DOI: 10.1002/cbic.201200402

Genome Mining of the Biosynthetic Gene Cluster of the Polyene Macrolide Antibiotic Tetramycin and Characterization of a P450 Monooxygenase Involved in the Hydroxylation of the Tetramycin B Polyol Segment

Bo Cao,[a] Fen Yao,[a] Xiaoqing Zheng,[a] Dongbing Cui,[a] Yucheng Shao,[a] Changxiong Zhu,[b] Zixin Deng,[a] and Delin You*[a]

A polyene macrolide antibiotic tetramycin biosynthetic gene cluster was identified by genome mining and isolated from *Streptomyces hygrospinosus* var. *beijingensis*. Genetic and in silico analyses gave insights into the mechanism of biosynthesis of tetramycin, and a model of the tetramycin biosynthetic pathway is proposed. Inactivation of a cytochrome P450 monooxygenase gene, *tetrK*, resulted in the production of a tetramycin B precursor: tetramycin A, which lacks a hydroxy group in its polyol region. TetrK was subsequently overexpressed heterologously in *E. coli* with a His_6 tag, and purified TetrK efficiently hydroxylated tetramycin A to afford tetramycin B. Kinetic studies revealed no inhibition of TetrK by substrate or product. Surprisingly, sequence-alignment analysis showed that TetrK, as a hydroxylase, has much higher homology with epoxidase PimD than with hydroxylases NysL and AmphL. The 3D structure of TetrK was then constructed by homology modeling with PimD as reference. Although TetrK and PimD catalyzed different chemical reactions, homology modeling indicated that they might share the same catalytic sites, despite also possessing some different sites correlated with substrate binding and substrate specificity. These findings offer good prospects for the production of improved antifungal polyene analogues.

Introduction

Polyene macrolide antibiotics include more than 200 members that exhibit broad biological activities, including antifungal, anticancer, and immunosuppressant activities.[1] As important antifungal agents, they are widely used in medicine, veterinary medicine, agriculture, and food preservation. However, there is still a need to generate novel polyene macrolide antibiotics for the treatment of fungi, due to the rapid development of drug resistance and to emerging infectious microorganisms.

Streptomyces hygrospinosus var. *beijingensis* produces a complex mixture with antifungal activity, containing tetramycin, anisomycin, and nystatin.[1b] Tetramycin consists of two components: tetramycin A and tetramycin B. Tetramycin B is a polyene macrolide with a 26-membered macrolactone ring, whereas tetramycin A lacks a hydroxy group at C-4 (Scheme 1 A). That tetramycin B has a higher antifungal activity than tetramycin A indicates that the C-4 hydroxy group of tetramycin B plays an important role in its bioactive properties.[2]

The biosynthetic mechanisms of several polyketides, such as pimaricin,[3] nystatin,[4] rimocidin,[5] candicidin (FR-008),[6] and amphotericin,[7] have been investigated in detail. The biosynthesis of a macrolide often requires not only large polyketide synthase subunits for the formation of the macrolactone ring, but also some enzymes for post-PKS tailoring steps. Three enzymes—PimD,[8] AmphL,[7] and NysL[4]—involved in the oxidation of the polyol regions as P450 monooxygenases have been studied in vivo and in vitro. The three homologous enzymes have been proved to catalyze different chemical reactions with distinct positional specificities: epoxidation of the C4=C5 double bond in pimaricin (PimD),[8] or hydroxylation either of the C-8 atom in amphotericin (AmphL)[7] or of the C-10 atom in nystatin (NysL).[4]

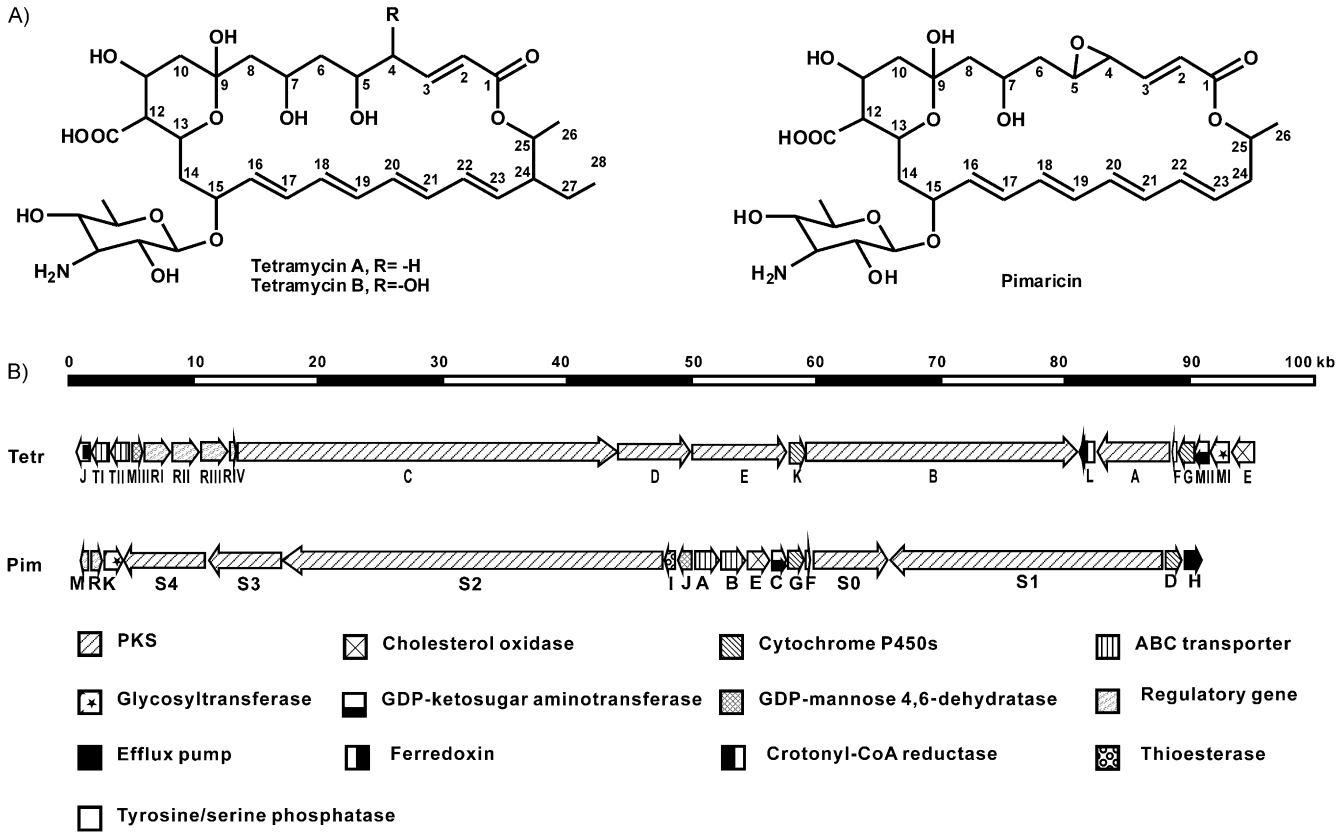
Additions of functional groups to polyene macrolide antibiotics, catalyzed by post-PKS modifications, are often crucial for their biological activities. 4,5-Deepoxypimaricin and 8-deoxyamphotericin B (intermediates obtained upon inactivation of *pimD* and *amphL*, respectively), for example, showed antifungal activities notably lower than those of pimaricin and amphotericin.[9] Studies of post-PKS modifying enzymes are therefore key to creating the structural diversity and biological activity of this class of natural products through combinatorial biosynthesis.

