Environmental Microbiology (2003) doi:10.1046/j.1462-2920.2003.00531.x

Targeting modular polyketide synthases with iteratively acting acyltransferases from metagenomes of uncultured bacterial consortia

Jörn Piel,[1]* Dequan Hui,[1] Nobuhiro Fusetani[2] and Shigeki Matsunaga[2]

[1]*Max Planck Institute for Chemical Ecology, Department of Bioorganics, Hans-Knöll-Str. 8, Beutenberg Campus, 07745 Jena, Germany.*

[2]*Laboratory of Marine Biochemistry, Faculty of Agriculture, The University of Tokyo, Bunkyo-ku, Tokyo 113, Japan.*

Summary

Bacterial type I polyketide synthases (PKSs) produce a wide range of biomedically important secondary metabolites. These enzymes possess a modular structure that can be genetically re-engineered to yield novel drug candidates not found in nature. Recently, we have reported the putative pederin PKS from an uncultured bacterial symbiont of *Paederus fuscipes* beetles. It belongs to an architecturally unusual PKS group, the members of which contain iteratively acting acyltransferases that are not integrated into the PKS modules but are encoded by isolated genes. As these systems are rare, often contain additional unusual features and are of smaller size than regular PKSs, the development of a method for the targeted isolation of new group members would be of great interest. Here, we present a phylogenetic approach to identify these systems rapidly in highly complex metagenomic DNA samples. To demonstrate its practical value, we located two pederin-type PKS systems putatively involved in the biosynthesis of antitumour polyketides in the metagenomic DNA of beetles, sponges and their uncultivated bacterial symbionts.

Introduction

Bacterial type I polyketide synthases (PKSs) generate a vast array of pharmaceutically important natural products. Clinically used examples are the antibiotic erythromycin, the immunosuppressant FK506 or the antiparasitic avermectin derivatives. During polyketide biosynthesis, the giant, modular PKS enzymes serve as a scaffold to which the growing polyketide chain is tethered (Rawlings, 2001). Each PKS module carries out one chain elongation step through attachment of a small acyl-CoA extension unit. The chain is thus transferred from module to module until it reaches its final size and is released from the enzyme. Owing to this repetitive catalytic mechanism, PKS modules are architecturally similar to each other and usually feature a ketosynthase (KS) domain catalysing the chain elongation, an acyltransferase (AT) domain that selects the correct acyl-CoA extender and an acyl carrier protein domain (ACP) serving as anchor to the polyketide chain. In addition to this minimal domain set, individual modules can carry various domains performing reduction, dehydration or methylation reactions for further substrate modification. The modular nature of the PKS biosynthetic machinery permits a rational engineering of these enzymes, yielding novel polyketides with largely predictable structures (Cane *et al*., 1998).

Recently, we have isolated a PKS with an architecture that deviates markedly from this general, erythromycin-type scheme. In an investigation on the biotechnological potential of uncultured bacterial symbionts of invertebrates, we have isolated the putative pederin biosynthesis (*ped*) gene cluster from a symbiont of *Paederus* spp. beetles (Piel, 2002). It encodes structurally unusual PKS proteins, the modules of which are entirely devoid of AT domains. Instead, the system contains two small enzymes resembling monodomain ATs. We proposed that these monofunctional ATs act *in trans* to acylate each PKS module (Piel, 2002). These novel enzymes would therefore be part of an unusual mixed iterative–modular PKS (for examples of modular PKSs using other types of iterative mechanisms, see Wilkinson *et al*., 2000; Gaitatzis *et al*., 2002; Mochizuki *et al*., 2003). Subsequently, the leinamycin biosynthesis gene cluster (*lnm*) from *Streptomyces globisporus* (Cheng *et al*., 2003), the mupirocin system (*mup*) from *Pseudomonas fluorescens* (El-Sayed *et al*., 2003) and the lankacidin (*lkc*) cluster from a *Streptomyces* sp. plasmid (Mochizuki *et al*., 2003) were reported as additional examples of a PKS with discrete ATs, and functional studies of the *lnm* ATs verified the *trans*-acting mechanism (Cheng *et al*., 2003). A sur-

Received 4 August, 2003; accepted 3 September, 2003. *For correspondence. E-mail piel@ice.mpg.de; Tel. (+49) 3641 571213; Fax (+49) 3641 571256.

