Construction of hybrid peptide synthetases by module and domain fusions

Henning D. Mootz, Dirk Schwarzer, and Mohamed A. Marahiel*

Philipps-Universität Marburg, Fachbereich Chemie/Biochemie, Hans-Meerwein-Str., 35032 Marburg, Germany

Edited by Christopher T. Walsh, Harvard Medical School, Boston, MA, and approved March 14, 2000 (received for review February 22, 2000)

Nonribosomal peptide synthetases are modular enzymes that assemble peptides of diverse structures and important biological activities. Their modular organization provides a great potential for the rational design of novel compounds by recombination of the biosynthetic genes. Here we describe the extension of a dimodular system to trimodular ones based on whole-module fusion. The recombinant hybrid enzymes were purified to monitor product assembly *in vitro*. We started from the first two modules of tyrocidine synthetase, which catalyze the formation of the dipeptide DPhe-Pro, to construct such hybrid systems. Fusion of the second, proline-specific module with the ninth and tenth modules of the tyrocidine synthetases, specific for ornithine and leucine, respectively, resulted in dimodular hybrid enzymes exhibiting the combined substrate specificities. The thioesterase domain was fused to the terminal module. Upon incubation of these dimodular enzymes with the first tyrocidine module, TycA, incorporating DPhe, the predicted tripeptides DPhe-Pro-Orn and DPhe-Pro-Leu were obtained at rates of 0.15 min^{-1} and 2.1 min^{-1}. The internal thioesterase domain was necessary and sufficient to release the products from the hybrid enzymes and thereby facilitate a catalytic turnover. Our approach of whole-module fusion is based on an improved definition of the fusion sites and overcomes the recently discovered editing function of the intrinsic condensation domains. The stepwise construction of hybrid peptide synthetases from catalytic subunits reinforces the inherent potential for the synthesis of novel, designed peptides.

hybrid enzyme | module exchange

A vast number of short peptides of pharmaceutical and biotechnological importance are synthesized on large enzyme complexes, termed nonribosomal peptide synthetases (NRPSs) (1–4). Prominent drugs like the antibiotics penicillin, vancomycin, actinomycin D, bacitracin, and the immunosuppressant cyclosporin A are examples derived from nonribosomal peptide biosynthesis. The large variety of biological activities within this group of peptides is a result of the enormous structural diversity that can be generated by NRPSs. These enzymes often incorporate unusual residues such as nonproteinogenic amino acids or α-hydroxy acids, which can be further modified by *N*-methylation, heterocyclic ring formation, or epimerization into the D-isomer. Many of these peptides are cyclic or branched-cyclic. Thus, NRPSs hold the key for the synthesis of peptides of complex structure that are chemically difficult to synthesize.

NRPSs exhibit a modular architecture (Fig. 1). In analogy to type I polyketide synthases (PKSs), a module is defined as the catalytic unit that incorporates one specific amino acid into the growing chain (1, 4). Order and specificity of the modules within the protein template determine the sequence of the product. A module can be further subdivided into different domains, each responsible for a certain biochemical reaction (1, 5). The adenylation domain (A-domain) controls the entry of the substrates into the peptide as it recognizes and adenylates its cognate substrate. The activated amino acid subsequently is tethered in thioester linkage at the cofactor 4′-phosphopantetheine of a thiolation-domain (T-domain), also referred to as peptidyl carrier protein (PCP), from where it is condensed with the aminoacyl or peptidyl moieties at the neighboring modules. The latter reaction is facilitated by the condensation domain (C-domain) (Fig. 1). These three domains, C, A, and T, constitute a minimal elongation module, the basic repetitive unit of multimodular NRPSs. Accordingly, the first module of a NRPS complex lacks a C-domain, whereas the last module usually contains in addition a termination domain (Te-domain) to release the product. In positions where a modified amino acid is incorporated, a modification domain is inserted into the corresponding module. This assembly-line arrangement of the catalytic modules and domains suggests straightforward strategies to construct hybrid NRPSs by recombination of gene fragments from different origins with the goal to obtain enzymes with designed novel specificities (1, 2, 4). However, only two such approaches have been reported that alter the amino acid sequence of a nonribosomally synthesized peptide (6, 7). The A- and T-domains of two modules of the surfactin synthetase were exchanged for heterologous A- and T-domains. At the C-terminal module of surfactin synthetase this specificity swap resulted in the predicted variants of the lipoheptapeptide, although at the cost of low product yield (6). However, further attempts to obtain variants at module two by domain swapping were unsuccessful (7). The reason(s) for this setback remained unclear. We suppose that by improper fusion domains were disrupted in their functional integrity and/or ability to interact with neighboring catalytic units. A more precise definition of the domain boundaries, relying on the significant progress achieved on the structure and organization of NRPSs within the last years (8–11), should greatly improve the qualitative and quantitative functionality of hybrid NRPSs. Moreover, a recent study describing the discovery of an editing function of the C-domain suggests that C- and A-domains in artificial junctions might exhibit an incompatibility that will prevent peptide synthesis to proceed (12). This finding demands a new and more general technique to construct novel NRPS, regarding C- and A-domains as an inseparable couple.

