites used for the recovery of the plasmid are indicated.

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 3. Proteins encoded on the recovered plasmid pTn-Rec_IE-2 and their putative function in disorazol biosynthesis.

Gene	Size [Da/bp]	Proposed Function of the Similar Protein	Source	Similarity/ Identity	Acc. No. of the Similar Protein
orf1	18008/522	arylesterase-related protein	*Caulobacter crescentus*	29%/43%	NP_422235
orf2	20979/591	SAM-dependent methyltransferase	*Gloeobacter violaceus*	48%/58%	Q7NEB0
orf3	46369/1284	putative esterase β-lactamase	*Rhodopirellula baltica* SH1 *Oceanobacillus iheyensis*	35%/51%	NP_691588
orf4	62063/1782	adenylate cyclase 2	*Stigmatella aurantiaca*	31%/51%	NP_711174
orf5	29564/854	outer membrane protein (incomplete)	*Myxococcus xanthus*	36%/46%	Q84FF8

transposon recovery (Table 1) revealed that no genetically identical mutants were found. Thus, no duplication of the transposon mutants occurred, which might have been expected due to the long incubation times prior to plating. Among the 1100 analyzed strains, four disorazol negative mutants could be identified which corresponds to the expected ratio in this statistical approach. The genetic verification of unique transposition sites in all of these mutants proves the reliability of the described transposon based approach.

Discussion

Development of a protocol for the electroporation of *S. cellulosum* So ce12

To date, a number of investigations have been carried out concerning the molecular basis for the biosynthesis of secondary metabolites in strains of the genus *Sorangium*.[1,18–20,23] However, this work is restricted to a few strains due to the lack of established tools for genetic manipulation of these organisms. A protocol for conjugational gene transfer into *S. cellulosum* strains was established by Jaoua et al.[42] Recently it was shown that this protocol is not applicable to numerous other *S. cellulosum* strains.[23] Furthermore, this method suffers from a low transformation efficiency compared to electroporation,[22,23,42] which has been shown to be a more convenient method for genetic manipulation. Electroporation has been used successfully for the transformation of members of the *Cystobacterineae* suborder of myxobacteria which does not include the genus *Sorangium*; the latter belongs to the *Sorangineae*.[24,27] Here we report the first electroporation protocol for strains from the genus *Sorangium* which is described in detail under the European patent application EP 04103546.0 (Irschik, K. Gerth, Kopp, Perlova, and Müller). Several parameters were optimized to establish the protocol described herein which generated *S. cellulosum* So ce12 transposon mutants with an efficiency of $6–8\times 10^{-7}$, based on the initial number of cells used. From the analysis of our experiments several factors appear to be important for successful electroporation. At least 20 h (equal to more than two doubling times) of phenotypical expression were used to enable transposition into the *S. cellulosum* So ce12 genome and expression of hygromycin resistance. Shorter incubation times which were used to prevent the cells from subsequent duplication did not give any transformants. In contrast, other myxobacterial strains could be electroporated with shorter incubation times (e.g., 3–6 h for *M. xanthus*, which equals approximately one doubling period). Problems related to unspecific growth with hygromycin B as selection marker could be overcome by incubating the cells in selection medium after the phenotypical expression.[23] Analysis of the generated mutants by Southern blot analysis (Figure 2) and

disA–D encode a hybrid NRPS/PKS cluster of the trans-AT type

The genes *disA–C* encode ten PKS modules and one NRPS module. Sequence analysis suggests that malonyl-CoA and serine are incorporated extender molecules.[35,39,40] This is in good agreement with previous feeding experiments that determined disorazol to be formed from acetate and serine.[5] The most unusual feature of the disorazol biosynthetic gene cluster is the presence of only one discrete AT domain on *disD*. Similar PKS organizations in so-called trans-AT type I PKS were previously described for some other biosynthetic gene clusters that have been completely sequenced, like the pederin,[43,44] leinamycin,[32,45,46] lankacidin, and mupirocin biosynthetic gene clusters.[33,47]

The discrete acyltransferase in these biosynthetic systems is assumed to load the malonyl-CoA extender units to all of the ACP domains in the PKS modules, which has been demonstrated for the AT domain of the leinamycin biosynthetic gene cluster.[45]

It is believed that such trans-AT gene clusters are extremely rare. Nevertheless, including the disorazol genes, nine trans-AT biosynthetic gene clusters have been reported—at least in part—since the first was completely sequenced two years ago.[32,33,43,45–52]