vey of GenBank entries and published partial gene clusters reveals a more widespread occurrence of PKSs with iterative ATs. Among such '*trans*-AT', pederin-type systems are the *pksX* cluster from *Bacillus subtilis* with unknown function (Scotti *et al*., 1993), the TA (= myxovirescin) biosynthesis cluster (*ta*) from *Myxococcus xanthus* (Paitan *et al*., 1999) and most probably the albicidin synthase (*xab*) from *Xanthomonas albilineans* (Huang *et al*., 2001). *Trans*-AT clusters are therefore found in a wide range of bacteria, including Gram-negative α- and δ-proteobacteria and Gram-positive bacteria of the low- and high-GC groups. In addition to the AT architecture, these systems often contain further unusual features not known from cis-AT, erythromycin-type PKSs. Examples are enzymes catalysing the attachment of carbon units at the extender carbonyl position (*ta*, *mup* and *lnm* clusters), novel or unusual sets of catalytic domains (PedF module 5, LnmJ modules 3, 4 and 8, many *pksX* modules) or extremely small-sized clusters with a highly repetitive enzyme usage (*lkc*). Additionally, they are of smaller size than regular PKSs. These systems should therefore provide new insights into the enzymology of PKS machineries and be useful tools in the production of novel polyketides by genetic engineering of biosynthetic pathways. A strategy for the targeted isolation of additional members of this rare type would thus be of great interest.

As reported by Neilan and coworkers, phylogenetic data from type I PKS systems reveal that the phylogeny of erythromycin-type proteins largely mirrors that of their host bacteria (Moffitt and Neilan, 2003). As *trans*-AT systems were not included in the analysis, we were interested in their evolutionary relationship to regular, 'cis-AT' PKSs. Did the enzymes evolve independently in various bacterial phyla or did this event take place only once, and species barriers were overcome by horizontal gene transfer? In the latter case, sequence information from a short PKS gene fragment generated by polymerase chain reaction (PCR) would probably be sufficient to identify a type I PKS as *cis*- or *trans*-AT type.

As we exemplify in this study, the answer to this question can indeed provide important guidance in the targeted isolation of rare *trans*-AT gene clusters. We demonstrate that these systems can be located rapidly even within highly complex and uncultivated microbial consortia. As a practical application of this approach, we have studied two associations of invertebrates and their bacterial symbionts. Symbiotic bacteria are postulated to be the true producers of a large number of drug candidates isolated from animals (Haygood *et al*., 1999; Proksch *et al*., 2002), but none of these producing microorganisms has so far been cultivated successfully. Long-term supplies of animal-derived drug candidates could be generated by cloning of biosynthetic gene clusters and their heterologous expression in culturable bacteria (Haygood *et al*., 1999; Piel, 2002). However, the search for specific genes is hampered by the biological and metabolic complexity of invertebrate–bacterial communities. In the work presented here, we have applied a phylogenetic approach to clone genes putatively involved in the biosynthesis of antitumor compounds from the metagenomic DNA of beetles and sponges and their bacteria.

Results

Phylogenetic analysis reveals a monophyletic origin of trans*-AT PKSs*

For the construction of phylogenetic trees, various PKS sequences from *cis*- and *trans*-AT-type proteins of diverse origin were retrieved from the GenBank database. The sequences were truncated to the KS region between the EPIAIV (corresponding to residues 31–36 of DEBS2 from *Saccharopolyspora erythraea*) and the HGTGT motifs (DEBS2 residues 337–341). This region was chosen because the KS domain exhibits the highest degree of conservation with only a few hypervariable stretches. Furthermore, it largely corresponds to the region that is usually amplified by PCR from genomic DNA as an initial step in the isolation of novel PKS gene clusters (Beyer *et al*., 1999; Piel, 2002; Moffitt and Neilan, 2003). Newly obtained PCR products can therefore be correlated directly with the phylogenetic tree.

As shown in Fig. 1, the resulting tree shows two distinct clades containing exclusively *cis*- or *trans*-AT sequences. The tight clustering of *trans*-AT PKSs regardless of the bacterial phylum indicates that these enzymes have a single evolutionary origin. As reported previously, *cis*-AT enzymes cluster into groups that largely mirror the phylogeny of their host organism (Moffitt and Neilan, 2003). In unexpected contrast, the evolutionary gene tree of *trans*-AT enzymes does not correlate with the bacterial species tree in spite of good bootstrapping values. For example, the well-separated clade containing PksL3 and LnmJ3 harbours KSs from a high-GC Gram-positive *Streptomyces* sp., a low-GC Gram-positive *Bacillus* sp. and two Gram-negative *Pseudomonas* spp. This may indicate a higher rate of horizontal transfer and/or a higher tendency of the proteins to be expressed and folded correctly in new host strains after such transfers, compared with proteins of the *cis*-AT type. Nevertheless, the distinct clustering of KS sequences into two PKS clades now allows us to predict the PKS type of novel gene fragments amplified by PCR without the necessity to clone larger gene regions. In the following sections, we applied this approach to the identification of *trans*-AT systems in the metagenomic DNA of uncultured bacterial symbionts.