Here we describe a strategy to obtain functional hybrid NRPSs based on whole-module fusions. The extension of a dimodular to trimodular NRPS system was undertaken through fusion of entire elongation modules. Furthermore, a Te-domain was linked at the carboxyl-terminal end of the third module to achieve release of the enzyme-bound tripeptides. As an important step toward understanding the readiness of certain domains

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: A-domain, adenylation domain; Abu, α-amino butyric acid; C-domain, condensation domain; NRPS, nonribosomal peptide synthetase; Orn, ornithine; PKS, polyketide synthase; Ppant, 4′-phosphopantetheine; T-domain, thiolation domain; Te-domain, thioesterase domain.

*To whom reprint requests should be addressed. E-mail: marahiel@chemie.uni-marburg.de.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.100075897. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.100075897

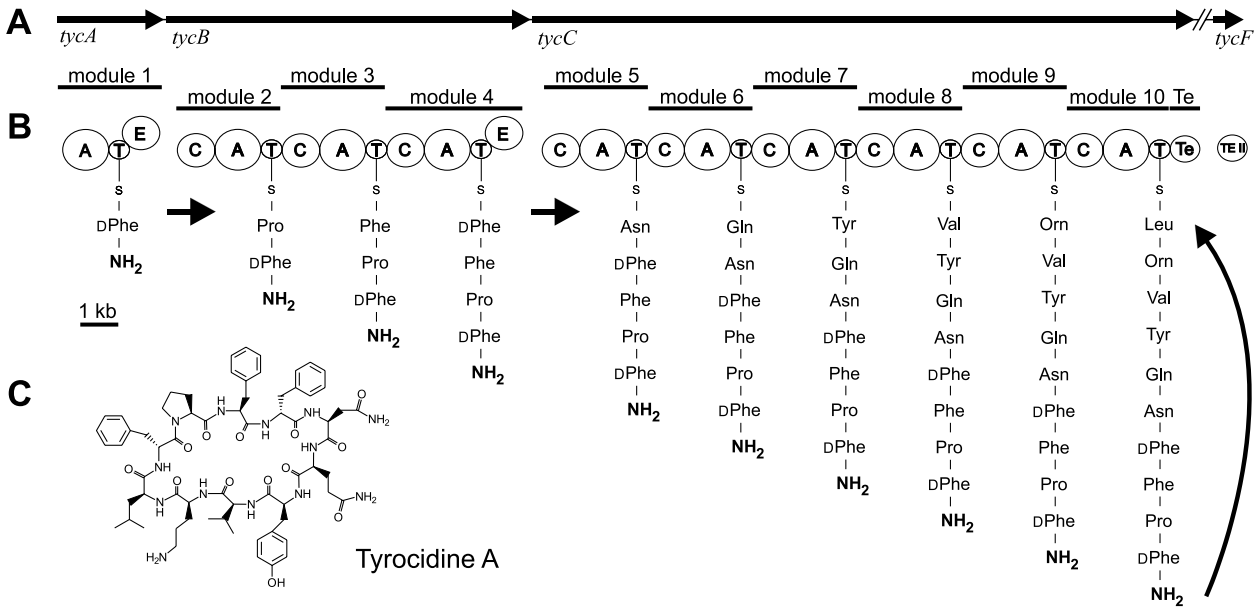

Fig. 1. Tyrocidine biosynthesis. Tyrocidine is a cyclic decapeptide (-DPhe-Pro-Phe-DPhe-Asn-Gln-Tyr-Val-Orn-Leu-)$_{cyc}$ produced by *B. brevis* ATCC 8185 (*C*). The three genes (*tycA*, *tycB*, and *tycC*) (*A*) are transcribed and translated into the NRPSs TycA (124 kDa), TycB (405 kDa), and TycC (724 kDa), which comprise one, three and six modules, respectively, and accomplish the ordered synthesis via the linear decapeptide (*B*).

and modules to accept and process unnatural intermediates, we studied the performance of the hybrid NRPSs *in vitro*.

Materials and Methods

Construction of Expression Plasmids. All recombinant gene fragments were amplified by PCR from chromsomal DNA of *Bacillus brevis* ATCC 8185 by using Vent polymerase (NEB, Schwalbach, Germany). Modified oligonucleotides (MWG Biotech, Ebersberg, Germany) were used to introduce restriction sites for cloning. The gene *tycA* was amplified by using oligonucleotides HM98–12 (5′-ATACCATGGTAGCAAATCAGGCCAATC-3′; restriction site underlined) and HM98–13 (5′-TATGGATC-CGCGCAGTCTATTTGCAAG-3′), digested with *Nco*I and *Bam*HI and ligated into the *Nco*I and *Bam*HI sites of pQE60 (Qiagen, Hilden, Germany) to give pTycA. The plasmid pProCAT was similarly constructed by using oligonucleotides HM98–1 (5′-ATACCATGGGTGTATTTAGCAAAGAA-CAAGTTC-3′) and HM98–2 (5′-TATGGATCCTTCCACAT-ACGCTGCCAG-3′). The gene fragments encoding the modules OrnCAT and LeuCATTe were amplified by using HM98–5 (5′-ATAGGATCCGCATTCGAGCAGTTCGAG-3′) and HM98–6 (5′-ATAGGATCCTTCGATGAACGCCGCCAG-3′), and HM98–7 (5′-ATAGGATCCGCCAAAGGGAAT-GTCTTCTCG-3′) and HM98–8 (5′-ATAGGATCCTTTCAG-GATGAACAGTTCTTG-3′), respectively, digested with *Bam*HI and cloned into pTZ18R linearized with *Bam*HI. After *Bam*HI digestion of the pTZ18R constructs, the fragments were cloned into pProCAT linearized with *Bam*HI, yielding pPro-CAT-OrnCAT and pProCAT-LeuCATTe. The correct orientations were identified by restriction analysis. pProCAT-OrnCAT-Te was obtained through cloning of the gene fragment encoding the Te-domain, which was amplified by using oligonucleotides HM98–9 (5′-ATAAGATCTCATAAGCGCTTT-GAGAGCAG-3′) and HM98–8 and digested with *Bam*HI and *Bgl*II, into the *Bgl*II site of pProCAT-OrnCAT. pProCAT-LeuCAT was obtained by inverse PCR using oligonucleotides HM98–9 and 3′-LeuTeinv (5′-ATAAGATCTCGTGAT-GAAATCGGCCAC-3′) and pProCAT-LeuCATTe as template followed by *Bgl*II digest and religation.