As we have additional data that show the biosynthesis of at least two more myxobacterial secondary metabolites is directed by trans-AT biosynthetic gene clusters (M.K., O. Perlova, and R.M. unpublished results), we believe that this type of PKS gene cluster represents a transition between the different canonical forms of PKS types, the classification of which needs to be viewed with caution.[53]

Analysis of the sequence downstream of each KS domain in *disA–C* revealed the presence of regions that exhibit high homologies among themselves and also to AT domains. Recently this was also shown for the biosynthetic gene clusters of leinamycin, pederin, and mupirocin and seems to be a characteristic feature of trans-AT type PKSs.[46] It is assumed that these regions are derived from functional AT domains from multiple deletions, which lead to the loss of the highly conserved active sites and their catalytic activities. Due to their homology to functional AT domains it was speculated that they could serve

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

as AT docking sites for the interaction with trans-AT. However, to date there has been no functional evidence presented to support this hypothesis.

The function of the oxidoreductase domain joined to this AT domain remains unclear. It is notable that trans-AT clusters often offer such oxidoreductases as separate proteins (pederin) or as a domain in multidomain proteins, as found in leinamycin or mupirocin.

Hypothetical pathway to disorazol in *S. cellulosum* So ce12

Theoretically, seven malonyl-CoA units have to be attached to the starter acetate unit, and serine incorporation leads to one half of the disorazol bis-lactone core unit. Due to the amino-acid substitutions in the conserved core regions of KSs 8, 9, and 10 (Figure 5), we assume that these domains and their corresponding modules are inactive. Thus, the ACP of module 7 most likely transfers the polyketide intermediate directly to the NRPS module of DisC (Figure 4), a hybrid PKS/NRPS. Interestingly, most of the published hybrid proteins are derived from myxobacteria, for example, EpoB, MtaD, MelD, and TubD.[19,27,54,55]

Analysis of the enzymatic domains in module 1–7 leads to a biosynthetic hypothesis (see Figure 4) that is in good agreement with one half of the disorazol bis-lactone. DisA does not harbor a loading domain, which would be expected to transfer acetyl-CoA or malonyl-CoA (and subsequently decarboxylate the thioester-bound molecule) to a loading ACP. We assume that biosynthesis starts with the condensation of the loading acetate, which is attached to an unidentified ACP, to malonyl-S-ACP1 on DisA, catalyzed by the KS1 of DisA. Alternatively, KS1 might use acetyl-CoA directly. The reduction of the β-keto intermediate is catalyzed by ketoreductase (KR) 1 and dehydratase (DH) 1 domains of module 1 which gives rise to the corresponding acrylyl intermediate. The identified domains of module 2 (KS-KR-MT-ACP-ACP) catalyze the incorporation of the second malonyl-CoA and the subsequent reduction to the β-hydroxy intermediate. The SAM-dependent MT present in module 2 is expected to incorporate the two methyl groups that are attached to the C15 and C15′ of disorazol, and which were shown to be derived from SAM.[5] Since ACP2b lacks the active-site serine, the intermediate will presumably be attached to ACP2. Module 3 harbors KS, KR, and ACP domains, which would lead to a β-hydroxy intermediate.

The next module (4) is split between DisA and DisB—KS and DH domains are found on DisA and seem to be complemented on DisB with the initial KR and ACP domains. Such a split domain organization between two proteins is unusual, but not unprecedented.[56] A similar organization was found in the myxalamid biosynthetic gene cluster where MxaB1 and MxaB2 harbor the module 7 domains (KS and ACP on MxaB1, DH, ER, KR, and ACP on MxaB2).

We have seen more examples of such noncanonical PKSs in other myxobacterial biosynthetic systems (R.M., unpublished results). Module 4 in the *dis*-system catalyzes the introduction of another malonyl-CoA which gives rise to the next acrylyl intermediate that is used by the downstream modules of DisB. Module 5, 6, and 7 on DisB share the same assembly compared to module 4. Due to these three further extensions, polyketide biosynthesis gets to the stage of an acrylyl intermediate that offers four conjugated double bonds; that is, bound to ACP7. We hypothesize that this intermediate is transferred to the NRPS part of the cluster. Necessary modifications of bis-lactone will be discussed below.