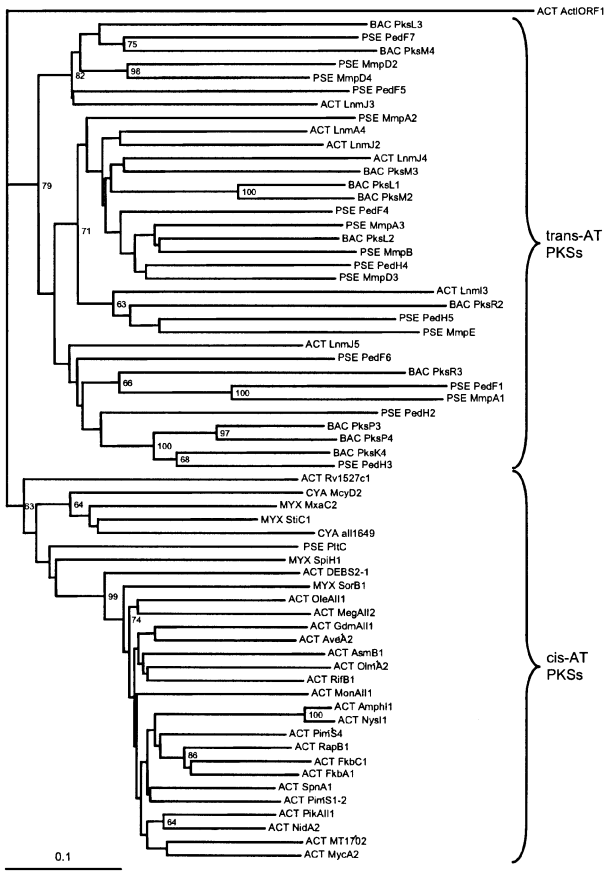

Fig. 1. Phylogenetic analysis of bacterial type I KS domains. The bacterial origin of the sequence is designated as ACT, Actinobacteria; BAC, Bacilli; PSE, Pseudomonadales; CYA, Cyanobacteria; MYX, Myxococcales. The last digit of each entry specifies the module number from which the KS sequence was derived. Bootstrap values larger than 60% are shown. The type II KS ActI ORF1 was used as outgroup.

Identification of trans*-AT PKS genes from a metagenomic library of bacteria isolated from* Paederus fuscipes *beetles*

We have recently cloned a set of putative pederin biosynthesis genes from a metagenomic library of *Paederus fuscipes* beetles and its associated bacteria (Piel, 2002). The beetles use the antitumour agent pederin (Fig. 2) as a chemical defence (Kellner and Dettner, 1996), but sequence analysis revealed that the gene cluster resides on the genome of an as yet unculturable *Pseudomonas* sp. bacterium associated with the insects (Piel, 2002). According to its architecture, the *ped* region does not encode the three PKS modules necessary for the biosynthesis of the starter triketide. Genes for this starter PKS should therefore be found on a separate genomic region. According to the strategy outlined above, we performed a PCR analysis of the DNA of the bacterial community isolated from *P. fuscipes* beetles. As the PKS should lack AT domains like the remaining part of the system, any novel PCR product that can be placed into the rare *trans*-AT clade would with high likelihood belong to the wanted gene(s). Analysis of 35 cloned PCR products revealed three novel sequences, designated PS1 to PS3, that did not belong to the previously sequenced gene cluster. Figure 3 shows the phylogenetic tree generated from an alignment containing these sequences. As the PCR products were obtained with primers based on the DPQQRLL (DEBS2 residues 118–124) and HGTGT motifs, they were slightly shorter than the GenBank sequences. Although the exclusion of several amino acids from the tree calculation resulted in lowered bootstrapping values, the tree topology was almost identical to that of the previously calculated tree. Therefore, the PKS type of even the shortened PCR products can be determined by phylogenetic analysis.

According to the phylogenetic data, two of the three KS fragments from beetle bacteria indeed clearly associated with the *trans*-AT clade (sequences PS1 and PS2; Fig. 3), whereas the third (PS3) was more related to *cis*-AT PKSs. As the amplified PCR product did not cover the region of an eventual AT domain, we wished to verify the PKS type by cloning and sequencing of an extended region of one of the pederin-type sequences, PS2. To this purpose, we screened a metagenomic