Production of Recombinant Enzymes. Expression plasmids were transformed into *Escherichia coli* BL21-gsp, which carries the plasmid pREP4-gsp providing the *lacI* gene for efficient repression of pQE60 transcription and the gene *gsp* coding for the 4′-phosphopantetheine (Ppant)-transferase Gsp (13). Expression and purification using Ni^{2+}-affinity chromatography was performed as described (13). Pooled enzyme fractions were dialyzed against assay buffer (50 mM Hepes/100 mM NaCl/1 mM EDTA, pH 8.0) with 2 mM dithioerythritol, and protein concentrations were determined by using the calculated extinction coefficients for the adsorbance at 280 nm.

Enzyme Assays and Product Detection. To test the activity of A- and T-domains, the ATP/PP_i exchange reaction and the thioester formation assay were performed as described in assay buffer with 10 mM $MgCl_2$ and 2 mM dithioerythritol (9, 13). Alkaline cleavage of thioesters was basically performed as described (14). [^{32}P]-PP_i was purchased from NEN. Radiolabeled amino acids [^{14}C]-Phe (specific activity 450 mCi/mmol), [^{14}C]-Pro (246 mCi/mmol), [^{14}C]-Leu (292 mCi/mmol), [^{14}C]-Orn (ornithine) (53 mCi/mmol), and [^{14}C]-Lys (308 mCi/mmol) were purchased from Hartmann Analytics, Braunschweig, Germany.

The product formation assays were performed in two ways, by using radiolabeled substrates with subsequent TLC analysis and by using nonradiolabeled substrates with subsequent HPLC/MS analysis. For the radioactive assay a total of 100 μl contained each enzyme, TycA and the respective dimodular hybrid protein, at 500 nM, 10 mM $MgCl_2$, and 2 mM dithioerythritol in assay buffer. Nonradiolabeled amino acids were added at 1 mM and the radiolabeled amino acid at 2.2 μM ([^{14}C]-Phe), 4.1 μM ([^{14}C]-Pro), 3.4 μM ([^{14}C]-Leu), 19 μM ([^{14}C]-Orn) and 3.2 μM ([^{14}C]-Lys). The assay was prewarmed at 37°C and started with the addition of ATP to a concentration of 5 mM. For kinetic studies, the assay was carried out in a larger scale and the concentration of the amino acid carrying the radioactive label was increased by adding the nonradiolabeled form at concentrations of 75 μM Pro (for TycA/ProCAT-LeuCATTe), 15 μM Pro (for TycA/ProCAT-OrnCAT and TycA/ProCAT-OrnCAT-Te), 75 μM Leu, and 15 μM Lys. At various time points 100-μl aliquots were removed and stopped by mixing with 50 μl of

BIOCHEMISTRY

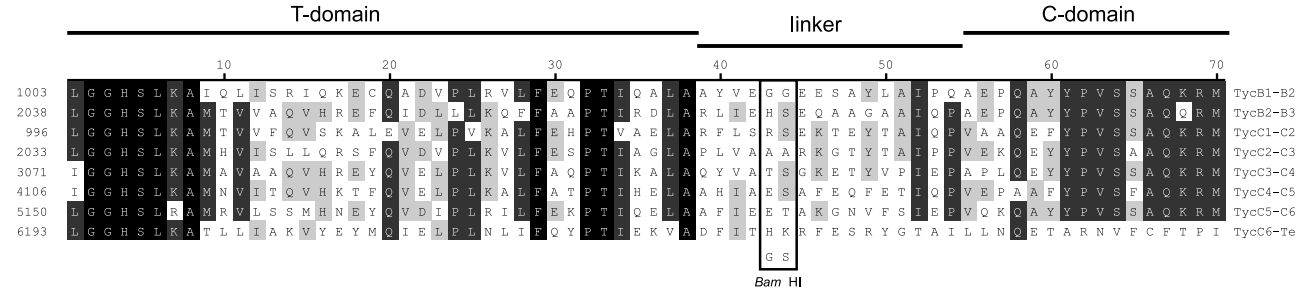

Fig. 2. Linker sequence between T- and C-domains. An alignment of the seven sequences in tyrocidine synthetases TycB and TycC that connect T- and C-domains, as well as the terminal T- and the Te-domains, reveals a stretch with very little conservation. This stretch was defined as linker between the domains. All fusions were performed at the site indicated by introducing a *Bam*HI restriction site in the corresponding gene fragments, which itself codes for the amino acids glycine and serine.