The corresponding NRPS module on DisC has an unusual tandem heterocyclization (HC) domain, monomers of which were shown to catalyze the formation of heterocycles from cysteine or serine.[57] According to the nonribosomal code, the A domain of DisC should incorporate serine which corresponds well with oxazole ring formation.[39,40] After extension with cysteine, thiazole rings can be formed, which was demonstrated during epothilone biosynthesis.[58] Condensation of the disorazol intermediate with the serine molecule leads to an amide bond. This might be catalyzed by one HC domain, whereas the second HC domain would cyclize this intermediate to give rise to an oxazoline ring. The unusual existence of two HC domains was previously shown in VibF of the vibriobactin biosynthetic gene cluster and LnmI of the leinamycin biosynthetic gene cluster.[45,59] Site directed and deletion mutagenesis of VibF showed that both HC domains work independently on the different steps of heterocyclization.[60] Whether the two HC domains of DisC function similarly remains to be analyzed. Alternatively, both steps might be performed by one HC domain and the second one would be inactive. To assign a putative function of condensation and/or cyclization/dehydration to the HC domains, alignments with both C and HC domains of other biosynthetic gene clusters (e.g., *tub, epo, mta, vib*) were performed. The alignments did not indicate a greater degree of similarity of C domains to either one of the DisC HC domains (data not shown).

The Ox domain located downstream of the PCP domain of DisC is expected to subsequently oxidize the oxazoline intermediate to the final oxazole ring. In addition to DisC, examples for the oxidation of an intermediary thiazoline or oxazoline ring can be found in the epothilone synthetase, EposP/EpoB,[19,20] the bleomycin synthetase, BlmIII,[41] the myxothiazol synthetases, MtaC and MtaD,[55] the leinamycin synthetase, LnmI, and the tubulysin synthetase, TubD.[27,45] The Ox domains of EposP and MtaD are located within the A domain of the NRPS module. Since the Ox domains of BlmIII and MtaC are located downstream of the PCP domain and the functionality for these could not be proved,[55] it was speculated that Ox domains located in this position might be inactive. Nevertheless, analysis of the leinamycin and tubulysin biosynthetic gene clusters identified two further Ox domains that are located downstream of the LnmI and TubD PCP domain, and the structure of both secondary metabolites demands that these Ox domains be active. These results suggest that Ox domains in both positions could be active.

After the incorporation and cyclization of the serine moiety no further extension of the carbon chain is needed to form one half side of the disorazol bis-lactone. Since KS domains 9 and 10 are most likely inactive (Figure 5), the PCP bound intermediate might not be extended further and could be released from the chain by the thioesterase (TE) domain encoded on

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

disC. The biosynthetic formation of the final product is currently unknown.

Whether the TE of DisD will dimerize the primary biosynthetic product as shown for enterobactin or bacillibactin biosyntheses,[61,62] or if an alternative enzyme will catalyze this bis-lactone formation, remains to be investigated. Within the core biosynthetic gene cluster no enzyme could be identified that might be involved in the formation of the two ester bonds found in the final product. In light of the results reported below, it might well be possible that other enzymes encoded elsewhere in the chromosome are involved in the heterodimerization process.

The modifications required to build the known 29 derivatives of disorazol cannot be explained with the set of enzymes encoded within the core biosynthetic gene cluster. The derivatives offer epoxidations, hydroxylations, and methylations of the resulting OH-groups as well as acylations.[5] These are expected to be introduced after PKS assembly. Analysis of the DH domains in DisA and DisB predicts all of them to be active which would lead to an initial intermediate with four conjugated double bonds, which can be correlated with all known derivatives by post PKS biosynthetic steps. Even derivatives with four conjugated double bonds were reported. Taken together, these observations make a lack of DH functionality, unlikely.

Another interesting feature is the unusual formation of *cis*-double bonds by the action of KR and DH domains as present in rifamycin and epothilone.[63,64]

Analysis of KR domains revealed conserved amino-acid residues to be responsible for the stereochemistry of the resulting alcohol.[65] Based on these results, in silico investigation of the Dis KR domains indicate that KR2, KR4, KR5, and KR7 catalyze the formation of an L-configured alcohol. The subsequent dehydration catalyzed by modules 4, 5, and 7 leads to the *cis*-double bonds found in most of the disorazoles. The fact that derivatives representing other *cis–trans* isomers are known is presumably due to unspecific epimerization reactions.

Analysis of the 5′- and 3′- region of the biosynthetic gene cluster

The chromosomal regions adjacent to the disorazol core-biosynthetic gene cluster, were examined with respect to a potential involvement in the "missing" steps of disorazol biosynthesis. The results of the in silico analysis are shown in Table 2; a function in disorazol biosynthesis could not be determined from comparisons to similar proteins in the database. Several genes were found for which functional prediction based on their homology is not possible. Further work has to be done to examine the role of these proteins. This will hopefully be facilitated by the results of the genome sequencing project from the closely related strain *S. cellulosum* So ce56 since genes with significant similarities have been found (M.K. and R.M., unpublished results).