[a] *B. Cao,[+] F. Yao,[+] X. Zheng, D. Cui, Y. Shao, Z. Deng, D. You*
State Key Laboratory of Microbial Metabolism and
School of Life Sciences and Biotechnology, Shanghai Jiao Tong University
Shanghai 200030 (China)
E-mail: dlyou@sjtu.edu.cn

[b] *C. Zhu*
Institute of Environment and Sustainable Development in Agriculture
Chinese Academy of Agricultural Sciences
Beijing 100081 (China)

[+] *These authors contributed equally to this work.*

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201200402.



Scheme 1. Comparison of A) the structures, and B) the gene clusters of tetramycin and pimaricin.

Here we report the cloning and sequencing of the tetramycin biosynthetic gene cluster from *S. hygrospinosus* var. *beijingensis*. Inactivation of *tetrK*, which was believed to encode a P450 monooxygenase, resulted in abolition of tetramycin B production and the accumulation of tetramycin A. TetrK was then heterologously overexpressed and the kinetic parameters of the recombinant protein were studied. Homology modeling suggests that TetrK adopts a catalytic mechanism similar to that of epoxidase PimD.

Results and Discussion

Cloning of the tetramycin biosynthetic gene cluster by genome scanning

Genome scanning was employed to identify genes from *S. hygrospinosus* var. *beijingensis* that were involved in tetramycin biosynthesis. The genomic DNA was subjected to Roche 454 sequencing; this yielded an 8.83 Mb sequence on 764 assembled contigs. One identified contig contained the encoding homologues of glucosyl transferases RimE and PimK from the rimocidin and pimaricin gene clusters, flanked by other polyketide synthase (PKS) genes; it displayed high homology to the pimaricin gene cluster in sequences and organization. In order to obtain the entire proposed tetramycin biosynthetic gene cluster, we screened approximately 3000 clones of the *S. hygrospinosus* var. *beijingensis* genomic library. Through the use of the homologues of glucosyl transferase PimK and mycosamine dehydratase PimJ as probes, nine cosmids (1E12, 7G6, 11A11, 15H10, 27H2, 5C1, 5D8, 16C10, 16G8) were located in four contigs. The three gaps between the four contigs were bridged by PCR amplification. Overall, we ascertained that these four contigs contained putative tetramycin biosynthetic genes, which covered almost 92.5 kb of the chromosomal DNA. Scheme 1 shows a detailed comparison between tetramycin and pimaricin in terms both of their chemical structures (Scheme 1 A) and of their biosynthetic gene organizations (Scheme 1 B).

Sequence analysis and organization of the tetramycin biosynthetic gene cluster

Computer-assisted analysis of the sequence revealed 20 potential open reading frames (ORFs); the putative functions of the deduced gene products of the ORFs are summarized in Scheme 1 and Table 1. Consistent with the structure of tetramycin, these 20 ORFs within the *tetr* cluster comprise five genes encoding typical multifunctional type I PKSs, four genes encoding transcriptional regulators, two genes encoding cytochrome P450 monooxygenases, two genes encoding ABC transporters, three genes responsible for mycosamine biosynthesis and its attachment to the tetramycin aglycone, and four ORFs encoding a crotonyl-CoA reductase, tyrosine phosphatase, ferredoxin oxidase, and cholesterol oxidase (Scheme 1 B).

Table 1. Deduced functions of proteins encoded by the tetramycin biosynthetic gene cluster.

Protein	Amino acids	Proposed function	Sequence similarity	Overall identity/ similarity [%]	Accession number
TetrJ	450	probable crotonyl-CoA reductase	Srm4*c, *Streptomyces ambofaciens*	91/96	CAM96574
TetrTI	603	putative ABC transporter	PimB, *S. natalensis*	86/91	CAC20925
TetrTII	662	putative ABC transporter	PimA, *S. natalensis*	82/88	CAC20924
TetrMI	343	putative mycosamine dehydratase	PimJ, *S. natalensis*	92/96	CAC20923
TetrRI	965	transcriptional regulator	NysRI, *Streptomyces noursei* ATCC 11455	51/65	AAF71778
TtrRII	928	putative transcriptional regulator	NysRII, *S. noursei* ATCC 1145	45/57	AAF71779
TetrRIII	888	LAL transcriptional regulator	ShygA5, *Streptomyces hygroscopicus* ATCC 53653	53/66	ZP 0552063
TetrRIV	208	positive regulator	*ScNRII*, *Streptomyces chattanoogensis*	60/73	ACM45445
TetrC	9504	type I polyketide synthase	PimS2, *S. natalensis*	84/89	CAC20921
TetrD	1794	type I polyketide synthase	PimS3, *S. natalensis*	69/78	CAC20920
TetrE	2343	type I polyketide synthase	FscE, *S. sp.* FR-008	61/73	AAQ82567
TetrK	397	cytochrome P450 monooxygenase	PimD, *S. natalensis*	68/80	CAC20932
TetrB	6786	type I polyketide synthase	PimS1, *S. natalensis*	77/84	CAC20931
TetrL	374	tyrosine phosphatase	RimC, *Streptomyces diastaticus*	64/75	AAR16515
TetrA	1736	type I polyketide synthase	RimA, *S. diastaticus*	73/81	AAR16521
TetrF	64	ferredoxin	RimH, *S. diastaticus*	74/87	AAR16520
TetrG	392	cytochrome P450 monooxygenase	PimG, *S. natalensis*	81/90	CAC20928
TetrMII	353	aminotransferase	RimF, *S. diastaticus*	86/92	AAR16518
TetrMIII	463	putative glucosyl transferase	RimE, *S. diastaticus*	81/90	AAR16517
TetrO	554	cholesterol oxidase	PimE, *S. natalensis*	88/94	CAC20926

In comparison to pimaricin biosynthesis,[3] this gene cluster has a different extender unit, one post-PKS modification step, and a more complicated regulation system, whereas it lacks an efflux pump and a discrete thioesterase (TE).