1-butanol/chloroforme (4:1; vol/vol). The mixture was evaporated and the pellet was resuspended in 50 μl methanol, 3 μl of which was applied onto a silica gel 60 TLC plate (Merck), which was developed in H_2O/butanol/acetic acid/ethyl acetate (1:1:1:1; vol/vol/vol/vol). Developed TLC plates were analyzed by autoradiography or radioactivity scanning (counting time 20 min) using radio scanner Ritastar (Raytest, Straubenhardt, Germany) and the supplied RITA software. Product formation was linear for at least 40 min. The nonradiolabeled assays were performed in a 2- to 10-fold scale-up and incubated overnight at 37°C. After identical work-up, the pellet was resolved in 100 μl of 10% buffer B [buffer A: H_2O/0.05% formic acid (vol/vol); buffer B: methanol/0.04% formic acid (vol/vol)], of which 20 μl were applied to HPLC/MS (Hewlett Packard 1100 Series; 250/3-Nucleosil-C18 reversed phase column from Macherey & Nagel). The gradient applied has been described (13). Chemical standards of DPhe-Pro-Orn and DPhe-Pro-Leu were purchased from Applied Biosystems.

Results

Strategy and Construction of Hybrid Peptide Synthetases. Our model systems are based on the tyrocidine synthetases (Fig. 1). The cyclic decapeptide is assembled on a complex of three NRPSs (TycA, TycB, and TycC), which incorporate in sequential order one, three, and six residues, respectively (9). We recently demonstrated that the formation of the first intermediate, the dipeptide DPhe-Pro, can be reconstituted by using the first two modules, GrsA and ProCAT, as purified recombinant enzymes (13). The dipeptide DPhe-Pro, bound as thioester to ProCAT, spontaneously cleaved off the enzyme by formation of the cyclic DPhe-Pro-diketopiperazine. This reaction resulted from the proline-induced higher extent of cis-conformation of the dipeptide that favors intramolecular cyclization. GrsA is the starter module of the gramicidin S synthetase complex, highly homologous to the starter module of the tyrocidine synthetase complex, TycA, which was used in this study.

We defined elongation modules as C-A-T-units comprising C-, A-, and T-domains (2). To link modules in the fashion CAT–CAT, the region between T- and C-domains was examined for sequences indicative of flexible linker loops. By comparison of the T–C junctions within the tyrocidine synthetases, a stretch of about 15 aa was identified 34–49 aa downstream the conserved Ppant-attachment motif LGG(H/D)S(I/L) of the T-domain and upstream of the first signature motifs of the C-domain (1). As shown in Fig. 2, this stretch was highly variable and relatively rich in typical linker amino acids such as Ala, Pro, Gly, Ser, and Glu. All fusions in this study were performed on the genetic level at the position 38 aa downstream of the invariable serine of the T-domain. This junction represents a significant difference to previous studies in which a site 46 aa downstream was used (6, 7). A very recent study on domain boundaries in TycA using partial proteolytic digest proposed a linker region between T- and epimerization domains (E-domains), which is in good agreement with our analysis (10).

To test the idea of whole module fusion, we fused the gene fragment encoding the proline-activating module of TycB, ProCAT, with the gene fragments for the Orn- and Leu-activating modules of the tyrocidine synthetases, as shown in Fig. 3. The gene fragment encoding ProCAT was cloned into the expression vector pQE60 according to the definition of the boundaries between T- and C-domains as mentioned above to give pProCAT. We first inserted into pProCAT in-frame the fragment for the terminal module of TycC, LeuCATTe, which possesses the Te-domain, giving rise after expression to the hybrid enzyme ProCAT-LeuCATTe (264 kDa, hyphen to indicate the fusion site). The same construct without the Te-domain, ProCAT-LeuCAT (235 kDa), was created to investigate the role of the Te-domain in product release. Second, to test the generalizability of this approach, an internal elongation module was chosen for comparison. The gene fragment encoding the second last module of TycC, the Orn-activating OrnCAT, was used to give ProCAT-OrnCAT (237 kDa). In a second fusion step with the Te-domain the corresponding enzyme, ProCAT-OrnCAT-Te (265 kDa), was obtained. The gene *tycA* also was cloned into pQE60 for production of the starter module TycA (124 kDa). *E. coli* BL21 was used as heterologous host for the expression of the recombinant genes. Coexpression of the *gsp* gene coding for the 4′-phosphopantetheine transferase Gsp from a second plasmid ensured the production of the active holo-enzymes (13, 15). Fig. 3 shows the C-terminally His_6-tagged proteins after one-step Ni^{2+}-affinity chromatography as they were used in this study.

Catalytic Activities of the Single Domains. The dimodular enzymes ProCAT-LeuCATTe and ProCAT-LeuCAT hydrolyzed ATP in response to the Pro (91%) and Leu (100%) and bound the two amino acids as thioesters, indicating that the A- and T-domains of the fused modules retained their specific activity. The Pro-module exhibited side specificities in the ATP/PP_i exchange reaction for the imino acid sarcosine (12%, *N*-methyl-glycine), α-amino butyric acid (Abu) (6%), and Ala (4%). Interestingly, the Leu module also activated the D-stereoisomer of its natural substrate to a considerable extent (77%). Ile (13%), norvaline (14%), and Lys (13%) also were recognized to some extent.

