Two new tailoring enzymes, a glycosyltransferase and an oxygenase, involved in biosynthesis of the angucycline antibiotic urdamycin A in *Streptomyces fradiae* Tü2717

B. Faust,1 D. Hoffmeister,1 G. Weitnauer,1 L. Westrich,1 S. Haag,2 P. Schneider,1 H. Decker,2 E. Küenzel,3 J. Rohr3 and A. Bechthold1

Author for correspondence: A. Bechthold. Tel.: +49 7071 2975483. Fax: +49 7071 295250. e-mail: andreas.bechthold@uni-tuebingen.de

1 Universität Tübingen, Pharmazeutisches Institut, Auf der Morgenstelle 8, D-72076 Tübingen, Germany
2 Hoechst Marion Roussel Deutschland GmbH, Process Development, D-65926 Frankfurt, Germany
3 Medical University of South Carolina, Department of Pharmaceutical Sciences, 171 Ashley Avenue, Charleston, SC 29425-2303, USA

Urdamycin A, the principal product of *Streptomyces fradiae* Tü2717, is an angucycline-type antibiotic and anticancer agent containing C-glycosidically linked D-olivose. To extend knowledge of the biosynthesis of urdamycin A the authors have cloned further parts of the urdamycin biosynthetic gene cluster. Three new ORFs (urdK, urdJ and urdO) were identified on a 3-35 kb fragment, and seven new ORFs (urdL, urdM, urdJ2, urdZ1, urdGT2, urdG and urdH) on an 8-05 kb fragment. The deduced products of these genes show similarities to transporters (urdJ and urdJ2), regulatory genes (urdK), reductases (urdO), cyclases (urdL) and deoxysugar biosynthetic genes (urdG, urdH and urdZ1). The product of urdM shows striking sequence similarity to oxygenases (N-terminal sequence) as well as reductases (C-terminal sequence), and the deduced amino acid sequence of urdGT2 resembles those of glycosyltransferases. To determine the function of urdM and urdGT2, targeted gene inactivation experiments were performed. The resulting urdM deletion mutant strains accumulated predominantly rabelomycin, indicating that UrdM is involved in oxygenation at position 12b of urdamycin A. A mutant in which urdGT2 had been deleted produced urdamycin I, urdamycin J and urdamycin K instead of urdamycin A. Urdamycins I, J and K are tetracyclic angucyclinones lacking a C-C connected deoxysugar moiety. Therefore UrdGT2 must catalyse the earliest glycosyltransfer step in the urdamycin biosynthetic pathway, the C-glycosyltransfer of one NDP-D-olivose.

Keywords: glycosyltransferase, oxygenase, angucycline, urdamycin A, *Streptomyces*

INTRODUCTION

Urdamycin A (Fig. 1), an angucycline-type antibiotic and anticancer agent (Drautz et al., 1995), is the principal product of *Streptomyces fradiae* Tü2717. It consists of aquayamycin (Fig. 1) and three O-glycosidically linked deoxysugars: two α-rhodinoses and D-olivose. Aquayamycin contains a C-glycosidically linked D-olivose and is the most common aglycone among the angucycline group of antibiotics. Thus it is an intermediate in several biosynthetic pathways, including that for urdamycin (Rohr & Thiericke, 1992; Krohn & Rohr, 1997). Aquayamycin is derived from a decapolyketide that, after cyclization and reduction, becomes modified by oxygenation and glycosylation reactions. In early genetic studies of the angucyclines jadomycin and landomycin, all of the biosynthetic genes first identified were shown to be typical genes of type II polyketide synthases (PKSs) (Han et al., 1994; Westrich et al., 1999). Genes that may be involved in tailoring reactions have since been found in both clusters. The function of Jad-ORF6 as an oxygenase involved in ring opening of
rabelomycin, an intermediate in jadomycin biosynthesis, has been shown by gene disruption experiments (Yang et al., 1996). Genes encoding the PKS involved in urdamycin formation have also been cloned and sequenced (Decker & Haag, 1995). Six genes have been characterized, five of them directly involved in forming the polyketide moiety and one involved in its oxygenation (Decker & Haag, 1995). Our goal in the present work was to clarify which gene products of the urdamycin pathway would catalyse tailoring reactions during the biosynthesis of aquayamycin.