Analysis of two additional chromosomal regions involved either in disorazol biosynthesis or regulation

Sequence information from plasmids recovered from other chromosomal regions has shown that genes that encode β-lactamases/esterases were targeted in one of the transposon mutants (Table 1). Referring to the bis-lactone structure of disorazol, a potential involvement of this type of functionality in the final formation of disorazoles from the monomeric polyketide intermediate could be supposed. Penicillin-binding proteins, such as the β-lactamases, are known to become acylated by bacterial cell-wall biosynthetic intermediates, which are further transferred giving rise to peptidoglycan. During disorazol biosynthesis such a protein might bind one monomer transferred from the PCP and assist the TE in forming the heterodimer. HPLC-MS analysis of the secondary metabolites produced from the β-lactamase mutant could not identify masses corresponding to one of the monomeric polyketides that form the disorazol derivatives. Nevertheless, this is not surprising, as the monomers would have to be set free from the biosynthetic proteins and no turnover can be expected in the mutant strain. To analyze all the ORFs adjacent to the inactivated gene, the recovered plasmid was completely sequenced. This enabled the analysis of genes that might be inactivated in the transposon mutant due to a possible polar effect (see Table 3 and Figure 6). The genes found downstream of the inactivated esterase/β-lactamase gene cannot be affected by the insertion of the transposon because they are transcribed in the opposite direction. Upstream of the inactivation site, genes encoding a protein with similarities to MTs could be identified. At this stage it remains unclear whether this gene together with the esterase/β-lactamase is indeed involved in bis-lactone formation and O-methylation of the oxygen at C6 or C6′. Further investigations are necessary to elucidate these steps of disorazol biosynthesis. Additionally, it remains unclear how the epoxidation is achieved.

It could be possible to identify the direct product of the core biosynthetic gene cluster by heterologous expression of DisA–D, which are located on a single BAC. Expression of the additional proteins encoded in the second chromosomal region identified by transposon mutagenesis, should help to understand their influence in disorazol biosynthesis.

The function of the third chromosomal region linked to disorazol biosynthesis is still unclear. Because of the similarity of the 3′-region to carbamoyltransferase genes a direct involvement in disorazol biosynthesis is unlikely. A putative regulatory effect on the biosynthetic machinery remains to be proven.

Experimental Section

Strains and media: *S. cellulosum* So ce12 was grown in HS medium and on solid PM_{12} plates as described previously.[23] Transformants of *S. cellulosum* So ce12 that carry the transposable element pMiniHimarHyg were grown in HS medium and on PM_{12} plates that contained hygromycin B (final concentration 150 $\mu g\,mL^{-1}$).

E. coli DH5α(λpir) cells were used for transposon recovery experiments and were grown in Luria–Bertani (LB) medium that contained hygromycin B (100 $\mu g\,mL^{-1}$), at 37 °C.

The yeast *Rhodotorula glutinis* was grown in liquid Myc medium at 30 °C in a gyratory shaker at 165 rpm as described previously.[23]

Construction and screening of a BAC-library: High molecular weight DNA was prepared as described by Riethman et al. and dia-

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

lyzed in 2 mL tubes at RT against 1× Tris/EDTA (TE) buffer (10 mM Tris base, 1 mM EDTA pH 8.0).[66] Partial digest of agarose plugs was performed with *Bam*HI (0.5 U, 0.075 mM Mg^{2+}). To size-fractionate the partially-digested DNA, preparative pulsed field gel electrophoresis (PFGE) was performed according to published procedures.[67] The size fraction of 100–150 kb was used for subsequent cloning into pIndigoBAC-5 ("*Bam*HI cloning ready"; Epicentre). The partially digested DNA was purified by using the enzyme β-Agarase and was then ligated to pIndigoBac-5 in a 10:1 vector:insert DNA ratio with T4 ligase (20 U; NEB) at 16 °C, overnight.[67,68] The ligation reaction was dialyzed against 0.5×TE buffer and then electroporated into *E. coli* DH10B in a 0.1 cm cuvette at a capacity of 25 μF, resistance of 100 Ω, and current at 1.8 kV. After 45 min of phenotypical expression in SOC medium the cells were plated onto LB agar with chloramphenicol (12.5 $\mu g\,mL^{-1}$) and incubated at 37 °C, overnight.