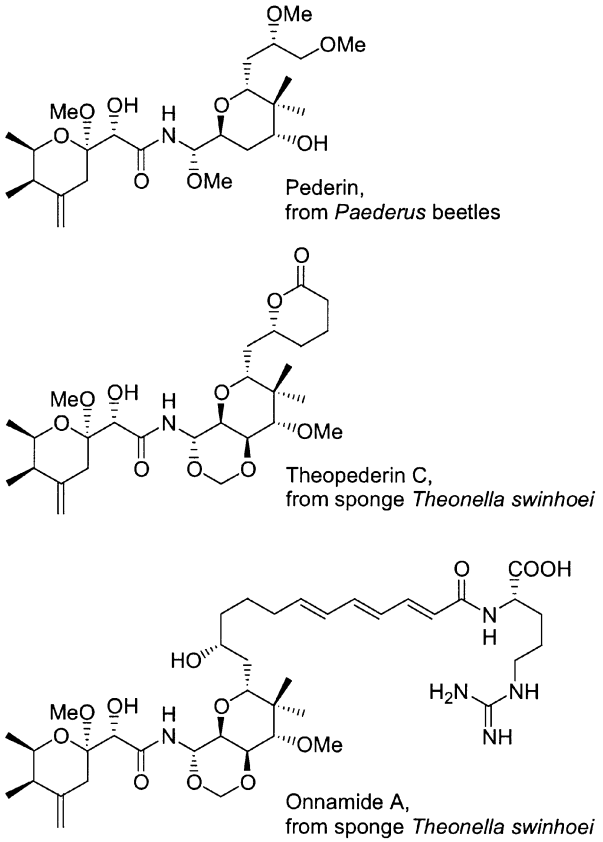

Fig. 2. Antitumour compounds isolated from *Paederus* spp. beetles and the sponge *Theonella swinhoei*.

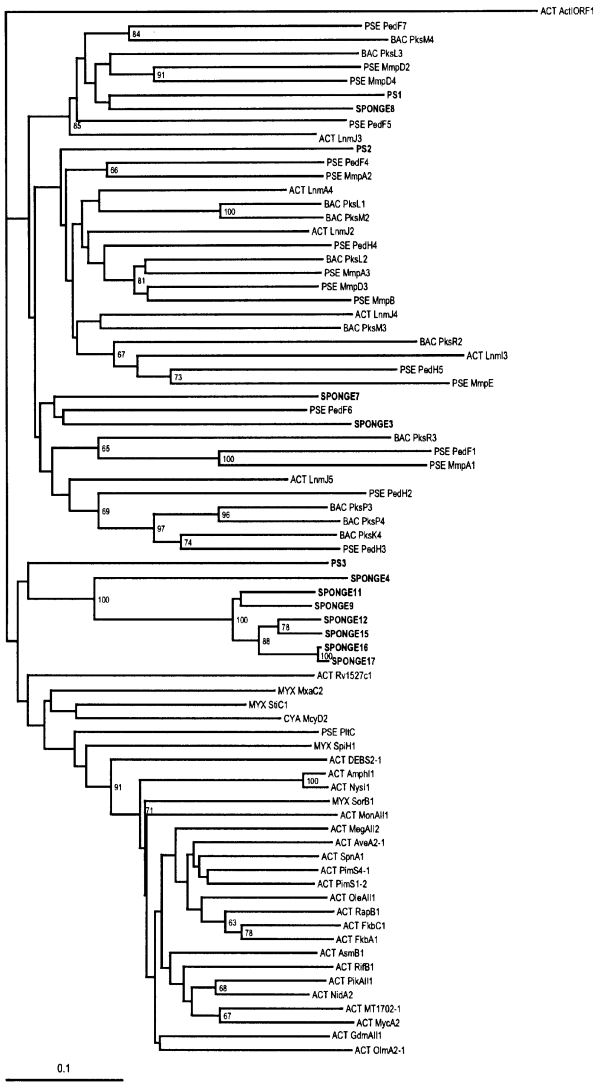

Fig. 3. Phylogenetic analysis of known KS domains and those amplified by PCR from metagenomic DNA. Entries starting with PS are derived from *Paederus fuscipes* bacteria; those starting with SPONGE are from the total DNA of *Theonella swinhoei*. Eleven additional sponge sequences with very high sequence identity (>95%) to members of the group containing SPONGE11 and SPONGE17 were not included in the tree diagram.

cosmid library constructed from bacterial DNA and isolated several cosmids containing the gene. As expected, sequencing of a common 3.1 kb *Sph*I fragment revealed a KS next to a KR domain, but no homology to an AT domain (Fig. 4). Preliminary sequence data indicate that the PKS region is of small size and located close to genes with high homology to *Pseudomonas aeruginosa* (data not shown). The gene should therefore belong to the same *Pseudomonas* sp. symbiont as the previously identified *ped* cluster. Current studies will clarify whether it encodes the PKS generating the pederin triketide starter.