The five tetramycin PKSs encoded by *tetrA–tetrE* show significant similarity to the corresponding pimaricin PKSs, as might be expected from the structural similarity between the two tetraenes. Analysis of the five multimodular PKSs by searching the database for nonribosomal peptide-synthetase (NRPS) and PKS (http://www.nii.res.in/nrps-pks.html) showed that they encoded exactly 13 PKS modules, consistent with 13 steps in the synthesis of the carbon skeleton of tetramycin.

TetrA, like PimS0,[3] represents a discrete loading module with acetyl-CoA as the starter unit. Replacement of the active site cysteine by a serine residue in the condensing ketosynthase (KS) domain is detected in the KS^S of TetrA. In TetrB, module 1 contains a butyryl-CoA-specific acyl transferase (AT) domain, whereas in the pimaricin PKS module 1 a malonate-specific AT is present.[3] The AT domain of module 7 in TetrC incorporates propionate to contribute the exocyclic methyl group, which later undergoes oxidation to form a carboxylic acid at C-12, whereas the other nine ATs belong to the class of acetate extenders. Module 12 contains one more enoyl reductase (ER) domain than PIM PKSs, but it is probably inactive, leaving a double bond at C2=C3. Meanwhile, disruption of *tetrK* showed that the DH domain of module 11 should be inactive, although the conserved motif HxxxGxxxxP was found.[10] In addition, although the two PKSs each have a TE domain in the last module, tetramycin synthetases lack a discrete TE, which is present in the pimaricin gene cluster.[3] In view of the fact that butyryl-CoA serves as the first extender unit for TetrB, TetrJ was predicted to be a crotonyl-CoA reductase, catalyzing the formation of butyryl-CoA from two acetyl-CoA units. TetrJ has a high similarity (76%) to RimJ, the function of which has been confirmed by gene disruption to be to provide butyryl-CoA as the starter unit.[5]

As for the genes *tetrMI*, *tetrMII*, and *tetrMIII*, their products are predicted to be a GDP-mannose 4,6-dehydratase, a GDP-ketosugar aminotransferase, and a glycosyltransferase, respectively. Their homologous genes in nystatin biosynthesis had been confirmed to be involved in the biosynthesis and attachment of the mycosamine moiety.[11] The protein products of two genes in the cluster, *tetrTI* and *tetrTII*, display high degrees of similarity to those proteins belonging to the ATP-dependent ABC transporter superfamily, so both of them are believed to be involved in tetramycin secretion from the cell. The counterpart of *pimH*,[3] an extra gene coding for a putative efflux pump, is absent in the tetramycin gene cluster.

The tetramycin gene cluster seems to have a regulation system more complicated than that of pimaricin.[3] Four putative regulatory genes—*tetrRI*, *tetrRII*, *tetrRIII*, and *tetrRIV*, all of which are transcribed in the same direction—were identified upstream of *tetrC*. Interestingly, the *tetrRI–RIII* genes have organization similar to that of the regulatory genes *nysRI–RIII*,[12] and their protein products have sequence identities of 51, 45, and 50%, respectively. The regulators belong to the LAL (large ATP binding regulator of the LuxR) family, which is usually considered to be involved in maintaining the stabilities of the extremely long mRNAs of the large PKS genes. The counterpart genes are absent, however, in the pimaricin gene cluster. *TetrRIV* has a sequence identity of 59% to *pimM*, which functions as a transcriptional activator in the pimaricin gene cluster.[13] In addition, *tetrO*, at the border of the gene cluster, encodes a cholesterol oxidase, which was predicted to have no effect on polyene macrolide biosynthesis until Mendes proved that the cholesterol oxidase PimE acted as a signaling protein for the biosynthesis of pimaricin.[14]

TetrG shares 81% sequence identity with the cytochrome P450 monooxygenase PimG from *Streptomyces natalensis*. *TetrF*, lying downstream of *tetrG*, encodes a ferredoxin. A *tetrG* analogue, *amphN*,[15] has been demonstrated to be responsible for generating the carboxylic acid group of amphotericin. TetrG and TetrF are therefore probably responsible for the oxidation of the methyl group at C-12 of tetramycin.

Confirmation of the function of the *tetr* gene cluster by gene replacement and complementation

To test whether these genes are involved in the biosynthesis of tetramycin, a PKS gene (*tetrB*) was disrupted by replacement of an approximately 3-kb region with the apramycin resistance cassette (Figure S1 in the Supporting Information). As expected, the production of tetramycin in the mutant CB1 was completely abolished; this confirmed that the deleted gene was indeed essential for tetramycin biosynthesis (Figure 1).

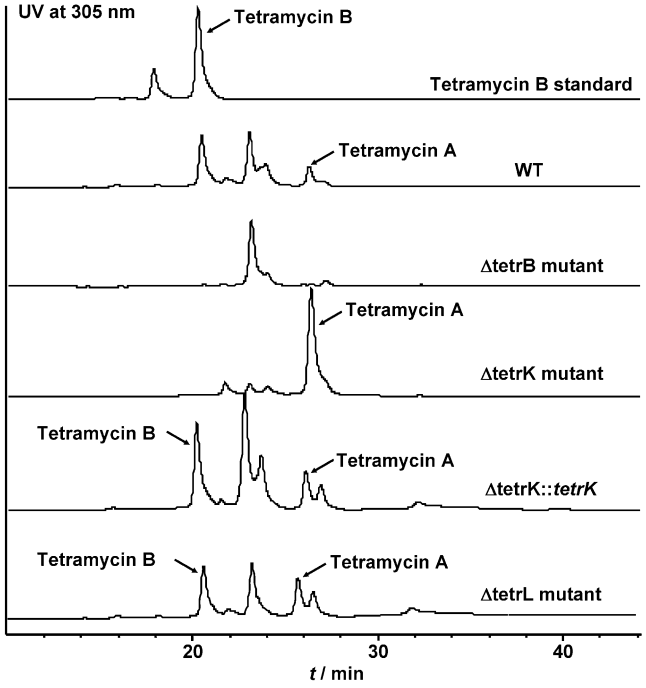


Figure 1. LC-MS analysis of tetramycin production in wild-type *S. hygrospinosus* var. *beijingensis* and in mutants Δ*tetrB*, Δ*tetrK*, Δ*tetrK* complemented by *tetrK*, and Δ*tetrL*.

It was thus proposed that TetrK, with high homology to cytochrome P450 hydroxylases, catalyzes the hydroxylation of tetramycin A (Scheme 2). To verify that TetrK is required for the formation of tetramycin B, we inactivated *tetrK* by replacing it with an apramycin resistance cassette (Figure S2). As shown in Figure 1, tetramycin B production was completely abolished in the resultant mutant strain CB2, whereas tetramycin A was still produced and accumulated. Genetic complementation of *tetrK* back into the mutant CB2 restored the production of tetramycin B (Figure 1). This strongly suggested that TetrK was indeed involved in the hydroxylation of the tetramycin polyketide moiety at C-4. Production of both tetramycins A and B in the wild-type strain can be explained in terms of TetrK being unable, for some unknown reason, to convert tetramycin A completely into tetramycin B in vivo.

Between *tetrA* and *tetrB* is *tetrL*, encoding for a 374 amino-acid protein, predicted to be a protein tyrosine/serine phosphatase, with considerable homology to RimC (58% identity). Deletion of *rimC* suggested that RimC might be related to the choice of starter units by the loading module, or might modify the balance of carboxylic acids inside the cell.[5] When aligned with PimS0, about 100 amino acid residues near the C terminus of TetrL shared 75% identity with the acyl carrier protein (ACP) domain of PimS0, whereas the rest shared 44% identity with protein tyrosine/serine phosphatases. This protein is therefore likely a fused enzyme of ACP and tyrosine/serine phosphatase formed during the evolution of the tetramycin biosynthetic pathway. However, the disruption of *tetrL* had no effect on tetramycin production, indicating that TetrL likely plays no role in tetramycin biosynthesis (Figure 1).

Characterization and kinetic properties of recombinant TetrK

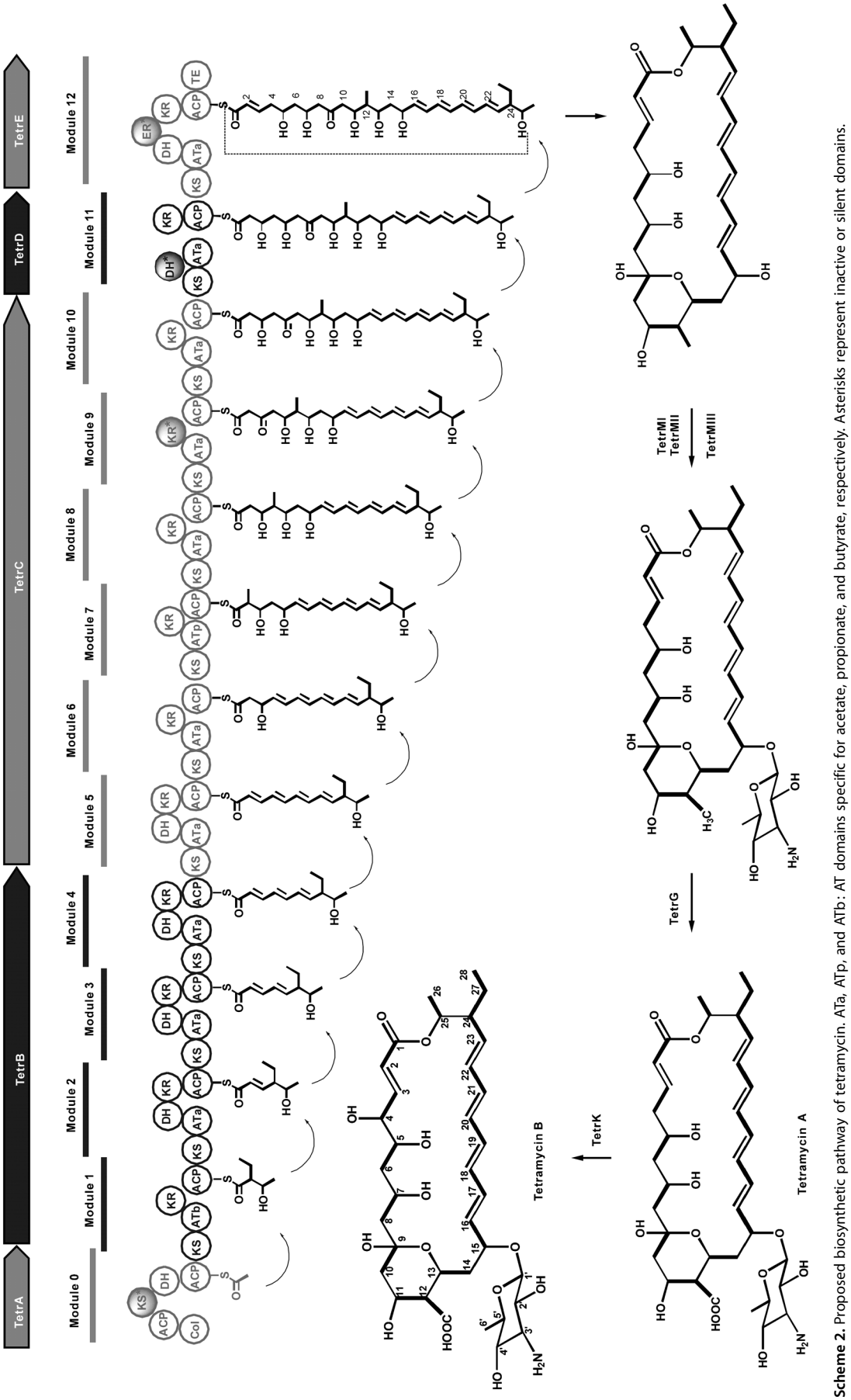
The TetrK protein was successfully heterologously expressed in *E. coli*; 9 mg of purified protein with an N-terminal His_6 tag was obtained from 1 L fermentation liquor. In vitro experiments showed that the recombinant TetrK efficiently converted tetramycin A into tetramycin B at 30 °C in the presence of spinach ferredoxin, spinach ferredoxin-$NADP^+$ reductase, and NADPH, with glucose-6-phosphate and glucose-6-phosphate dehydrogenase participating as a NADPH-regenerating system. Figure 2A shows the conversion result, as measured by HPLC, over different reaction times.

After the optimal conditions for the TetrK activity assay had been determined to be a reaction temperature of 30 °C and pH 7.5, the kinetic parameters of the recombinant TetrK-catalyzed reaction were studied. With an enzyme concentration of 2 nM and a reaction time of 20 min, the reaction could be fitted well to the Michaelis–Menten equation at concentrations of tetramycin A in the range of 1 to 200 μM (Figure 2B). The K_m and V_{max} values were then determined to be 64 ± 11 μM and 4.6 ± 0.9 μM min^{-1}, respectively, corresponding to a k_{cat} value of 2.3 ± 0.21 s^{-1}. This represents a K_m value twice that of the similar reaction catalyzed by PimD and a correspondingly greater V_{max}, resulting in a k_{cat} value that was three times higher.[8]

Sequence alignments and homology modeling of TetrK

Three P450 enzymes involved in oxidative modifications of the polyol segment of polyene macrolides—PimD, AmphL, and NysL—have so far been characterized in vivo and in vitro. Sequence alignment analysis demonstrated that TetrK had greater homology with the epoxidase PimD than with hydroxylases NysL and AmphL, and the X-ray structure of PimD was reported recently.[8,16]

Homology models of TetrK were built by using the 3D structure of PimD as a reference. Comparative modeling structures



Scheme 2. Proposed biosynthetic pathway of tetramycin. ATa, ATp, and ATb: AT domains specific for acetate, propionate, and butyrate, respectively. Asterisks represent inactive or silent domains.

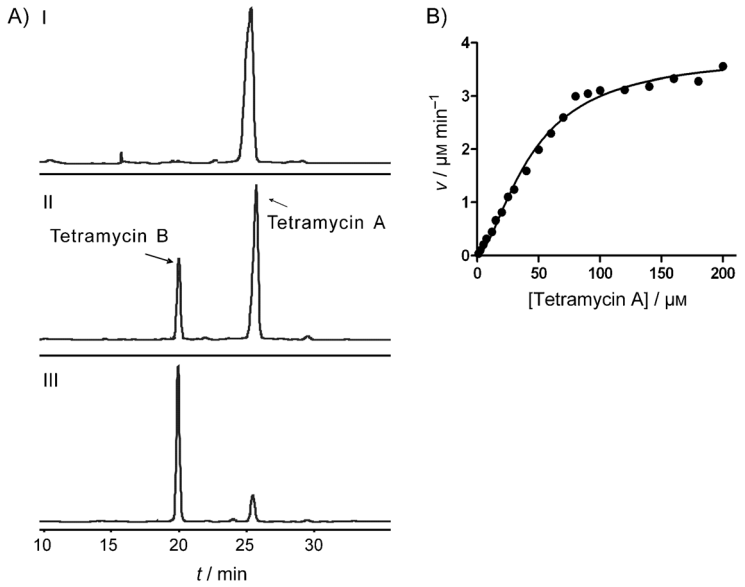


Figure 2. In vitro reactions and kinetics of conversion of tetramycin A into tetramycin B in the presence of TetrK. A) Results of HPLC analysis (detection wavelength 304 nm) after reaction times of: 1) 0 min, 2) 15 min, and 3) 30 min. B) Kinetic study of this reaction at substrate concentration ranges of 1 to 200 μM.

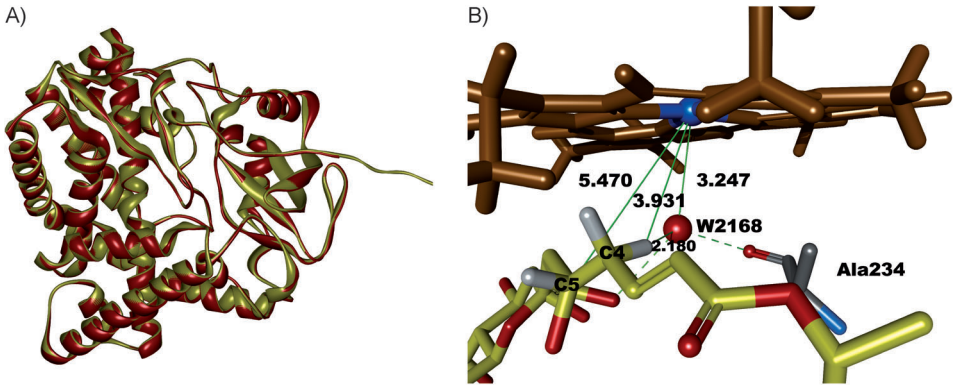


Figure 3. Homology-modeled structures of TetrK. A) Alignment analysis of overall structures of TetrK (green) and PimD (red). B) Docking of tetramycin A and heme iron in the model of the TetrK active site.

for TetrK showed few differences in the overall structure of PimD (Figure 3A), especially in the substrate binding and catalytic site. Sequence and secondary structure alignment, with the aid of the ENDscript web server,[17] between the predicted secondary structures of the TetrK model and the reference protein PimD showed that they shared a high degree of similarity both in secondary structures and in amino acid sequences (Figure 4). In TetrK, however, a Pro288 is absent in the loop region proposed to correlate with the size of substrates in PimD (Figure 4).

The structure of substrate-bound PimD showed that the protein provided two H-bonds to 4,5-desepoxypimaricin: between the hydroxy group at C-11 and Ser241, and between the exocyclic carboxylic acid group at C-12 and Ser283.[16] In TetrK, Ser241 was replaced with a threonine residue, which was also the case in AmphL. Although Ser283 was conserved in the four cytochrome P450 monooxygenases (Figure 4), it has proved unnecessary for NysL-catalyzed hydroxylation.[18]

Despite its high structural similarity with PimD, TetrK was believed to function through a different catalytic mechanism. It had been suggested that PimD-catalyzed epoxidation of 4,5-desepoxypimaricin involves a hydroperoxoferric intermediate (Fe^{III}–OOH, compound 0), which acts as an oxidant, inserting its distal oxygen atom into the C4=C5 double bond.[16] However, as the C4–H σ-bond points directly towards the heme Fe (3.931 Å, Figure 3B), hydroxylation of tetramycin A by TetrK could be explained by the consensus hydroxylation mechanism,[19] in which P450 compound I (Fe^{IV}=O intermediate) would abstract hydrogen from C-4 of tetramycin A to form an iron(IV) hydroxide that would subsequently hydroxylate the substrate radical to yield tetramycin B.

Conclusions

Polyene macrolide antibiotics exhibit potent antifungal activities. We have sequenced and identified a complete tetramycin biosynthetic gene cluster from *S. hygrospinosus* var. *beijingensis* and have also obtained a mutant that produced only one compound, which accumulated owing to a blocked biosynthetic pathway. Furthermore, the hydroxylase TetrK was studied in detail by measuring the kinetic parameters and homology modeling. Surprisingly, the tetramycin biosynthetic cluster was very different in its genetic organization from those of pimaricin and rimocidin, thus extending our knowledge on the biosynthetic mechanisms of polyene macrolides, providing a foundation for genome mining of polyene macrolide antibiotics, and opening up new prospects for obtaining a series of polyene derivatives with improved pharmacological properties, increased activities, reduced toxicities, or improved water solubilities by combinatorial biosynthesis.