Approximately 2000 colonies were spotted onto nylon membranes and screened with probes generated from the recovered plasmids pTn-Rec_2793 and pTn-Rec_13-21. To generate labeled probes, a PCR labeling kit (Roche) was used with the primers dis_2793_fwd: 5′-CGTCCAGGATCCACTCGT-3′, dis_2793_rev: 5′-CGTCGTCCACCACGACAT-3′, dis_13-21_fwd: 5′-CTTCTTCCAGGCGACCTCTC-3′, and dis_13-21_rev: 5′-CCGCCGGTGATCAGATAG-3′. Hybridizations were carried out at 65 °C, overnight. After stringent washing with 0.5-fold SSC (2×), the signals were detected by using CDP star (Roche) according to manufacturer's protocol.

DNA preparations, manipulations, analysis, and PCR: *S. cellulosum* So ce12 genomic DNA was prepared with the Purgene® Genomic DNA Purification Kit (Gentra) according to the manufacturer's protocol. Plasmid DNA purification was performed by using the NucleoSpin Plasmid Kit (Macherey-Nagel). BAC DNA was isolated as described by Birnboim and Doly, and modified according to Qiagen (Hilden, Germany).[69]

PCR was performed with *Taq* polymerase (MBI Fermentas). The reactions contained 5 % DMSO and were carried out in an Eppendorf Mastercycler gradient (Eppendorf) under the following conditions: initial denaturation for 5 min at 95 °C prior to the addition of polymerase; 30 cycles consisting of 30 s of denaturation at 95 °C, 20 s of annealing at 60 °C, and 30 s of extension at 72 °C; and a final extension of 10 min at 72 °C. Primers for the control reaction using the BACs as template, were dis_2793_fwd and dis_2793_rev, dis_13-3_fwd (5′-GCGCCGCATTTCGTAGAG-3′) and dis_13-3_rev (5′-CTCGTTCCGAGGTGCTTC-3′), dis_13-21_fwd and dis_13-21_rev, dis_IE-2_fwd (5′-CATGCCGCCGTTCGAGAAG-3′) and dis_IE-2_rev (5′-CAGGACGAGCAGGACTTCG-3′).

All other DNA manipulations were performed according to standard protocols.[70]

Sequencing of the BAC that encoded the disorazol biosynthetic gene cluster was performed in a shotgun approach by Qiagen (Hilden).

DNA and amino-acid sequence analyses were carried out by using the VectorNTI software package (Invitrogen). Comparison with Genbank data was performed with Blast programs.[71]

Construction of the myxobacterial transposon: A 2 kb *Hind*III fragment harboring the *hyg*^{R} gene, which is under the control of the *aph*II-promotor, from pJKB10 was subcloned into *Hind*III predigested plasmid pBCSK(−) (Stratagene).[22] The *hyg*^{R} cassette was then excised from the resulting plasmid, pFG113, as a 2.2 kb *Bss*HII fragment. The kanamycin resistance gene was removed from pMiniHimar by using *Mlu*I and replaced by the 2.2 kb *hyg*^{R} cassette, which resulted in pMiniHimarHyg.

Electroporation: *S. cellulosum* So ce12 was grown to a cell density of 1.0–1.5×10^8 cells$\,mL^{-1}$. Cells from a 300 mL culture were harvested by centrifugation with a GS3 rotor, at 4650 g and 4 °C. The transformation was essentially carried out by using the conditions described in the European patent application EP 04103546.0 (Irschik, Gerth, Kopp, Perlova, and Müller). The pellet was washed with ice cold washing buffer (300 mL). The cells were collected by centrifugation according to the previous step. Two further washing steps with washing buffer (30 mL each, see patent application mentioned above) were carried out. The cells were collect by using a SS34 rotor at 12 100 g at 4 °C. The final pellet was resuspended in ice-cold electroporation buffer to give a final cell density of approximately 5×10^{10} cells$\,mL^{-1}$. This cell suspension (100 μL) was mixed on ice with transposon pMiniHimarHyg plasmid DNA (approximately 1 μg) and transferred into a chilled 1 mm electroporation chamber. Electroporations were then carried out and time constants of 6–8 ms were obtained. The cells were transferred into HS medium (10 mL) and grown at 30 °C and 165 rpm for phenotypical expression. This culture was added to HS medium (20 mL) that contained hygromycin B (final concentration 150 $\mu g\,mL^{-1}$). The cells were incubated at 30 °C and 165 rpm and collected after 8 h. The pellet was resuspended in 1 mL HS medium and equally spread on six PM_{12} plates with hygromycin B (150 $\mu g\,ml^{-1}$). The plates were incubated for 10–12 days at 32 °C.