**METHODS**

**Bacterial strains, plasmids and culture conditions.** *Streptomyces fradiae* Tu2717 (Drautz et al., 1986), obtained from the Department of Microbiology, University of Tübingen, was grown on 1% malt extract, 0.4% yeast extract, 0.4% glucose and 1 mM CaCl₂, pH 7.2 (HA medium), at 28 °C. DNA manipulation was performed in *Escherichia coli* XL-1 Blue MRF (Stratagene). Plasmids were passed through *E. coli* ET12567 (dam dcm hsdS Cm⁺) to generate unmethylated DNA before their use to transform *S. fradiae* Tu2717. *E. coli* strains were grown under standard conditions. pBluescript SK− (pSK−) was obtained from Stratagene and plasmid pSP1, carrying the erythromycin-resistance gene (Pelzer et al., 1997), was provided by S. Pelzer, Department of Microbiology, University of Tübingen. pKC1132, carrying the apramycin-resistance gene, was a gift from Eli Lilly.

**General genetic manipulation and SDS-PAGE.** Isolation of *E. coli* plasmid DNA, digestion of DNA with restriction endonucleases and Southern hybridization were carried out according to the directions of the supplier of kits, enzymes and reagents (Amersham). Southern hybridization was performed with Hybond N nylon membranes (Amersham). Probes were labelled with digoxigenin by using a DIG (digoxigenin) labelling and detection kit (Boehringer Mannheim). Restriction mapping, other routine molecular biology methods and SDS-PAGE (Coomassie blue staining) were performed as described by Sambrook et al. (1989). Protoplast formation, transformation, and regeneration of protoplasts from *S. fradiae* Tu2717 were carried out by standard procedures (Hopwood et al., 1985).

**Library construction and screening.** A complete genomic library of *S. fradiae* Tu2717 was prepared in cosmid pOJ446 as described by Westrich et al. (1999). The cosmid library was screened for hybridization using an internal fragment of a dNDP-glucose 4,6-dehydratase gene as a probe. This fragment had been obtained by PCR amplification (Decker et al., 1996).

**DNA sequencing and computer-assisted sequence analysis.** DNA was sequenced by the dideoxynucleotide chain-termination method with thermosequenase. Universal and reverse primers were used. Sequencing reactions were performed on an automated sequencer (Vistra 725) from Molecular Dynamics and on an ABI sequencer from 4-base lab. DNA sequences were analysed using the DNASIS software package (version 2, 1995; Hitachi Software Engineering). BLASTX (Altschul et al., 1997) was used to search the GenBank CDC translations + PDB + SWISS-PROT + Spupdate + PIR, release 2.0 for matching sequences.

**Generation of a chromosomal urdGT2 mutant of S. fradiae Tu2717.** For generation of a chromosomal urdGT2 mutant of
**RESULTS**

**Screening a cosmid library and restriction mapping of cosmid purd12**

Three colonies contained DNA hybridizing to the TDP-glucose 4,6-dehydratase gene probe. Cosmid DNA isolated from these colonies was analysed by Southern hybridization and restriction mapping. All three cosmids contained overlapping DNA. Cosmid purd12 (Fig. 2) was used for further restriction mapping, hybridization and sequencing experiments.