Experimental Section

Bacterial strains, plasmids, and culture conditions: *S. hygrospinosus* var. *beijingensis*, the wild-type producer for tetramycin, was obtained from the Chinese Academy of Agricultural Sciences (ACCC40068). The cosmid pJTU2463,[20] a derivative of pOJ446 with *int* and *attP* from pSET152, was used for construction of a genomic library. pJTU968 (pRSET-B derivative *bla PermE**)[21] and pPM927 (*tsr oriT int attP*)[22] were used for *tetrK* complementation. Streptomyces and its derivatives were grown at 30 °C in TSBY liquid medium

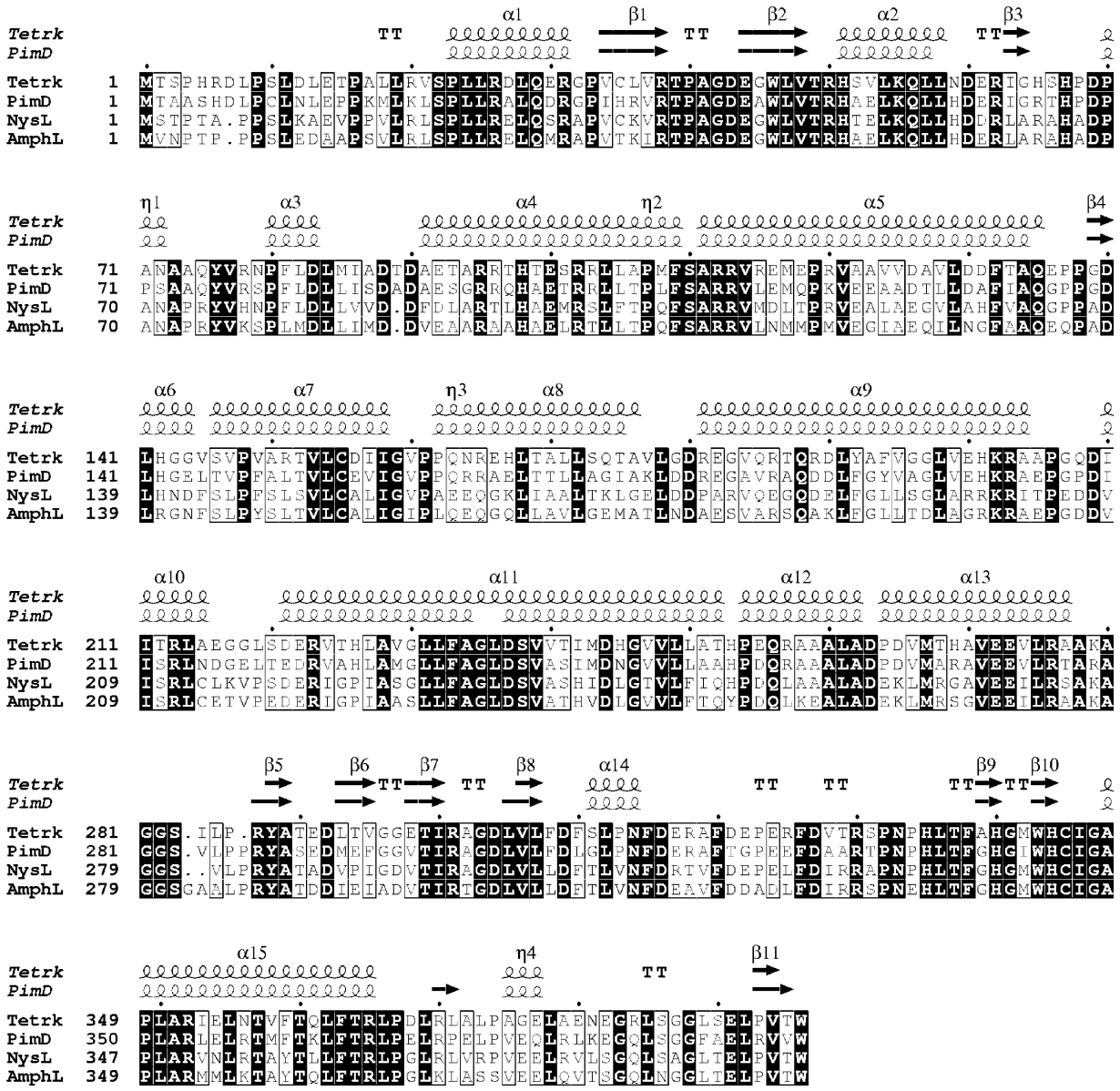


Figure 4. Sequence and secondary structure alignments of polyene macrolide monooxygenases.

[tryptic soy broth powder (Difco, 3%), sucrose (10.3%), yeast extract (Difco, 1%), pH 7.2] or SFM solid medium [agar (2%), mannitol (2%), soybean powder (2%), pH 7.2]. Lysogeny broth (LB) and agar were used for culturing *E. coli* strains. All plasmid subcloning experiments were performed in *E. coli* DH10B by standard protocols. The final antibiotic concentrations used for the selection of *E. coli* and *Streptomyces* were: thiostrepton (5 $\mu g\,mL^{-1}$) and apramycin (15 $\mu g\,mL^{-1}$) for *Streptomyces*, and apramycin (30 $\mu g\,mL^{-1}$), kanamycin (50 $\mu g\,mL^{-1}$), ampicillin (100 $\mu g\,mL^{-1}$), and chloramphenicol (12.5 $\mu g\,mL^{-1}$) for *E. coli*.

Genomic library construction and screening: A pOJ446-derived cosmid, pJTU2463, was used to construct the genome library of *S. hygrospinosus* var. *beijingensis*, according to a standard protocol,[23,24] and the competent EPI300-T1^R *E. coli* was selected as host.[25] To locate the tetramycin gene cluster, two pairs of primers were used for screening of the genome library (*tetrJ–S*: 5′-CCTCGG TGACGA GTTCCG CCTGCTG-3′, *tetrJ–A*: 5′-GGTGCC CCGTCC CGAACT GCTCATG-3′, and *tetrK–S*: 5′-CCGACC GGGCGA TGCTGC GCTGGAC-3′, *tetrK–A*: 5′-GAGGAC GGACGG GGTTCG GCGAGC AC-3′). The whole gene cluster sequence was completed after subcloning and sequencing of the PCR product.

DNA sequencing and analysis: The genome was sequenced at Roche's 454 Sequencing Center, and 704 large contigs were obtained. Potential ORFs were identified by use of GLIMMER 3.0 and BLAST analysis. Sequence analysis with the FramePlot 3.0 beta online program (http://watson.nih.go.jp/~jun/cgi-bin/frameplot-3.0b.pl) identified four contigs covering the genes homologous with the pimaricin gene cluster. After screening of the cosmid library by PCR, the remaining gaps were filled by targeted subcloning and the use of specific primers. The structure prediction of the five multi-modular PKS was carried out by database searching for NRPS and PKS (http://www.nii.res.in/nrps-pks.html).

Construction and complementation of the mutants: All of *tetrB*, *tetrL*, and *tetrK* were deleted by insertional inactivation through double crossing over by the *aac(3)IV/oriT* cassette (Figures S1–S3).

The mutated plasmids were introduced into *S. hygrospinosus* var. *beijingensis*, by conjugation with use of *E. coli* ET12567/pUZ8002. The double crossover mutants were screened against apramycin and confirmed by PCR. After hundreds of double crossover mutant colonies had been obtained, four single spores of each mutant were independently used for LC-MS analysis of their products.