Detection of disorazol-negative mutants: *S. cellulosum* So ce12 transposon mutants were transferred to PM_{12} plates without hygromycin and incubated at 32 °C until colonies became visible. An *R. glutinis* overnight culture (200 μL) was added to Myc soft agar (100 mL) and 5 mL of this suspension was poured over the *S. cellulosum* So ce12 mutant colonies. The plates were incubated overnight at 30 °C and growth-inhibition zones were compared to the wild type.

Transposon recovery: Transposon recovery from *S. cellulosum* So ce12 mutants and sequencing of the resulting plasmids was performed as described previously.[23]

Southern hybridization: The 2.2 kb *hyg*^{R} cassette from pFG113 was labeled with digoxigenin by using DIG High Prime kit (Roche). Genomic DNA was digested with *Mlu*I and blotted onto positively charged nylon membranes (Roche). After 2 h of prehybridization in standard hybridization buffer at 65 °C, the denatured probe was added to hybridization buffer (25 mL) according to the manufacturer's recommendations and incubated with the membrane. Hybridization was carried out at 65 °C, overnight. After stringent washing with 0.5-fold SSC (2×), the signals were detected with NBT solution according to the manufacturer's protocol.

Analysis of secondary metabolites: Analysis of the secondary metabolites produced by *S. cellulosum* So ce12 was carried out as described previously.[23]

Nucleotide accession number: The nucleotide sequence reported here is available under accession number AJ874112 (EMBL database).

Acknowledgements

The authors would like to acknowledge helpful comments of B. Lelarge, H. B. Bode and M. Zabriskie regarding this manuscript. pMiniHimar was a kind gift from H. Kaplan. We would like to thank F. Gross for providing pMiniHimarHyg and I. Kristen and C.

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Körner for skilful technical assistance. Y. Elnakady kindly provided immunofluorescense pictures of L929 cells.