Localization of trans*-AT PKS genes in the highly complex metagenome of the sponge* Theonella swinhoei

The sponge *Theonella swinhoei* is known for its extraordinary diversity in bioactive complex polyketides and nonribosomal peptides (Narquizian and Kocienski, 2000; Matsunaga and Fusetani, 2003). As these natural product groups are typical of the microbial secondary metabolism, it is likely that they are produced by bacteria rather than by the sponge. Strikingly, some of these substances, the onnamides and theopederins (Fig. 2) (Narquizian and Kocienski, 2000), are structurally almost identical to pederin from *Paederus* spp. beetles. We speculated that metabolites of this complexity should have evolved from a common biosynthetic pathway rather than by a convergent mechanism. The onnamide PKS should therefore represent a *trans*-AT system.

For the study of natural product symbiosis in sponges, *T. swinhoei* represents an ideal model, because the similarity between the onnamide and pederin biosynthesis pathways could provide a relatively rapid access to sponge symbiont genes. We suspected, however, that the physical separation of the pathways in beetles and sponges over millions of years has resulted in a high sequence divergence between the two PKS systems. This would probably prohibit a direct PCR amplification of onnamide genes relying on sequence similarity. In accordance, PCR with various degenerate primer pairs specific

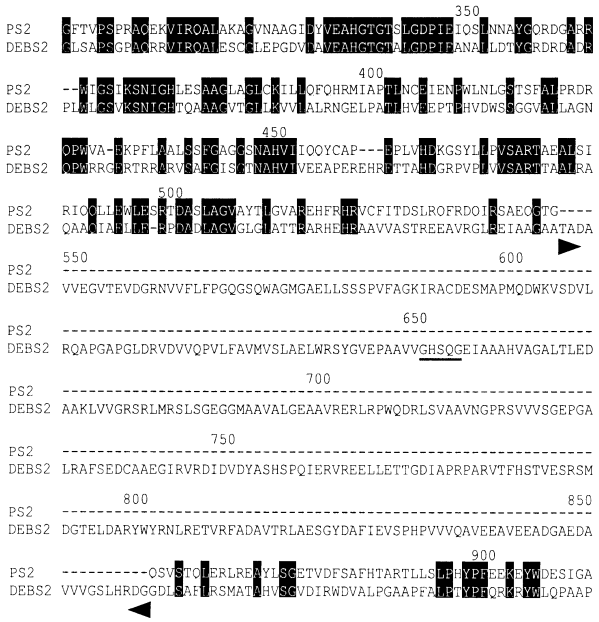

Fig. 4. Local alignment of the erythromycin biosynthesis protein DEBS2 with the protein sequence encoded by the cloned 3.1 kb *Sph*I fragment (PS2). Only the region around the AT domain is shown. The arrows mark the borders of the AT domain lacking in PS2. The conserved active site motif GHSxG is underlined. The amino acid numbering refers to DEBS2.

to the *ped* PKS did not yield the expected products from sponge DNA (data not shown). A more promising approach is the phylogenetic analysis of PCR products obtained with general PKS primers, which could reveal sequences belonging to gene clusters of the *trans*-AT type. Owing to the scarcity of these clusters, the likelihood is high that an identified *trans*-AT sequence belongs to the onnamide system.

As expected for a highly complex microbial community, PCR with KS-specific primers amplified a large number of different products from sponge metagenomic DNA. Restriction fragment length polymorphism and sequence analysis of 60 PCR products revealed 21 different sequences, all of which exhibited the highest homology to the KS domains of type I PKSs. When these were placed into the general PKS tree, most located near the cyanobacterial/myxobacterial region of the *cis*-AT clade (Fig. 3). These sequences formed a distinct subgroup, the members of which were often highly similar to each other. However, three sequences, SPONGE3, SPONGE7 and SPONGE8, clustered into the *trans*-AT clade and associated close to sequences derived from beetle bacteria. The metagenomic DNA of *T. swinhoei* should therefore contain at least one rare *trans*-AT PKS system, which can now be isolated directly from a metagenomic library with the aid of specific gene probes to determine its possible role in onnamide biosynthesis.

Discussion

More than 99% of all bacteria are unculturable with existing techniques (Kaeberlein *et al*., 2002). These microorganisms represent a vast and virtually unexplored resource of pharmacologically active natural products. However, because of the inherent technical difficulties, only a few studies have successfully attempted to unlock their enormous biosynthetic potential (Lorenz *et al*., 2002). One approach is the random heterologous expression of gene clusters present on environmental DNA fragments (Wang *et al*., 2000; Brady *et al*., 2001; Gillespie *et al*., 2002). As the size of such fragments is currently limited to around 70–80 kb (Rondon *et al*., 2000), medically important substance classes generated by large gene clusters, such as complex polyketides or non-ribosomal peptides, have not yet been produced by such expression libraries. A second strategy to target metabolites of uncultured microbes is the amplification of PCR products covering small gene fragments of a specific enzyme family, followed by the isolation of the entire gene clusters. Although even large PKS systems have been isolated by this approach (Piel, 2002), the systematic search for genes of interest is usually hampered by the presence of numerous related genes in a metagenomic DNA sample. In the case of PKSs, strategies have been developed that permit a more directed search by classifying a bacterial PCR product according to functional types (PKS or hybrid PKS-non-ribosomal peptide synthetase, type II PKS subgroups) and, to some extent, host lineages (Metsa-Ketela *et al*., 1999; Moffitt and Neilan, 2003). Here, we report that bacterial type I PKSs can be phylogenetically subdivided further into *cis*-AT and *trans*-AT systems, allowing for the rapid identification of specific and unusual PKS types. A preliminary phylogenetic comparison of such systems on the basis of a small set of AT sequences has been reported recently (Cheng *et al*., 2003). However, these enzymes lack suitable conserved regions that could be used in PCR-based screenings.

At present, our phylogenetic data do not allow a conclusion regarding which group of PKS represents the evolutionary older type. A general trend in the evolution of metabolic enzymes is the formation of multidomain proteins by fusion of separate genes (Snel *et al*., 2000). The development of the PKS-related animal multidomain fatty acid synthases from individual proteins, as found in prokaryotes, is an example (McCarthy and Hardie, 1984). Such fusion events lead to a balanced translation of catalytic domains and optimize the spatial arrangement for substrate shuttling. An insertion of an AT domain into a PKS module could have resulted in similar evolutionary advantages. Furthermore, all iterative *trans*-ATs known so far incorporate exclusively malonyl-CoA into every position of the polyketide chain, which is sometimes further modified by methylation. *Cis*-ATs have a wider substrate range and can control polyketide structure regiospecifically at every extension step. The resulting enhanced structural diversity could have fostered an evolution of *cis*-AT systems.

Type I PKSs are currently among the most important enzymes that produce natural and rationally designed drug candidates (Cane *et al*., 1998). Our data bear several implications for drug discovery from bacteria. *Trans*-AT systems deviate far more frequently from the general PKS rules than *cis*-AT PKSs. Most of them contain unusual and novel features that should not only be useful tools in the generation of structural diversity through engineered polyketides by combinatorial biosynthesis, but should also improve our understanding of PKS enzymology. The data presented in this work now allow for rapid detection of such clusters in even complex, metagenomic DNA mixtures as the initial step of a targeted isolation.

A further practical application has been demonstrated in this work by the identification of *trans*-AT PKS candidates involved in the biosynthesis of the symbiont antitumour compound pederin. Various PKS sequences were obtained by PCR from the microbial consortium of *Paederus* spp. beetles. The phylogenetic analysis revealed that these belong to at least two different PKS systems, and that one of them should exhibit an architec-

ture unrelated to pederin biosynthesis. Thus, genes of interest were identified among related sequences without the need to isolate and sequence larger portions of every gene cluster present in the library. This approach was taken one step further in the very large metagenomic library of the sponge *T. swinhoei*. Sponges contain a remarkable number of as yet uncultured bacteria (Hentschel *et al*., 2002), which can comprise up to half the total biomass (Vacelet, 1971). Several metabolites from sponges that are currently in preclinical or clinical trial phases are suspected symbiont products, such as discodermolide or halichondrin (Proksch *et al*., 2002). The isolation of biosynthesis gene clusters from sponge symbionts in order to create a long-term supply by heterologous expression is therefore of great biomedical interest (Haygood *et al*., 1999). Such an approach, however, is severely complicated by the complexity of the sponge metagenome. The method presented here has for the first time permitted us to target a specific gene cluster from sponge metagenomic DNA that could encode the biosynthesis of a drug candidate. Future work aims at isolating and sequencing the *trans*-AT PKS genes to study their role in onnamide biosynthesis.