The dimodular enzymes ProCAT-OrnCAT and ProCAT-OrnCAT-Te were specific for Pro and Orn. Both amino acids were activated by the A-domains in the ATP/PP_i exchange reaction (93% and 100%) and covalently tethered to the enzyme as thioesters. Orn was bound only substoichiometrically; however, that finding was expected because of the chemical instability of Orn-esters (16) (see product formation). Sarcosine (19%), Abu (8%), and Ala (6%) also were accepted by the Pro module, whereas D-Orn (91%), Lys (73%), and Arg (18%)

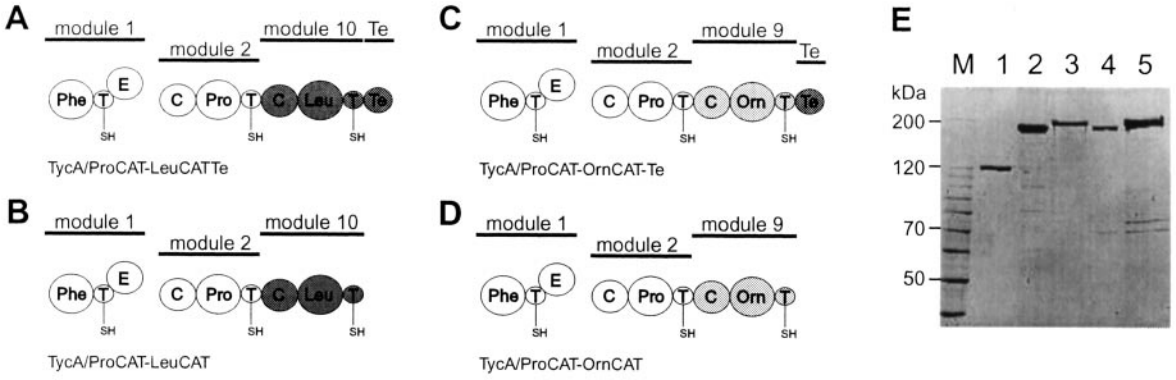

Fig. 3. Construction of artificial trimodular NRPS systems. Artificial trimodular NRPS systems were obtained by extension of the first two modules of tyrocidine synthetases, TycA and TycB1 (here referred to as ProCAT). Fusion of the last module of TycC, containing the Te-domain, to ProCAT yielded TycA/ProCAT-LeuCATTe (*A*). TycA/ProCAT-LeuCAT is devoid of the Te-domain (*B*). Fusion of the second last module of TycC, OrnCAT, resulted in the system TycA/ProCAT-OrnCAT (*D*), which was extended in a second fusion step to TycA/ProCAT-OrnCAT-Te (*C*). The purified His_6-tagged enzymes are shown on a Coomassie blue-stained SDS/PAGE (7.5% polyacrylamide) in *E*. M = marker. Lanes 1–5: TycA, ProCAT-OrnCAT, ProCAT-OrnCAT-Te, ProCAT-LeuCAT, and ProCAT-LeuCATTe, respectively. The gel could not sufficiently resolve the dimodular enzymes, which are 234–265 kDa in size.

represented the most significant side specificities of the Orn module. [^{14}C]-Lys was also efficiently bound in thioester linkage.

The starter module TycA was found to be almost identical in its catalytic properties to those described previously for a similar recombinant construct lacking a His_6-tag (17, 18). L-Phe (97%) and D-Phe (100%) were the preferred substrates of the adenylation reaction. Trp (6%), Met (4%), Tyr (3%), and Leu (2%) gave little activity. [^{14}C]-Phe was bound in thioester linkage.

In summary, the activities of all hybrid enzymes suggested that their integrated A- and T-domains were unaffected in their particular activities and behaved like the sum of the single domains.

Product Formation. We next investigated the artificial tripeptide systems, i.e., the dimodular hybrid enzymes together with the starter module TycA, for their ability to synthesize the predicted tripeptide products.

When the system TycA/ProCAT-LeuCATTe was incubated with the substrates ATP, Phe, Pro, and [^{14}C]-Leu, a radiolabeled product could be identified after TLC analysis, which comigrated with a chemically synthesized standard of the tripeptide DPhe-Pro-Leu. This product was not obtained in any of the control reactions omitting one of the substrates or one of the enzymes (not shown). A product with an identical R_F value could be detected when [^{14}C]-Phe or [^{14}C]-Pro were used instead of [^{14}C]-Leu. Subsequently, the same reaction was performed with nonradiolabeled amino acid substrates, and the products were analyzed by HPLC/MS. Again, a product with identical retention time as the standard DPhe-Pro-Leu was observed only in the presence of all substrates and enzymes. The $[M + H]^+$ mass of this compound was 376 Da (376 Da calculated for DPhe-Pro-Leu, Table 1), further substantiating that the new product was the expected tripeptide.

The same assays were performed with the system TycA/ProCAT-LeuCAT lacking the Te-domain to investigate the role of this domain in product release. No free DPhe-Pro-Leu could be detected, neither in the radioactive nor in the HPLC/MS assay. However, after incubation with all substrates and [^{14}C]-Leu, the tripeptide DPhe-Pro-[^{14}C]-Leu could be chemically cleaved off the enzyme by alkaline treatment and identified by TLC analysis (not shown). From these results we conclude that the Te-domain is catalytically capable and necessary to release the enzyme-bound tripeptide.