Keywords: biosynthesis · natural products · polyketides · transferases · tubulin

[1] K. Gerth, S. Pradella, O. Perlova, S. Beyer, R. Müller, *J. Biotechnol.* **2003**, *106*, 233.
[2] H. Reichenbach, G. Höfle in *Drug Discovery from Nature* (Eds.: S. Grabley, R. Thiericke), Springer, Berlin, **1999**, pp. 149.
[3] H. Irschik, R. Jansen, K. Gerth, G. Höfle, H. Reichenbach, *J. Antibiot.* **1995**, *48*, 31.
[4] Y. Elnakady, F. Sasse, H. Lünsdorf, H. Reichenbach, *Biochem. Pharmacol.* **2004**, *67*, 927.
[5] R. Jansen, H. Irschik, H. Reichenbach, V. Wray, G. Höfle, *Liebigs Ann. Chem.* **1994**, 759.
[6] H. Irschik, R. Jansen, K. Gerth, G. Höfle, H. Reichenbach, *J. Antibiot.* **1987**, *40*, 7.
[7] H. Irschik, R. Jansen, K. Gerth, G. Höfle, H. Reichenbach, *J. Antibiot.* **1995**, *48*, 886.
[8] H. Irschik, R. Jansen, K. Gerth, G. Höfle, H. Reichenbach, *J. Antibiot.* **1995**, *48*, 962.
[9] B. Silakowski, B. Kunze, R. Müller, *Gene* **2001**, *275*, 233.
[10] W. C. Nierman, D. Kaiser, H. B. Kaplan, B. Goldman, et al., **2005**, unpublished results.
[11] B. Kunze, R. Jansen, F. Sasse, G. Höfle, H. Reichenbach, *J. Antibiot.* **1995**, *48*, 1262.
[12] F. Sasse, B. Kunze, T. M. Gronewold, H. Reichenbach, *J. Natl. Med. Assoc. J. Nat. Cancer Inst.* **1998**, *90*, 1559.
[13] F. Sasse, H. Steinmetz, G. Höfle, H. Reichenbach, *J. Antibiot.* **1993**, *46*, 741.
[14] F. Sasse, H. Steinmetz, J. Heil, G. Höfle, H. Reichenbach, *J. Antibiot.* **2000**, *53*, 879.
[15] H. Steinmetz, N. Glaser, E. Herdtweck, F. Sasse, H. Reichenbach, G. Höfle, *Angew. Chem.* **2004**, *116*, 4996; *Angew. Chem. Int. Ed.* **2004**, *43*, 4888.
[16] K. Gerth, N. Bedorf, G. Höfle, H. Irschik, H. Reichenbach, *J. Antibiot.* **1996**, *49*, 560.
[17] D. M. Bollag, P. A. McQueney, J. Zhu, O. Hensens, L. Koupal, J. Liesch, M. Goetz, E. Lazarides, M. Woods, *Cancer Res.* **1995**, *55*, 2325.
[18] T. Schupp, C. Toupet, B. Cluzel, S. Neff, S. Hill, J. J. Beck, J. M. Ligon, *J. Bacteriol.* **1995**, *177*, 3673.
[19] I. Molnar, T. Schupp, M. Ono, R. Zirkle, M. Milnamow, B. Nowak-Thompson, N. Engel, C. Toupet, A. Stratmann, D. D. Cyr, J. Gorlach, J. M. Mayo, A. Hu, S. Goff, J. Schmid, J. M. Ligon, *Chem. Biol.* **2000**, *7*, 97.
[20] B. Julien, S. Shah, R. Ziermann, R. Goldman, L. Katz, C. Khosla, *Gene* **2000**, *249*, 153.
[21] J. Ligon, S. Hill, J. Beck, R. Zirkle, I. Molnar, J. Zawodny, S. Money, T. Schupp, *Gene* **2002**, *285*, 257.
[22] S. Pradella, A. Hans, C. Sproer, H. Reichenbach, K. Gerth, S. Beyer, *Arch. Microbiol.* **2002**, *178*, 484.
[23] M. Kopp, H. Irschik, F. Gross, O. Perlova, A. Sandmann, K. Gerth, R. Müller, *J. Biotechnol.* **2004**, *107*, 29.
[24] H. H. Kimsey, D. Kaiser, *J. Bacteriol.* **1991**, *173*, 6790.
[25] I. Stamm, A. Leclerque, W. Plaga, *Arch. Microbiol.* **1999**, *172*, 175.
[26] S. Beyer, B. Kunze, B. Silakowski, R. Müller, *Biochim. Biophys. Acta* **1999**, *1445*, 185.
[27] A. Sandmann, F. Sasse, R. Müller, *Chem. Biol.* 2004, 11, 1071.
[28] E. Rubin, B. Akerley, V. Novik, D. Lampe, R. Husson, J. Mekalanos, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 1645.
[29] S. Donadio, M. J. Staver, J. B. McAlpine, S. J. Swanson, L. Katz, *Science* **1991**, *252*, 675.
[30] T. Schwecke, J. F. Aparicio, I. Molnar, A. König, L. E. Khaw, S. F. Haydock, M. Oliynyk, P. Caffrey, J. Cortes, J. B. Lester, G. A. Böhm, J. Staunton, P. F. Leadlay, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 7839.
[31] D. Konz, M. A. Marahiel, *Chem. Biol.* **1999**, *6*, R39.
[32] Y. Q. Cheng, G. L. Tang, B. Shen, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3149.
[33] A. K. El-Sayed, J. Hothersall, S. M. Cooper, E. Stephens, T. J. Simpson, C. M. Thomas, *Chem. Biol.* **2003**, *10*, 419.
[34] S. Haydock, J. F. Aparicio, I. Molnar, T. Schwecke, A. König, A. F. A. Marsden, I. S. Galloway, J. Staunton, P. F. Leadlay, *FEBS Lett.* **1995**, *374*, 246.