Experimental procedures

Cloning and sequencing of PKS gene fragments from P. fuscipes *and* T. swinhoei

Preparation and screening of a cosmid library prepared from metagenomic DNA of *P. fuscipes* beetles collected in Aydın, Turkey, was performed as described previously (Piel, 2002). Metagenomic DNA from *T. swinhoei* sponges collected at Hachijo-Jima, Japan, was prepared from an ethanol-stored specimen using the QIAamp DNA mini kit (Qiagen) according to the directions given in the tissue protocol. Alternatively, bacteria were isolated by processing a freshly collected sponge in Ca^{2+}- Mg^{2+}-free artificial sea water in a Panasonic MJ-C28 juicer, centrifuging at 100 *g* for 10 min and passing the supernatant through a 40 μm nylon mesh (Millipore). Bacteria were then collected by centrifugation at 5000 *g* for 10 min. Genomic DNA from sponge bacteria was isolated manually according to a standard SDS lysis protocol with added lysozyme. KS sequences were amplified with the primers KSDPQQF (5′-MGN GAR GCN NWN SMN ATG GAY CCN CAR CAN MG-3′) and KSHGTGR (5′-GGR TCN CCN ARN SWN GTN CCN GTN CCR TG-3′) and Platinum *Taq* DNA polymerase High Fidelity (Invitrogen). PCR products were ligated into the pGEM-T Easy vector (Promega) and digested with *Rsa*I. Plasmids showing a unique restriction pattern were sequenced using the BigDye Terminator Ready Mix (Applied Biosystems) and an ABI 3700 sequencer (Applied Biosystems).

Phylogenetic analysis

Protein sequences were retrieved from GenBank and truncated manually to the KS domains between the EPIAIV and the HGTGT motifs. From PCR-amplified sequences, the regions were removed that corresponded to the primer sequences. After alignment using CLUSTALX, manual correction with BIOEDIT and removal of gap columns, phylogenetic trees were created using the neighbour-joining algorithm of CLUSTALX, the maximum parsimony and maximum likelihood algorithms of the PHYLIP package and the maximum likelihood algorithm with quartet puzzling of TREE-PUZZLE. One thousand bootstraps or 50 000 puzzling steps were performed to assess tree robustness. The type II PKS ActI-ORF1 was used as outgroup. Trees were visualized with the TREEVIEW program.

Accession numbers

These sequence data have been submitted to the GenBank database under accession numbers AY353960–AY353973.

Acknowledgements

The authors would like to thank H. Başpınar and E. Şavk for helping with the collections of Turkish beetle specimens, K. Takada and H. Kobayashi for assistance during sponge collections, and D. Schnabelrauch for technical assistance. We are grateful to W. Boland for support, and to K. Schmid, C. Hertweck and G. Pohnert for critical reading of the manuscript. This work was supported by research grants from the Deutsche Forschungsgemeinschaft (PI 430/1-1 and PI430/2-1) and the JSPS to J.P.

References

Beyer, S., Kunze, B., Silakowski, B., and Müller, R. (1999) Metabolic diversity in myxobacteria: identification of the myxalamid and the stigmatellin biosynthetic gene cluster of *Stigmatella aurantiaca* Sg a15 and a combined polyketide-(poly) peptide gene cluster from the epothilone producing strain *Sorangium cellulosum* So ce90. *Biochim Biophys Acta* **1445:** 185–195.

Brady, S.F., Chao, C.J., Handelsman, J., and Clardy, J. (2001) Cloning and heterologous expression of a natural product biosynthetic gene cluster from eDNA. *Organic Lett* **3:** 1981–1984.

Cane, D.E., Walsh, C.T., and Khosla, C. (1998) Biochemistry – harnessing the biosynthetic code: combinations, permutations, and mutations. *Science* **282:** 63–68.

Cheng, Y.Q., Tang, G.L., and Shen, B. (2003) Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. *Proc Natl Acad Sci USA* **100:** 3149–3154.

El-Sayed, A.K., Hothersall, J., Cooper, S.M., Stephens, E., Simpson, T.J., and Thomas, C.M. (2003) Characterization of the mupirocin biosynthesis gene cluster from *Pseudomonas fluorescens* NCIMB 10586. *Chem Biol* **10:** 419–430.

Gaitatzis, N., Silakowski, B., Kunze, B., Nordsiek, G., Blocker, H., Höfle, G., and Müller, R. (2002) The biosynthesis of the aromatic myxobacterial electron transport inhibitor stigma-

tellin is directed by a novel type of modular polyketide synthase. *J Biol Chem* **277:** 13082–13090.

Gillespie, D.E., Brady, S.F., Bettermann, A.D., Cianciotto, N.P., Liles, M.R., Rondon, M.R., *et al.* (2002) Isolation of antibiotics turbomycin A and B from a metagenomic library of soil microbial DNA. *Appl Environ Microbiol* **68:** 4301–4306.

Haygood, M.G., Schmidt, E.W., Davidson, S.K., and Faulkner, J.D. (1999) Microbial symbionts of marine invertebrates: opportunities for microbial biotechnology. *J Mol Microbiol Biotechnol* **1:** 33–43.

Hentschel, U., Hopke, J., Horn, M., Friedrich, A.B., Wagner, M., Hacker, J., and Moore, B.S. (2002) Molecular evidence for a uniform microbial community in sponges from different oceans. *Appl Env Microbiol* **68:** 4431–4440.

Huang, G.Z., Zhang, L.H., and Birch, R.G. (2001) Multifunctional polyketide-peptide synthetase essential for albicidin biosynthesis in *Xanthomonas albilineans*. *Microbiology* **147:** 631–642.

Kaeberlein, T., Lewis, K., and Epstein, S.S. (2002) Isolating 'uncultivable' microorganisms in pure culture in a simulated natural environment. *Science* **296:** 1127–1129.

Kellner, R.L.L., and Dettner, K. (1996) Differential efficacy of toxic pederin in deterring potential arthropod predators of *Paederus* (Coleoptera: Staphylinidae) offspring. *Oecologia* **107:** 293–300.

Lorenz, P., Liebeton, K., Niehaus, F., and Eck, J. (2002) Screening for novel enzymes for biocatalytic processes: accessing the metagenome as a resource of novel functional sequence space. *Curr Opin Biotechnol* **13:** 572–577.

McCarthy, A.D., and Hardie, D.G. (1984) Fatty acid synthase – an example of protein evolution by gene fusion. *Trends Biochem Sci* **9:** 60–63.

Matsunaga, S., and Fusetani, N. (2003) Nonribosomal peptides from marine sponges. *Curr Organic Chem* **7:** 945–966.

Metsa-Ketela, M., Salo, V., Halo, L., Hautala, A., Hakala, J., Mäntsälä, P., and Ylihonko, K. (1999) An efficient approach for screening minimal PKS genes from *Streptomyces*. *FEMS Microbiol Lett* **180:** 1–6.

Mochizuki, S., Hiratsu, K., Suwa, M., Ishii, T., Sugino, F., Yamada, K., and Kinashi, H. (2003) The large linear plasmid pSLA2-L of *Streptomyces rochei* has an unusually condensed gene organization for secondary metabolism. *Mol Microbiol* **48:** 1501–1510.

Moffitt, M.C., and Neilan, B.A. (2003) Evolutionary affiliations within the superfamily of ketosynthases reflect complex pathway associations. *J Mol Evol* **56:** 446–457.

Narquizian, R., and Kocienski, P.J. (2000) The pederin family of antitumor agents: structures, synthesis and biological activity. In *The Role of Natural Products in Drug Discovery* Mulzer, J., and Bohlmann, R. (eds). Heidelberg, Germany: Springer-Verlag, pp. 25–56.

Paitan, Y., Alon, G., Orr, E., Ron, E.Z., and Rosenberg, E. (1999) The first gene in the biosynthesis of the polyketide antibiotic TA of *Myxococcus xanthus* codes for a unique PKS module coupled to a peptide synthetase. *J Mol Biol* **286:** 465–474.

Piel, J. (2002) A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus* beetles. *Proc Natl Acad Sci USA* **99:** 14002–14007.

Proksch, P., Edrada, R.A., and Ebel, R. (2002) Drugs from the seas – current status and microbiological implications. *Appl Microbiol Biotechnol* **59:** 125–134.

Rawlings, B.J. (2001) Type I polyketide biosynthesis in bacteria (Part B). *Natural Product Rep* **18:** 231–281.

Rondon, M.R., August, P.R., Bettermann, A.D., Brady, S.F., Grossman, T.H., Liles, M.R., *et al.* (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl Environ Microbiol* **66:** 2541–2547.

Scotti, C., Piatti, M., Cuzzoni, A., Perani, P., Tognoni, A., Grandi, G., *et al.* (1993) A *Bacillus subtilis* large Orf coding for a polypeptide highly similar to polyketide synthases. *Gene* **130:** 65–71.

Snel, B., Bork, P., and Huynen, M. (2000) Genome evolution – gene fusion versus gene fission. *Trends Genet* **16:** 9–11.

Vacelet, J. (1971) Electron microscopic study on association between blue-green algae (Chroococcales) and sponge of genus *Verongia*. *J Microsc* **12:** 363–&.

Wang, G.Y.S., Graziani, E., Waters, B., Pan, W.B., Li, X., McDermott, J., *et al.* (2000) Novel natural products from soil DNA libraries in a streptomycete host. *Organic Lett* **2:** 2401–2404.

Wilkinson, B., Foster, G., Rudd, B.A.M., Taylor, N.L., Blackaby, A.P., Sidebottom, P.J., *et al.* (2000) Novel octaketide macrolides related to 6-deoxyerythronolide B provide evidence for iterative operation of the erythromycin polyketide synthase. *Chem Biol* **7:** 111–117.