The byproduct DPhe-Pro-DKP (diketopiperazine) also was formed by both systems, TycA/ProCAT-LeuCATTe and TycA/ProCAT-LeuCAT, as shown by radioactive labeling with [^{14}C]-Phe or [^{14}C]-Pro and HPLC/MS analysis.

The formation rates of DPhe-Pro-Leu were determined from the time-dependent incorporation of [^{14}C]-Leu or [^{14}C]-Pro into the tripeptide, which was monitored by TLC analysis and radioactivity scanning (Fig. 4). DPhe-Pro-[^{14}C]-Leu was formed at 2.1 min^{-1}, DPhe-[^{14}C]-Pro-Leu at 2.2 min^{-1}, and DPhe-[^{14}C]-Pro-DKP (diketopiperazine) at 0.8 min^{-1}.

In the case of the systems TycA/ProCAT-OrnCAT and TycA/ProCAT-OrnCAT-Te the product pattern was expected to be more complicated. Orn-esters have been reported to be chemically unstable because of intramolecular cyclization, resulting in the formation of the δ-lactam (3-amino-2-piperidone), here referred to as Orn_{cyc} (16, 19). Gramicidin S synthetase produces both Orn_{cyc} and the tetrapeptide DPhe-Pro-Val-Orn_{cyc} in *in vitro* assays (20). Indeed, using the constructs with the Orn module, we observed an alternative route for product release, independent on the presence of the Te-domain.

Incubation of the system TycA/ProCAT-OrnCAT with ATP, Phe, Pro, and [^{14}C]-Orn yielded two radiolabeled products as monitored by TLC analysis and autoradiography. The formation of the first product depended only on the presence of ProCAT-OrnCAT, ATP, and [^{14}C]-Orn and comigrated with a chemically

Table 1. Electrospray ionization-MS analysis of the tripeptide products

Product	Mass calculated $[M + H]^+$	Mass found $[M + H]^+$
	TycA/ProCAT-LeuCATTe	
DPhe-Pro-Leu	376	376
DPhe-Pro-Ile	376	376
DPhe-Pro-NVal	362	362
DPhe-Sar-Leu	350	350
DPhe-Abu-Leu	364	364
DTrp-Pro-Leu	415	415
	TycA/ProCAT-OrnCAT	
DPhe-Pro-Orn_{cyc}	359	359
	TycA/ProCAT-OrnCAT-Te	
DPhe-Pro-Orn_{cyc}	359	359
DPhe-Pro-Orn	377	377
DPhe-Pro-Lys	391	391

BIOCHEMISTRY

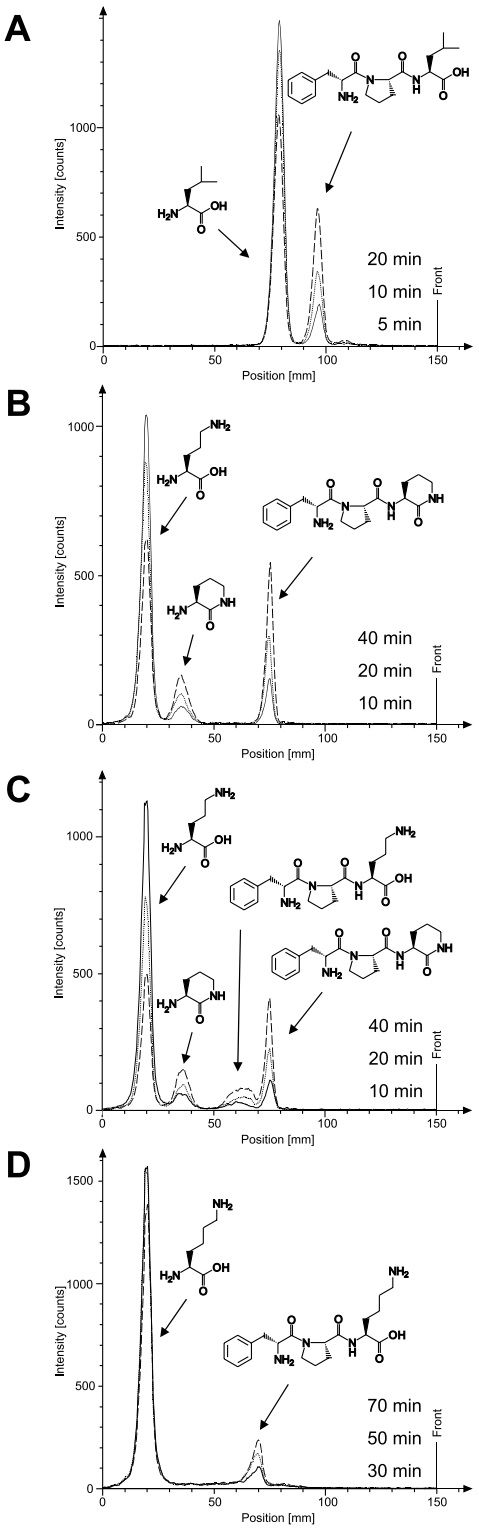

Fig. 4. Formation of the predicted peptides. The product pattern of the trimodular NRPS systems was monitored by using Phe, Pro, and the third amino acid in radiolabeled form, TLC separation of the products and subsequent radioactivity scanning. (*A*) TycA/ProCAT-LeuCATTe with [^{14}C]-Leu. (*B*) TycA/ProCAT-OrnCAT with [^{14}C]-Orn. (*C*) TycA/ProCAT-OrnCAT-Te with [^{14}C]-Orn and (*D*) [^{14}C]-Lys.

synthesized sample of Orn_{cyc}. In contrast, the second product required both enzymes and all substrates to be synthesized. The latter also could be radiolabeled by using [^{14}C]-Phe or [^{14}C]-Pro instead of [^{14}C]-Orn. HPLC/MS analysis confirmed the identity of this product as DPhe-Pro-Orn_{cyc} with a $[M + H]^+$ ion peak of 359 Da (calculated 359, see Table 1). The formation rates were determined at 0.20 min^{-1} for [^{14}C]-Orn_{cyc}, 0.28 min^{-1} for DPhe-Pro-[^{14}C]-Orn_{cyc}, and 0.27 min^{-1} for DPhe-[^{14}C]-Pro-Orn_{cyc} (Fig. 4). No formation of the "free" tripeptide DPhe-Pro-Orn could be detected with the system TycA/ProCAT-OrnCAT lacking the Te-domain.

Product analysis of the system TycA/ProCAT-OrnCAT-Te revealed the effect of the Te-domain fused to the Orn module. An additional product was found, whose formation depended on both enzymes and all substrates. It could be radiolabeled with any of the three amino acids [^{14}C]-Phe, [^{14}C]-Pro, and [^{14}C]-Orn and comigrated with a chemically synthesized standard of DPhe-Pro-Orn. HPLC/MS confirmed the product as the predicted tripeptide DPhe-Pro-Orn with the retention time identical to the standard and a mass of the $[M + H]^+$ ion peak of 377 Da (calculated 377, Table 1). The formation rates were determined at 0.15 min^{-1} for DPhe-Pro-[^{14}C]-Orn and 0.14 min^{-1} for DPhe-[^{14}C]-Pro-Orn. [^{14}C]-Orn_{cyc} was formed at 0.17 min^{-1}, DPhe-Pro-[^{14}C]-Orn_{cyc} at 0.25 min^{-1}, and DPhe-[^{14}C]-Pro-Orn_{cyc} at 0.23 min^{-1} (Fig. 4). We further took advantage of the side specificity of the Orn module for Lys and incubated the system TycA/ProCAT-OrnCAT-Te with ATP, Phe, Pro, and [^{14}C]-Lys. Lys is the higher homolog of Orn with an additional methylene group in the side chain disfavoring the formation of the corresponding ε-lactam Lys_{cyc}. Indeed, DPhe-Pro-[^{14}C]-Lys was the only radiolabeled product detected, although at the lower rate of 0.08 min^{-1} (Fig. 4), and also could be confirmed by HPLC/MS analysis with a $[M + H]^+$ ion peak of 391 Da (calculated 391, Table 1). The system without Te-domain failed to form free DPhe-Pro-Lys in agreement with the hypothesis that the Te-domain is necessary to mediate hydrolytic cleavage (21, 22).

Both systems, TycA/ProCAT-OrnCAT and TycA/ProCAT-OrnCAT-Te, also formed DPhe-[^{14}C]-Pro-DKP (diketopiperazine), the premature cleavage product of the enzyme-bound dipeptide, at rates of 0.74 min^{-1} and 0.39 min^{-1}, respectively.

Substrate Tolerances and Domain Selectivities. To further explore the artificially fused modules and domains for their tolerance in processing different intermediates, we varied the substrate amino acids and qualitatively analyzed the formation of the corresponding variants by HPLC/MS. The system TycA/ProCAT-LeuCATTe readily synthesized the tripeptides DTrp-Pro-Leu, DPhe-Sar-Leu, DPhe-Abu-Leu, DPhe-Pro-Ile, and DPhe-Pro-NVal (norvaline) as judged by the presence of products with the corresponding mass (Table 1). Thus, the C-domain of the Leu module not only accepts at its donor position a DPhe-Pro-S-Ppant intermediate instead of its natural substrate DPhe-Pro-Phe-DPhe-Asn-Gln-Tyr-Val-Orn-S-Ppant, but also the DTrp-Pro-, DPhe-Sar-, and DPhe-Abu-S-Ppant intermediates. On the other hand, when using D-Leu instead of L-Leu as a substrate, only DPhe-Pro-DKP (diketopiperazine) and no DPhe-Pro-DLeu tripeptide was formed. This result further supports the recently discovered selectivity at the acceptor position of the C-domain of NRPSs for the correct stereoisomer (12). D-Leu was bound in thioester linkage but obviously was not accepted at the catalytic center of the preceding C-domain specific for L-Leu. A similar result was obtained with the systems TycA/ProCAT-OrnCAT and TycA/ProCAT-OrnCAT-Te. D-Orn was loaded on the Ppant but did not promote formation of DPhe-Pro-DOrn or DPhe-Pro-$DOrn_{cyc}$. However, DPhe-Pro-S-Ppant was accepted at the donor position instead of the natural substrate DPhe-Pro-Phe-DPhe-Asn-Gln-Tyr-Val-S-Ppant, as was the sterically more demanding Lys at the acceptor position.