[35] G. Yadav, R. S. Gokhale, D. Mohanty, *J. Mol. Biol.* **2003**, *328*, 335.
[36] L. Tang, Y. J. Yoon, C. Y. Choi, C. R. Hutchinson, *Gene* **1998**, *216*, 255.
[37] R. M. Kagan, S. Clarke, *Arch. Biochem. Biophys.* **1994**, *310*, 417.
[38] C. P. Joshi, V. L. Chiang, *Plant Mol. Biol.* **1998**, *37*, 663.
[39] T. Stachelhaus, H. D. Mootz, M. A. Marahiel, *Chem. Biol.* **1999**, *6*, 493.
[40] G. L. Challis, J. Ravel, C. A. Townsend, *Chem. Biol.* **2000**, *7*, 211.
[41] L. Du, M. Chen, C. Sanchez, B. Shen, *FEMS Microbiol. Lett.* **2000**, *189*, 171.
[42] S. Jaoua, S. Neff, T. Schupp, *Plasmid* **1992**, *28*, 157.
[43] J. Piel, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14002.
[44] J. Piel, G. P. Wen, M. Platzer, D. Q. Hui, *ChemBioChem* **2004**, *5*, 93.
[45] Y. Q. Cheng, G. L. Tang, B. Shen, *J. Bacteriol.* **2002**, *184*, 7013.
[46] G. L. Tang, Y. Q. Cheng, B. Shen, *Chem. Biol.* **2004**, *11*, 33.
[47] S. Mochizuki, K. Hiratsu, M. Suwa, T. Ishii, F. Sugino, K. Yamada, H. Kinashi, *Mol. Microbiol.* **2003**, *48*, 1501.
[48] C. Scotti, M. Piatti, A. Cuzzoni, P. Perani, A. Tognoni, G. Grandi, A. Galizzi, A. M. Albertini, *Gene* **1993**, *130*, 65.
[49] J. Piel, D. Hui, G. Wen, D. Butzke, M. Platzer, N. Fusetani, S. Matsunaga, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16222.
[50] G. Huang, L. Zhang, R. G. Birch, *Microbiology* **2001**, *147*, 631.
[51] Y. Paitan, G. Alon, E. Orr, E. Z. Ron, E. Rosenberg, *J. Mol. Biol.* **1999**, *286*, 465.
[52] A. M. Albertini, T. Caramori, F. Scoffone, C. Scotti, A. Galizzi, *Microbiology* **1995**, *141*, 299.
[53] R. Müller, *Chem. Biol.* **2004**, *11*, 4.
[54] B. Silakowski, H. U. Schairer, H. Ehret, B. Kunze, S. Weinig, G. Nordsiek, P. Brandt, H. Blöcker, G. Höfle, S. Beyer, R. Müller, *J. Biol. Chem.* **1999**, *274*, 37391.
[55] S. Weinig, T. Mahmud, R. Müller, *Chem. Biol.* 2003, 10, 953.
[56] B. Silakowski, G. Nordsiek, B. Kunze, H. Blöcker, R. Müller, *Chem. Biol.* **2001**, *8*, 59.
[57] D. Konz, A. Klens, K. Schorgendorfer, M. A. Marahiel, *Chem. Biol.* **1997**, *4*, 927.
[58] H. W. Chen, S. O'Connor, D. E. Cane, C. T. Walsh, *Chem. Biol.* **2001**, *8*, 899.
[59] T. A. Keating, C. G. Marshall, C. T. Walsh, *Biochemistry* **2000**, *39*, 15513.
[60] C. G. Marshall, M. D. Burkart, T. A. Keating, C. T. Walsh, *Biochemistry* **2001**, *40*, 10655.
[61] A. M. Gehring, I. Mori, C. T. Walsh, *Biochemistry* **1998**, *37*, 2648.
[62] J. J. May, T. M. Wendrich, M. A. Marahiel, *J. Biol. Chem.* **2001**, *276*, 7209.
[63] P. R. August, L. Tang, Y. J. Yoon, S. Ning, R. Müller, T. W. Yu, M. Taylor, D. Hoffmann, C. G. Kim, X. Zhang, C. R. Hutchinson, H. G. Floss, *Chem. Biol.* **1998**, *5*, 69.
[64] L. Tang, S. Ward, L. Chung, J. R. Carney, Y. Li, R. Reid, L. Katz, *J. Am. Chem. Soc.* **2004**, *126*, 46.
[65] R. Reid, M. Piagentini, E. Rodriguez, G. Ashley, N. Viswanathan, J. Carney, D. V. Santi, C. R. Hutchinson, R. McDaniel, *Biochemistry* **2003**, *42*, 72.
[66] H. Riethman, B. Birren, A. Gnirke in *Genome Analysis. A Laboratory Manual*, *Vol. 1: Analyzing DNA* (Eds.: B. Birren, E. D. Green, S. Klapholz, R. M. Myers, J. Roskams,), Cold Spring Harbor, Laboratory Press, New York, **1997**, pp. 152.
[67] B. Birren, V. Mancino, H. Shizuya in *Genome Analysis. A Laboratory Manual*, *Vol. 3: Cloning Systems* (Ed.: B. Birren, E. D. Green, S. Klapholz, R. M. Myers, J. Roskams), Cold Spring Harbor Laboratory Press, New York, **1997**, pp. 241.
[68] K. Osoegawa, P. Y. Woon, B. Zhao, E. Frengen, M. Tateno, J. J. Catanese, P. J. de Jong, *Genomics* **1998**, *52*, 1.
[69] H. C. Birnboim, J. Doly, *Nucleic Acids Res.* **1979**, *7*, 1513.
[70] J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, NY, **1989**.
[71] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* **1990**, *215*, 403.

Received: December 30, 2004
Published online on May 13, 2005

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim