A Gene Cluster Encoding Resistomycin Biosynthesis in *Streptomyces resistomycificus*; Exploring Polyketide Cyclization beyond Linear and Angucyclic Patterns

Kathrin Jakobi and Christian Hertweck*

Hans-Knöll-Institute for Natural Products Research, Department of Bioorganic Synthesis, Beutenbergstrasse 11a, 07745 Jena, Germany

Received October 15, 2003; E-mail: christian.hertweck@hki-jena.de

Resistomycin (*rem*, **3**) is an unusual aromatic polyketide metabolite of *Streptomyces resistomycificus*[1] that exhibits a variety of pharmacologically relevant properties, for example, inhibition of HIV-1 protease,[2] as well as RNA and DNA polymerase, and activity against Gram-positive bacteria and mycobacteria.[3] More recently, **3** has also been implicated as a modulator of apoptosis and may thus serve as a valuable tool in cell biology.[4]

Scheme 1. Model for the Biosynthesis of Angular (**1**) and Linear (**2**) Dekaketides versus "Discoid" Resistomycin (**3**); Diverging Pathways of **2** and **3**
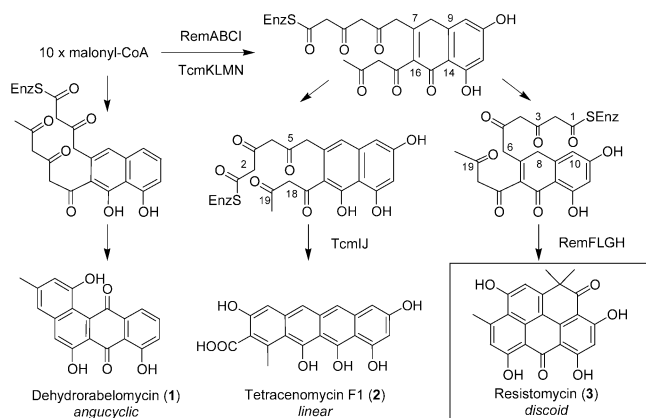
Streptomycetes synthesize aromatic polyketides by type II polyketide synthases (PKS), a complex of iteratively used individual proteins. Each aromatic PKS contains a "minimal" set of enzymes required for polyketide chain elongation, consisting of two β-keto-acylsynthase subunits, KS_α and KS_β, and an acyl carrier protein (ACP). Additional PKS subunits, including keto-reductases, cyclases, and aromatases, convert the elongated poly-β-ketoacyl thioester to the cyclized polyketide molecule prior to further tailoring reactions (e.g., oxidation, methylation, glycosylation).[5] A number of type II PKS gene clusters have been cloned and sequenced, and a large body of knowledge on polyketide biosynthesis has been obtained over the past years.[5] However, to date, the precise nature and mechanism of aromatic PKSs has not been unveiled, and pathway engineering is hampered by a lack of understanding on the detailed programming of chain propagation and cyclization of the nascent poly-β-keto intermediates, which suppresses spontaneous aldol reactions. Considering the vast number of bacterial aromatic polyketides, it is remarkable that virtually all polyphenols are formed by a U-shape polyketide folding and thus can be grouped into linear and angular polyketides, for example, dehydrorabelomycin (**1**)[6] and tetracenomycin F1 (**2**).[5,7] The cumulated ("discoid") ring system of **3** is a clear exception. Classical ^{13}C-labeling experiments suggested the naphthanthrone derivative **3** is generated by an S-shape folding of the decaketide (Figure 1).[8]
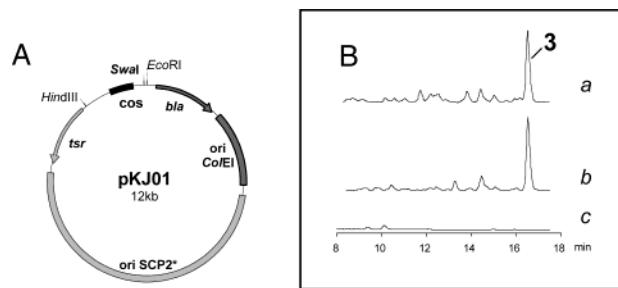

Figure 1. (A) Map of shuttle cosmid vector pKJ01; (B) HPLC-MS analyses of wild type (a), *S. lividans* TK23/pKJ05 (b), and host as the control (c).

Exploring the molecular basis of resistomycin biosynthesis may not only provide novel biosynthetic tools to generate metabolic diversity, but would also allow a new insight into the controlled cyclization of the highly reactive polyketide intermediates.

The strong orange fluorescence of **3** at 366 nm appeared ideal for a phenotype screening approach to clone the *rem* biosynthetic gene cluster, bypassing time-consuming genetic screening procedures. For this purpose, we constructed a self-replicating *E. coli–Streptomyces* shuttle cosmid vector (pKJ01). A similar construct already proved to be capable of harboring up to 40 kb genomic fragments,[9] more than sufficient to cover the predicted size of the *rem* gene cluster (<15 kb). A *S. resistomycificus* cosmid library was constructed by ligation of blunted 38−45 kb genomic fragments into the *Swa*I site of pKJ01, followed by transfection of *E. coli* EPI 100. All resulting clones were pooled, and the entire cosmid library was introduced into *S. lividans* TK23 by PEG induced protoplast transformation.[10] Positive transformants were selected for thiostrepton resistance and grown on R5 plates for 5 days. Out of about 800 clones, one transformant (*S. lividans* TK23/pKJ05) was detected by its orange fluorescence under UV light. Fermentation of this clone and HPLC−MS analysis of the extract unequivocally proved the identity of this product of *S. lividans* TK23/pKJ05 with **3** (Figure 1). Thus, the cosmid was expected to contain the entire set of *rem* biosynthetic genes.

Plasmid DNA was recovered and subjected to shotgun sequencing. Frame analysis of the 15 kb *rem* biosynthesis gene cluster (GenBank accession nr. AJ585192), which is flanked by genes involved in primary metabolism and housekeeping, revealed 18 open reading frames (ORFs) that consist of 11 structural genes and 7 genes for regulation and resistance. Putative functions of deduced gene products were assigned by homology with known protein sequences in the databases. Analysis of the *rem* gene cluster disclosed several very unusual features. The PKS gene cassette clearly deviates from the common $KS_\alpha/KS_\beta/ACP$ architecture. First, the probable *rem* ACP (RemC) has only weak homology to its counterparts from known aromatic PKSs, but rather resembles ACPs

Protein	Deduced function	Protein	Deduced function
RemA	KSα	RemJ	Kinase
RemB	KSβ	RemK	Unknown
RemC	ACP	RemL	Cyclase
RemD	PPTase	RemM	Regulator
RemE	MCAT	RemN	Efflux
RemF	Cyclase	RemO	Oxygenase
RemG	MT	RemP	Regulator
RemH	MT	RemQ	Regulator
RemI	Cyclase	RemR	Regulator

Figure 2. Organization of the resistomycin (*rem*) biosynthesis gene cluster and deduced functions (table). PKS and associated genes (black), post PKS (grey), regulation, resistance, unknown (white).

from fatty acid synthases. Second, with only a very few exceptions,[9,11] minimal PKS genes are grouped (Figure 2). In the *rem* cluster, genes encoding the *rem* $KS_α/KS_β$ heterodimer (*remAB*) are separated from the ACP gene *remC* by one inserted ORF (*remD*). The deduced gene product of *remD* has high homology to phosphopantetheinyl transferases (PPTases) that transform *apo*-ACPs into activated *holo*-ACPs. In addition, much to our surprise, a gene (*remE*) that codes for a malonyl-CoA:ACP transacylase (MCAT) was identified downstream of the minimal PKS genes. MCATs are usually not encoded by type II PKS gene clusters, and to date the enzymology of malonyl-CoA transfer in type II PKS has not been firmly established in vivo. In vitro experiments demonstrate that the *act* holo-ACP is capable of self-catalyzing its own malonation at high concentrations of malonyl-CoA.[12] An alternative model is the recruitment of MCATs from fatty acid synthases (FAS).[13] Out of the high number of cloned and sequenced type II PKSs, yet only a few putative MCATs have been detected.[14] However, our most unexpected and striking finding is that the deduced gene product of *remE* is more closely related to AT domains from modular polyketide synthases (similarity/identity: 51%/37%) than to MCATs from type II PKSs and FASs (<42%/30%).

The biosynthesis of **3** probably involves several cyclases/aromatases for the controlled formation of the multicyclic decaketide from the nascent poly-β-keto intermediate. On the basis of the structures of **3** and tetracenomycin (**3**, *tcm*), we assumed that both polyketides share the first biosynthetic steps prior to branching into a linear and a discoid decaketide. In agreement with our model, the protein encoded by *remI* shows high homology to first and second ring cyclases/aromatases in unreduced systems, such as TcmN and WhiE-ORFVI.[5] Accordingly, the *remABCI* cassette is functionally equivalent to *tcmKLMN* (Scheme 1). Not surprisingly, the downstream processing enzymes that are implicated in the unprecedented S-shape polyketide folding share much lower homology to known cyclases. The deduced gene products of *remF* and *remL*, which are related to TcmJ and the AknW/ZhuJ family of cyclases,[9] respectively, are the candidate cyclases for changing the direction of chain folding. It needs to be established whether the nascent poly-β-keto intermediate will be processed by individual enzymes, which may have dual functions, or alternatively, by a multienzyme complex; PKS and cyclases may form a cage, in which the correct folding of the polyketide chain is enhanced and stabilized, ultimately dictating the shape of the resulting molecule. One may also reason that upon formation of the fourth ring (3/8),

the fifth cyclization (1/10) occurs spontaneously by attack of the highly nucleophilic aromatic ring system on the activated acyl-ACP intermediate. Functional studies are in progress in our lab, which will soon shed more light on this.

In the biosynthetic scheme, the resulting bis-*nor*-resistomycin is subject to a yet unprecedented geminal bis-methylation.[13]C-labeling experiments by Höfle et al. revealed that both *gem* methyl groups of **3** are derived from *S*-adenosyl methionine (SAM).[8] In the deduced gene products of *remG* and *remH*, conserved motifs were identified that are characteristic for probable SAM-binding methyltransferases, strongly suggesting that two individual enzymes are involved in this unusual PKS tailoring reaction. A further post PKS transformation may be catalyzed by a FAD-dependent oxygenase, the deduced gene product of an ORF located on the second transcript (*remO*). A very likely role for RemO is hydroxylation of **3** yielding resistoflavin,[15] a minor metabolite of *S. resistomycificus*.

In summary, we have successfully identified a novel PKS gene cluster by heterologously expressing an entire cosmid library and screening for the fluorescence of the metabolite produced. The *rem* gene cluster exhibits several unusual features of the type II PKS involved, most remarkably, a putative MCAT with highest homology to AT domains from modular PKSs. In addition, while all bacterial aromatic polyketides are either linear or angular, we provide the first insight into the molecular basis of a unique mode of cyclization giving rise to a discoid polyketide. The work presented here thus sets the basis for gaining a deeper understanding of the mechanisms of polyketide cyclization and may ultimately provide new opportunities for engineering novel polyketide ring systems.

Supporting Information Available: Material and methods, table of genes of the *rem* cluster and deduced functions (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Brockmann, H.; Reschke, T. *Tetrahedron Lett.* **1968**, *27*, 3167−3170. Recently, a chlorinated derivative of **1** has been isolated from a *Bacillus sp.*: Magyarosy, A.; Ho, J. Z.; Rapoport, H.; Dawson, S.; Hancock, J.; Keasling, J. D. *Appl. Environ. Microbiol.* **2002**, *68*, 44095−44101.
(2) Roggo, B. E.; Petersen, F.; Delmendo, R.; Jenny, H. B.; Peter, H. H.; Roesel, J. *J. Antibiot.* **1994**, *47*, 136−142.
(3) Haupt, I.; Waehnert, U.; Pitra, C.; Loeber, G.; Luck, G.; Eckardt, K. *Z. Allg. Mikrobiol.* **1975**, *15*, 411−421.
(4) Shiono, Y.; Shiono, N.; Seo, S.; Oka, S.; Yamazaki, Y. *Z. Naturforsch.* **2002**, *57c*, 923−929.
(5) Rawlings, B. J. *Nat. Prod. Rep.* **1999**, *16*, 425−484.
(6) Han, L.; Yang, K.; Ramalingam, E.; Mosher, R. H.; Vining, L. C. *Microbiology* **1994**, *140*, 3379−3389.
(7) Summers, R. G.; Wendt-Pienkowski, E.; Motamedi, H.; Hutchinson, C. R. *J. Bacteriol.* **1993**, *1993*, 7571−7580.
(8) Höfle, G.; Wolf, H. *Liebigs Ann. Chem.* **1983**, 835−843.
(9) Marti, T.; Hu, Z.; Pohl, N. L.; Shah, A. N.; Khosla, C. *J. Biol. Chem.* **2000**, *275*, 33443−33448.
(10) Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A. *Practical Streptomyces Genetics*; The John Innes Foundation: Norwich, U.K., 2000.
(11) Li, A.; Piel, J. *Chem. Biol.* **2002**, *9*, 1017−1026. Grimm, A.; Madduri, K.; Ali, A.; Hutchinson, C. R. *Gene* **1994**, *151*, 1−10.
(12) Matharu, A.-L.; Cox, R. J.; Crosby, J.; Byrom, J.; Simpson, T. J. *Chem. Biol.* **1998**, *5*, 699−711.
(13) Summers, R. G.; Ali, A.; Shen, B.; Wessel, W. A.; Hutchinson, C. R. *Biochemistry* **1995**, *34*, 9389−9402.
(14) Novakova, R.; Bistakova, J.; Homerova, D.; Rezuchova, B.; Kormanec, J. *Gene* **2002**, *297*, 197−208. Fischer, C.; Lipata, F.; Rohr, J. *J. Am. Chem. Soc.* **2003**, *125*, 7818−7819. Rajgarhia, V. B.; Priestley, N. D.; Strohl, W. R. *Metab. Eng.* **2001**, *3*, 49−63. Meadows, E. S.; Khosla, C. *Biochemistry* **2001**, *40*, 14855−14861.
(15) Eckardt, K.; Fritzsche, H.; Tresselt, D. *Tetrahedron* **1970**, *26*, 5875−5883.

JA0390698