**Sequence analysis**

We sequenced 4 kb of DNA located upstream, and 8 kb of DNA located downstream of the PKS genes. Within the 4 kb of DNA, three ORFs (urdK, urdJ and urdO) could be identified; seven new ORFs (urdL, urdM, urdJ2, urdZ1, urdGT2, urdG and urdH) were detected within the 8 kb fragment (Fig. 2). All the ORFs possessed the characteristics of Streptomyces genes (overall G + C content 68–90 mol%; high bias towards G and C at the third codon position). Database comparisons with the deduced products of urdJ and urdJ2 using BLASTP and BLASTX revealed similarities between both proteins and various transporters. UrdJ is most similar to LanJ of *Streptomyces cyanogenus* S136 (Westrich et al., 1999) (55% identical amino acids), and UrdJ2 resembles a putative export protein from *S. roseofulvus* (database accession number AF058302) (37% identical amino acids). The hydropathy profiles of both proteins indicate several fragments of sufficient length and hydrophobicity to be membrane-spanning sequences (Kyte & Doolittle, 1982). UrdK is 53% identical to LanK from *S. cyanogenus* S136 and both resemble the putative regulatory proteins JadR2 from *S. venezuelae* (Yang et al., 1995) and TcmR from *S. glaucescens* (Guilfoile & Hutchinson, 1992). Three of the predicted protein sequences (UrdG, UrdH, UrdZ1) show homology to proteins known to be involved in deoxysugar biosynthesis. The highest resemblance was found to LanG (66% identical amino acids), LanH (75% identical amino acids).
Fig. 2. Restriction map of a 50 kb fragment of *S. fradiae* Tu2717 genomic DNA. Cosmids purd8 and purd12 contain about 25 kb of *S. fradiae* Tu2717 DNA. Black arrows represent the extent and direction of ORFs in the urdamycin biosynthetic gene cluster. Restriction sites: Ps, *Pst*I; Ba, *Bam*HI; Ec, *Eco*RI; S, *Sac*I.

B. FAUST and OTHERS

amino acids) and LanZ1 (59% identical amino acids), respectively (Westrich et al., 1999). *urdG* might code for a dNDP-β-glucose synthase, *urdH* for a dNDP-glucose 4,6-dehydratase and *urdZ1* for a dNDP-hexose 3,5-epimerase. The deduced UrdL protein resembles proteins assigned as aromatases in different organisms. The closest resemblance was to LanL (72% identical amino acids) (Westrich et al., 1999) and JadOrf4 (76% identical amino acids) (Han et al., 1994). Comparison of the putative product of *urdO* to proteins in the database suggests that UrdO belongs to the short-chain alcohol dehydrogenase family of proteins. The highest resemblance was to LanO (61% identical amino acids) and LanZ4 (53% identical amino acids). All these proteins contain a motif common in some NAD(P)H-dependent dehydrogenases. Comparison of the N-terminal sequence (amino acids 1–383) of UrdM with database sequences showed significant similarity to oxygenases. UrdM strongly resembled the products of *lanM* and *lanE* from *S. cyanogenus* S136 (Westrich et al., 1999) (57% identical amino acids in comparison with amino acids 1–377 of LanM, and 48% identical amino acids in comparison with amino acids 1–373 of LanE). Further, strong resemblance was found to UrdE from *S. fradiae* Tu2717 (Decker & Haag, 1995) (47% identical amino acids in comparison with amino acids 1–376 of UrdE), and to an oxygenase gene (*orf6*) from *S. venezuelae* (Yang et al., 1996) (57% identical amino acids in comparison with amino acids 1–376 of ORF6). The C-terminal sequence of UrdM (amino acids 415–672) resembled putative reductases. Most similar were LanV (54% identical amino acids), LanN (50% identical amino acids) (Westrich et al., 1999) and a putative glucose dehydrogenase from *S. venezuelae* (database accession number U24659) (48% identical amino acids). The deduced amino acid sequence encoded by *urdGT2* exhibits similarity to glycosyltransferases. It most resembles a glycosyltransferase (ORF14) from *S. violaceoruber* Tu22 (Ichinose et al., 1998) and a glycosyltransferase (LanGT2) from *S. cyanogenus* S136 (Westrich et al., 1999).