The gene *tetrK* was then complemented back into the *tetrK* mutant. The gene was amplified with KOD-plus DNA polymerase with use of chromosomal DNA of *S. hygrospinosus* var. *beijingensis* as a template and primers *tetrK–F* (5′-TTACAT ATGACC TCCCCC CACCGT GATCT-3′; NdeI sites underlined) and *tetrK–R* (5′-CACGAA TTCCTA CCAGGT CACCGG CAGCTC-3′; engineered EcoRI site underlined). The PCR amplification was carried out under the following conditions: initial denaturation at 94 °C for 5 min, then 28 cycles of 45 s at 94 °C, 45 s at 58 °C, and 1 min at 68 °C. A final elongation step was performed at 68 °C for 5 min. The 1190 bp purified NdeI/EcoRI-digested PCR product was ligated to NdeI/EcoRI-digested pJTU968 (pRSET-B derivative *bla PermE**). The 1.5 kb MfeI/EcoRI fragment, containing the *PermE** promoter and *tetrK*, was then cleaved from pJTU968 and ligated into EcoRI-digested pPM927 to give the *tetrK*-complemented plasmid. The recombinant plasmid was then introduced into the *tetrK* mutant stain by conjugation. Four single spores of the *tetrK*-complemented mutant were selected for fermentation and analysis by LC-MS.

Fermentation, extraction, and LC-MS analysis of tetramycin: *S. hygrospinosus* var. *beijingensis*, both wild type and recombinant mutants, were grown on SFM solid medium at 30 °C for four days. The culture medium and mycelia were then collected and extracted with acetone. The extract was concentrated in vacuo and finally dissolved in methanol (1 mL) for the LC-MS analysis, with use of the Agilent 1100 series LC/MSD Trap system. The LC was operated at a flow rate of 0.6 mL min^{-1} with an Agilent TC-C18 column (4.6× 2500 mm, 5 μm), and the parameters were as follows: 25 °C; solvent A: ammonium acetate in water (0.005 M); solvent B: acetonitrile; gradient: from 20% B to 80% B over 40 min; 95% B for 10 min; UV absorbance detection: 305 nm. The ion-trap mass spectrometer was acquired with the electrospray ionization source in the positive ion mode (Figure S4).

Heterologous expression of TetrK: The *tetrK* gene was amplified by PCR and inserted into the N-terminal His_6 tag expression vector pET-28a. The forward primer *tetrK-s* (5′-CTACAT ATGACC TCCCCC CACCGT GATCT-3′) introduced a unique NdeI site at the 5′-end of the gene, whereas the reverse primer *tetrK-anti* (5′- GGCGAA TTCCTA CCAGGT CACCGG CAGCTC-3′) carried a HindIII site downstream from the translational stop codon. The amplified DNA fragment was digested with NdeI and HindIII and was then cloned into the same sites of pET-28a, generating the recombinant plasmid pJTU6001. After confirmation of the fidelity of the inserted gene by DNA sequencing, pJTU6001 was transformed into *E. coli* BL21 (DE3) for protein expression. The transformants were grown at 37 °C in LB broth (1 L) containing selective antibiotics (50 μg mL^{-1}, kanamycin for pET-28a and chloromycetin for plusS) until an OD_{600} of 0.8–1.0 was reached. The cells were then induced with isopropyl D-thiogalactoside (IPTG, 0.1 mM) and allowed to continue to grow at 12 °C for 30 h. The induced cells were harvested by centrifugation and resuspended in lysis buffer [Tris·HCl (pH 8.0, 50 mM), NaCl (500 mM), imidazole (20 mM), 40 mL]. Cell disruption was accomplished on ice by sonication. The soluble fraction was collected by centrifugation (15000 *g*, 30 min at 4 °C) and applied onto a HisTrap HP column (GE Healthcare, 1 mL). The proteins were eluted with a linear gradient of buffer B [Tris·HCl (pH 8.0, 50 mM), NaCl (150 mM), imidazole (500 mM)] with an ÄKTA fast protein liquid chromatography system (GE health). Eluted fractions were analyzed by SDS-PAGE and the His-tagged TetrK was concentrated with Centrifugal Filter Units (Millipore, 30 kDa). Subsequent desalting was achieved by buffer exchange into desalting buffer [Tris·HCl (pH 8.0, 50 mM), NaCl (150 mM)].

Purification of tetramycin A: Crude tetramycin A was extracted from the *tetrK* mutant CB2 by the procedure described above for tetramycin B. Tetramycin A was then purified with an Agilent 1200 series HPLC and an Agilent TC-C18 column (Agilent, 4.6×250 mm, 5 μm) with elution with a gradient (1 mL min^{-1}) of methanol (70% to 80%) for 30 min. Tetramycin was detected at the retention time of 25 min. The pure tetramycin A was eluted in methanol after quantitative determination with an Agilent 2100 Bioanalyzer.

Enzyme assay and kinetic studies of TetrK: The assay of TetrK activity and kinetic studies were carried out on a 100 μL scale with tetramycin A (1 to 200 μM), His_6-TetrK (10 to 30 nM), spinach ferredoxin (100 μg mL^{-1}), spinach ferredoxin-NADP$^+$ reductase (0.2 U mL^{-1}), NADPH (1.4 mM), glucose-6-phosphate (10 mM), and glucose-6-phosphate dehydrogenase (8 U mL^{-1}) in Tris·HCl (pH 7.5, 50 mM). The reactions were carried out at 30 °C for 20 min and terminated by the addition of four volumes of butanol. The resulting organic extracts were dried and resuspended in methanol for HPLC analysis. The substrates and products were quantified by use of standard curves for tetramycin A and tetramycin B respectively, according to the absorbing area at 304 nm.

Sequence alignment and homology modeling of TetrK: Sequence analysis of identified homologues for TetrK protein sequences was carried out by searching PSI-BLAST (comparison matrix, BLOSUM62; E-threshold 10) in the NCBI (National Center for Biotechnology Information) website. The 3D model structure of TetrK was performed by homology modeling with use of PimD (PDB IDs: 2X9P and 2XBK), which shared 81% amino acid identity to TetrK, as a template. The structure of TetrK was predicted and evaluated by use of Discovery Studio (DS) v2.5 (Accelrys, Inc., http://www.accelrys.com). The template was aligned with the target and examined for conserved sequences. The aligned sequences were taken for the model construction and built by use of "build homology model" in DS. The coordinates for the heme were obtained from 2XBK and positioned as in the template. The structure was then refined by energy minimization. The model was refined by use of CHARMm in DS 2.5, which provides powerful mechanics and dynamics protocols for studying the energetics and motion of molecules, from small ligands to multicomponent physiological complexes. Accelrys CHARMm force field was used throughout the simulation. The refined model was validated with a ProSA2003[26] *Z*-score and a PROCHECK[27] Ramachandran plot.

Nucleotide sequences accession number: The sequence of the gene cluster for tetramycin biosynthesis was deposited in GenBank with accession number JN688154.

Acknowledgements

This work was supported by grants from the National Science Foundation of China, the Ministry of Science and Technology (973 and 863 Programs), China Ocean Mineral Resources R & D Association, and the Shanghai Municipal Council of Science and Technology.

Keywords: antifungal agents · cytochrome P450 monooxygenase · natural products · polyene macrolide antibiotics · polyketide synthases

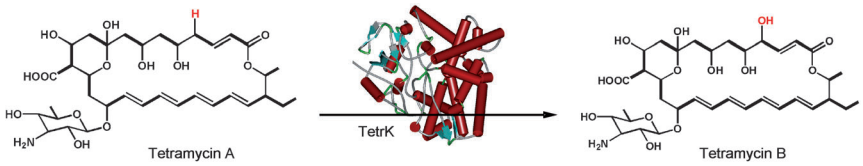
[1] a) D. J. Newman, G. M. Cragg, *J. Nat. Prod.* **2007**, *70*, 461 – 477; b) F. Yao, Q. J. Kang, S. J. Lin, C. X. Zhu, D. L. You, Z. X. Deng, *J. Shanghai Jiaotong Univ.* **2011**, *4*, 88 – 91.
[2] M. I. L. Radics, K. Dornberger, *Tetrahedron* **1982**, *38*, 183 – 189.
[3] J. F. Aparicio, R. Fouces, M. V. Mendes, N. Olivera, J. F. Martin, *Chem. Biol.* **2000**, *7*, 895 – 905.
[4] O. Volokhan, H. Sletta, T. E. Ellingsen, S. B. Zotchev, *Appl. Environ. Microbiol.* **2006**, *72*, 2514 – 2519.
[5] E. M. Seco, F. J. Perez-Zuniga, M. S. Rolon, F. Malpartida, *Chem. Biol.* **2004**, *11*, 357 – 366.
[6] S. Chen, X. Huang, X. Zhou, L. Bai, J. He, K. J. Jeong, S. Y. Lee, Z. Deng, *Chem. Biol.* **2003**, *10*, 1065 – 1076.
[7] P. Caffrey, S. Lynch, E. Flood, S. Finnan, M. Oliynyk, *Chem. Biol.* **2001**, *8*, 713 – 723.
[8] M. V. Mendes, N. Antón, J. F. Martín, J. F. Aparicio, *Biochem. J.* **2005**, *386*, 57 – 62.
[9] a) B. Byrne, M. Carmody, E. Gibson, B. Rawlings, P. Caffrey, *Chem. Biol.* **2003**, *10*, 1215 – 1224; b) M. V. Mendes, E. Recio, R. Fouces, R. Luiten, J. F. Martin, J. F. Aparicio, *Chem. Biol.* **2001**, *8*, 635 – 644.
[10] J. Wu, T. J. Zaleski, C. Valenzano, C. Khosla, D. E. Cane, *J. Am. Chem. Soc.* **2005**, *127*, 17393 – 17404.
[11] A. Nedal, H. Sletta, T. Brautaset, S. E. Borgos, O. N. Sekurova, T. E. Ellingsen, S. B. Zotchev, *Appl. Environ. Microbiol.* **2007**, *73*, 7400 – 7407.
[12] T. Brautaset, O. N. Sekurova, H. Sletta, T. E. Ellingsen, A. R. Strøm, S. Valla, S. B. Zotchev, *Chem. Biol.* **2000**, *7*, 395 – 403.
[13] N. Antón, J. Santos-Aberturas, M. V. Mendes, S. M. Guerra, J. F. Martín, J. F. Aparicio, *Microbiology* **2007**, *153*, 3174 – 3183.
[14] N. Anton, M. V. Mendes, J. F. Martin, J. F. Aparicio, *J. Bacteriol.* **2004**, *186*, 2567 – 2575.
[15] M. Carmody, B. Murphy, B. Byrne, P. Power, D. Rai, B. Rawlings, P. Caffrey, *J. Biol. Chem.* **2005**, *280*, 34420 – 34426.
[16] P. M. Kells, H. Ouellet, J. Santos-Aberturas, J. F. Aparicio, L. M. Podust, *Chem. Biol.* **2010**, *17*, 841 – 851.
[17] P. Gouet, X. Robert, E. Courcelle, *Nucleic Acids Res.* **2003**, *31*, 3320 – 3323.
[18] T. Brautaset, H. Sletta, K. F. Degnes, O. N. Sekurova, I. Bakke, O. Volokhan, T. Andreassen, T. E. Ellingsen, S. B. Zotchev, *Appl. Environ. Microbiol.* **2011**, *77*, 6636 – 6643.
[19] J. Rittle, M. T. Green, *Science* **2010**, *330*, 933 – 937.
[20] W. Chen, D. Qu, L. Zhai, M. Tao, Y. Wang, S. Lin, N. P. Price, Z. Deng, *Protein Cell* **2010**, *1*, 1093 – 1105.
[21] H. Xu, Y. Zhang, J. Yang, T. Mahmud, L. Bai, Z. Deng, *Chem. Biol.* **2009**, *16*, 567 – 576.
[22] T. Smokvina, P. Mazodier, F. Boccard, C. J. Thompson, M. Guerineau, *Gene* **1990**, *94*, 53 – 59.
[23] J. Sambrook, E. F. Fritsch, T. Maniatis in *Molecular Cloning*, 2nd ed., Cold Spring Harbor Laboratory Press, New York, **1989**, p. 502.
[24] B. M. Kieser, K. F. Chater, M. J. Butter, D. Hopwood in *Genetic Manipulation of Streptomyces: A Laboratory Manual*, John Innes Foundation, Norwich, **2000**, p. 311 – 334.
[25] W. Chen, T. Huang, X. He, Q. Meng, D. You, L. Bai, J. Li, M. Wu, R. Li, Z. Xie, H. Zhou, X. Zhou, H. Tan, Z. Deng, *J. Biol. Chem.* **2009**, *284*, 10627 – 10638.
[26] M. J. Sippl, *Proteins Struct. Funct. Bioinf.* **1993**, *17*, 355 – 362.
[27] R. A. Laskowski, M. W. MacArthur, D. S. Moss, J. M. Thornton, *J. Appl. Crystallogr.* **1993**, *26*, 283 – 291.

Received: June 13, 2012
Published online on ■■ ■■, 0000

*B. Cao, F. Yao, X. Zheng, D. Cui, Y. Shao, C. Zhu, Z. Deng, D. You**

■■ – ■■

Genome Mining of the Biosynthetic Gene Cluster of the Polyene Macrolide Antibiotic Tetramycin and Characterization of a P450 Monooxygenase Involved in the Hydroxylation of the Tetramycin B Polyol Segment



Genome mining has identified a polyene macrolide antibiotic tetramycin biosynthetic gene cluster. A biosynthetic pathway was proposed and confirmed by genetic and in silico analyses. In vitro study and homology modeling of a cytochrome P450 monooxygenase, TetrK, revealed its kinetic properties and catalytic mechanism. These findings could help to produce improved antifungal polyene analogues.