Identification of a chromosomal *urdGT2* mutant of *S. fradiae* Tu2717

To generate a chromosomal *urdGT2* mutant of *S. fradiae* Tu2717 by homologous recombination the plasmid pSP-urdGT2d was constructed (Fig. 3b). This was used to introduce an in-frame deletion into *urdGT2*. After transformation of *S. fradiae* Tu2717 by pSP-urdGT2d, several erythromycin-resistant colonies were obtained. Sall-digested chromosomal DNA of such mutants was probed with a 1±7 k b *Sma*I fragment from pSP1 containing parts of the erythromycin-resistance gene, and with the 1±9 k b *Sal*I fragment containing *urdGT2*. In mutant BF-1 the 1±7 kb probe detected a hybridizing fragment. When the 1±9 kb Sall fragment was used as a probe, a 1±55 kb band was detected. Integration of pSP-urdGT2d by a single cross-over event can take place in two different ways (upstream or downstream of the deleted fragment). Both integrations should give hybridization signals at 1±9 kb and 1±55 kb after Sall digestion when probed with the 1±9 kb Sall fragment. The absence of the 1±9 kb band in mutant
Biosynthesis of urdamycin A

Fig. 3. Generation of a chromosomal urdGT2 mutant of S. fradiae Tu2717 by an in-frame deletion. (a) Wild-type DNA. (b) Inactivation construct pSP-urdGT2d, containing a 327 bp deletion within urdGT2. (c) DNA of mutant BF-1 obtained by integration of pSP-urdGT2d into S. fradiae Tu2717 by an unusual integration event. (d) Mutant BF-1-1, obtained by a double crossover, containing the expected deletion within the chromosome.

Fig. 4. Southern hybridization of total DNA of the wild-type S. fradiae Tu2717 (1) and mutant BF-1-1 (2). DNA was restricted with SalI and probed with a 1.9 kb SalI fragment. Lanes M, molecular markers.

BF-1 is not in accordance with one of these usual single cross-over events but it can be explained by the unusual integration event depicted in Fig. 3(c). Mutant BF-1-1 was obtained from mutant BF-1 after screening for loss of resistance to erythromycin. When chromosomal DNA of BF-1-1 was digested with SalI and probed with the 1.7 kb SmaI fragment no signal was detected. Using the 1.9 kb fragment as a probe a 1.55 kb band was detected, instead of the 1.9 kb signal detected in the wild-type strain (Fig. 4). Using as a probe the 207 bp StyI fragment that had been deleted in pSP-urdGT2d no signal was detected whereas a 1.9 kb fragment was detected in the wild-type strain. These results confirmed that the expected deletion had occurred in the chromosome of BF-1-1 (Fig. 3d). To investigate the structure of intermediates produced by mutant BF-1-1, compounds accumulated by cultures were isolated and their structures were elucidated by NMR spectroscopy (data not shown). Mutant BF-1-1 accumulated one major product, urdamycin I, and minor amounts of urdamycin J and urdamycin K (Fig. 1) (Künzel et al., 1999).

Complementation of BF-1-1 with urdGT2

urdGT2 on a 1.9 kb SalI fragment was ligated into the expression vector pEM4 (Quiros et al., 1998) to create pBF-EM4+1. After transformation of BF-1-1 with this plasmid several colonies were obtained and grown at 28 °C in liquid medium containing thiostrepton. Extracts of the cultures were analysed by HPLC. Seven of 10 colonies produced urdamycin A, indicating that the function of genes downstream of urdGT2 was not affected by the deletion.
Identification of a chromosomal urdM mutant of *S. fradiae* Tü2717

When *S. fradiae* Tü2717 was transformed by pKC-urdM, several apramycin-resistant colonies were obtained. Integration of pKC-urdM into the chromosome at the expected position was verified by Southern hybridization. To allow for the second recombination event, integrants were screened for apramycin sensitivity. The chromosomal mutation in mutant BF-2-1, which showed apramycin sensitivity, was analysed by Southern hybridization. A 3-4 kb SacI fragment, carrying the 3'-terminal portion of *urdM*, was used to probe SacI-digested chromosomal DNA. Analysis of the wild-type *S. fradiae* showed the expected 3-4 kb fragment after hybridization. When chromosomal DNA from clone BF-2-1 was treated similarly, one 2.9 kb fragment was detected, verifying the deletion of *urdM*. To confirm the presence of the deletion within the chromosome of BF-2-1 the mutant strain and the wild-type strain were individually subjected to PCR analysis. The size of the amplified fragment (0.8 kb) detected in the mutant strain was identical to the one generated with the plasmid pKC-urdMd, while the wild-type strain gave a PCR fragment of the expected higher mobility (1.3 kb). Analysis for antibiotic production showed that strain BF-2-1 produced predominantly rabelomycin instead of urdamycin A (Fig. 1).

Complementation of mutant BF-2-1 with urdM and expression of urdM in *E. coli*

urdM, on a 7.1 kb fragment, was ligated into pUWL201 (U. Wehmeier & W. Piepersberg, personal communication) under the control of the ermE-up promoter to create pUWL-urdM. When pUWL-urdM was expressed in mutant BF-2-1, production of urdamycin A should be restored, provided that no other gene of the urdamycin cluster was influenced by the deletion introduced into urdM.

UrdM was also expressed in *E. coli* by using an inducible T7-RNA-polymerase-dependent expression system. Before induction with IPTG, T7 lysozyme, provided by *E. coli* BL21(DE3)/pLysS/pRSETb, inhibits the T7 RNA polymerase and decreases production of UrdM. Induction with IPTG increases the amount of T7 RNA polymerase and thereby the amount of the expressed protein. SDS-PAGE was used to monitor the expression of UrdM before and after induction with IPTG (Fig. 5). Upon induction, the level of a 70.5 kDa protein in extracts of *E. coli* BL21(DE3)/pLysS/pRSETb-urdM increased substantially. This prominent band was not detectable in extracts of *E. coli* BL21(DE3)/pLysS/pRSETb, either before or after induction.

**DISCUSSION**

So-called post-PKS modifying (tailoring) enzymes play a crucial role in the formation of interesting and unique molecular structures. Oxygenases and glycosyltransferases are especially important in providing structural elements essential for their specific biological activity (Weymouth-Wilson, 1997; Kirschning et al., 1997; Rohr, 1998, Trefzer et al., 1999). Thus, these enzymes and their corresponding genes may be very useful for the design of novel hybrid molecules. To continue our genetic work on cloning the genes of urdamycin biosynthesis we isolated and sequenced DNA located upstream and downstream of genes involved in biosynthesis of the urdamycin polyketide moiety. Predicted gene products from ten ORFs identified were similar to proteins known to be involved in antibiotic biosynthesis, antibiotic transport and regulation. Two transporter genes (*urdJ* and *urdJ2*) have been sequenced; both resemble hydrophobic membrane proteins. The deduced amino acid sequence of *urdK* is very similar to LanK, detected in the landomycin biosynthetic gene cluster. As in the landomycin cluster, a putative regulatory gene has been found in the urdamycin cluster, and *lanK* and *urdK* are transcribed divergently from *lanJ* and *urdJ*, respectively, indicating that both transporter proteins are regulated by LanK and UrdK at the level of transcription. Among all sequences showing similarity to UrdL, the resemblance was particularly strong to LanL, JadORF4,
AcrVII, and other cyclases (aromatases) believed to catalyse aromatization of the first ring of landomycin, jadomycin and actinorhodin, respectively. The function of UrdO remains unknown; it may act as a reductase during the biosynthesis of the polyketide moiety of urdamycin A. The biosynthesis of d-olivose and l-rhodinose should proceed through a NADP-4-keto-6-deoxyglucose intermediate. UrdG and UrdH are probably involved in the conversion of glucose-1-phosphate to this key intermediate. So far, one additional putative deoxyxugar biosynthetic gene, urdZ1, has been identified which might act as a 3,5-epimerase involved in l-rhodinose biosynthesis. Assignment of urdM as a second oxygenase gene (besides urdE) is consistent with the origin of two oxygenas at C-12 and C-12b in urdamycin A from molecular oxygen (Udvarnoki et al., 1992). In contrast to usual oxygenases found in antibiotic biosynthetic gene clusters, the putative UrdM protein consists of two parts: an oxygenase portion and a reductase portion. Within the sequence of the N-terminal portion a putative ADP-binding domain and a putative ribityl-binding domain are present which can be found in FAD- and NADPH-dependent enzymes. Like the C-terminal portion of TcmN, a multifunctional protein with cyclase/dehydratase and methyltransferase activity (Summers et al., 1992; Shen et al., 1995), the C-terminal portion of UrdM (reduction portion) seems to contain an internal start site. To be sure that we did not fail to detect a stop codon that would separate both portions, we expressed urdM in E. coli. As monitored by SDS-PAGE analysis a 70.5 kDa protein was produced by E. coli BL21(DE3)/pLysS/pRSETb-urdM and the level of expressed protein increased substantially upon induction with IPTG (Fig. 5). No similar band was detectable in extracts of E. coli BL21(DE3)/pLysS/pRSETb, indicating that the overexpression of the 70.5 kDa protein was dependent on urdM. Based on the size of the expressed protein, urdM undoubtedly encodes a 70.5 kDa protein consisting of two portions that might have different catalytic activities. After disruption of urdM the mutant BF-2-1 accumulated rabelomycin, which does not contain the tertiary alcohol groups in positions 4a and 12b. This indicates clearly that UrdM catalyses the oxygenation at position 12b during urdamycin biosynthesis, and also that urdE, the second oxygenase gene in the urdamycin cluster, is involved in oxygenation at position 12 and not at position 12b as has been discussed by Decker & Haag (1995). Rabelomycin has also been obtained after disruption of Orf6 from S. venezuelae involved in jadomycin biosynthesis. Thus, early steps in the biosynthetic pathways to urdamycin A and jadomycin are identical, and rabelomycin seems to be a central intermediate in both pathways (Yang et al., 1996; Rohr et al., 1993; Künzle et al., 1999).

Once a cyclized polyketide-derived aglycone moiety is available, two molecules of d-olivose and two of l-rhodinose must be added to it at specific sites. As aquayamycin is a central intermediate in urdamycin A biosynthesis, attachment of a d-olivose moiety at position C-9 should precede the addition of l-rhodinose at position C-12b (Rohr et al., 1993). The conclusion that urdGT2 encodes a glycosyltransferase was first based upon sequence similarities between its product and authentic bacterial glycosyltransferases including GraOrf14 and LanGT2. A targeted in-frame deletion of urdGT2 has now been accomplished and confirmed by Southern analysis. By introducing an in-frame deletion into urdGT2, negative effects on genes located downstream of urdGT2 could be avoided. An urdGT2 mutant accumulated urdamycins I, J and K, which are tetracyclial angucyclinones to which no C-glycosidic moiety is attached, thus implying that UrdGT2 catalyses the earliest glycosyltransfer step in the urdamycin biosynthetic pathway. This step is the C-glycosyltransfer of an activated d-olivose, since this precedes all other glycosylation steps. The structure of urdamycin I excludes its being the real substrate for the glycosyltransferase UrdGT2, and indicates that it is a shunt product derived from a hypothetical intermediate (Künzle et al., 1999). This is further converted into urdamycin I through the influence of oxidoreductases. GraOrf14 was assigned to the transfer of l-rhodinose to granaticin and not to the transfer of d-olivose to the aglycone during granaticin B biosynthesis. GraOrf14 and UrdGT2 are indeed very similar, indicating similar functions. Therefore we might speculate that GraOrf14 is responsible for attaching d-olivose to the aglycone or may be involved in both glycosyltransfer steps.

In conclusion, the work described here has allowed the unambiguous assignment of the functions of the gene products of urdGT2 and urdM.

ACKNOWLEDGEMENTS

This work was supported by EU grant BIO4-CT96-0068 to J.R. and A.B., by start-up grants of the Medical University of South Carolina (MUSC) to J.R. and by the Deutsche Forschungsgemeinschaft (SFB 323) to A.B.

REFERENCES


Received 26 July 1999; revised 14 September 1999; accepted 22 September 1999.