The Te-domain was catalytically competent to cleave tripeptides that are not their natural substrates. Not only Leu was accepted as the carboxyl-terminal residue, but also Ile, norval-

ine, and, fused to the Orn module, the hydrophilic Orn and Lys. Similar observations have been published for the Te-domain of the surfactin synthetase (21). It is important to note that the Te-domain in the natural tyrocidine synthetase most probably catalyzes the formation of the cyclic decapeptide, thus acting as a cyclase (9, 23), whereas in the artificial systems described here it acts as a hydrolase. It is conceivable that the tripeptides do not sufficiently fill the catalytic center of the Te-domain, enabling water as alternative nucleophile to intrude. On the other hand, the formation of a cyclic tripeptide might demand too much sterical constraint. It will be interesting to see whether a further stepwise modular extension of these systems will lead to the preferred synthesis of cyclic peptides.

Discussion

We have shown that functional hybrid NRPSs can be generated by whole-module fusion. Modules were defined as a unit comprising C-, A-, and T-domains, representing the minimal catalytic equipment of a repetitive building block in the linear organization of NRPSs (2). The approach of whole-module fusion was chosen for two reasons. First, it probably most closely resembles nature's means to produce structural diversity by developing multimodular biosynthetic genes through gene duplication or deletion. Second, the recently discovered editing function of the C-domain at the acceptor site ("downstream site") suggested to consider C-domains as potential barriers when only the specificity-conferring A-domains are exchanged (12). This problem might cause a drastic reduction of catalytic competence or even a complete failure to synthesize the desired peptide by the engineered NRPS. This can be circumvented when C- and A-domains of one module are maintained as a functional pair. In general, the construction of catalytically efficient hybrid NRPSs is likely to be achievable, because modular NRPSs are invented by nature to produce structural diversity. Such systems should be highly flexible, capable of rearrangements within a relative short period, to randomly evolve the compound that provides the evolutionary advantage against the competing microorganisms. Thus, catalytic units of NRPSs should *per se* be amenable to genetic manipulations. In fact, our artificial systems produced exactly the predicted peptides.

The reported low substrate specificity of the C-domain at the donor site ("upstream" site) also argued for our fusion strategy (12). Both modules tested were capable to process the unnatural dipeptide substrate DPhe-Pro-S-Ppant across the fusion site. Indeed, the formation of the predicted tripeptide DPhe-Pro-Leu on the artificial enzyme template TycA/ProCAT-LeuCATTe took place at the very promising rate of 2.1 min^{-1}, which is in the range of wild-type NRPS systems. The slower formation of DPhe-Pro-Orn at 0.15 min^{-1} also might arise from the second artificial junction of the Orn module with the thioesterase domain and from the likely preference of the Te-domain for a leucyl-peptide. Also the other products of TycA/ProCAT-OrnCAT-Te that result from intramolecular lactamization of the Orn residue must be considered. DPhe-Pro-Orn_{cyc} was formed at 0.25 min^{-1}, increasing the total flux of tripeptides through the Orn module to 0.40 min^{-1}. Further, the formation of Orn_{cyc} at 0.17 min^{-1} probably competes with tripeptide production by consumption of thioester-activated Orn. The turnover rate of wild-type tyrocidine synthetase has not been determined precisely (24), but the formation of the enzyme-bound DPhe-Pro dipeptide was observed at 1.8 min^{-1} by using ProCAT and GrsA from the gramicidin S NRPS (12, 13). The ergotamine NRPS of *Claviceps purpurea* was measured at about 1 min^{-1} (14).

Thus, using the described technique, the prospect of constructing from scratch entire NRPSs templates for a given peptide seems attainable. Because of the C-domain selectivity, this approach also should prove superior to a domain swapping strategy when manipulating existing biosynthetic clusters. The fusion site in the linker downstream of the T-domain also might be convenient for junctions between NRPS and PKS modules. First, the T-domain as peptidyl carrier represents the common feature of NRPS and PKS in terms of biosynthetic mechanism, the latter using the homologous acyl carrier protein (ACP) domain (4). Second, PKS modules usually are defined as a unit comprising the catalytic domains in the order KS-AT-ACP, where KS is ketosynthase and AT is acyl transferase (4, 25), which is functionally equivalent to our definition of C-A-T domains in a module of NRPSs. Indeed, in all yet reported natural examples of mixed NRPS/PKS systems, the organization has been of the kind C-A-T-KS-AT-ACP or vice versa (26–28).

We thank Veit Bergendahl for instructions on the use of HPLC/MS, Torsten Stachelhaus for comments on the manuscript, and Henner Knust for preparing a chemical standard of Orn_{cyc}. H.D.M. received a Ph.D. fellowship from Stiftung Stipendien-Fonds des Verbandes der Chemischen Industrie. Work in M.A.M.'s laboratory was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

1. Marahiel, M. A., Stachelhaus, T. & Mootz, H. D. (1997) *Chem. Rev.* **97,** 2651–2673.
2. Mootz, H. D. & Marahiel, M. A. (1999) *Curr. Opin. Biotechnol.* **10,** 341–348.
3. Kleinkauf, H. & von Döhren, H. (1996) *Eur. J. Biochem.* **236,** 335–351.
4. Cane, D. E., Walsh, C. T. & Khosla, C. (1998) *Science* **282,** 63–68.
5. Konz, D. & Marahiel, M. A. (1999) *Chem. Biol.* **6,** 39–48.
6. Stachelhaus, T., Schneider, A. & Marahiel, M. A. (1995) *Science* **269,** 69–72.
7. Schneider, A., Stachelhaus, T. & Marahiel, M. A. (1998) *Mol. Gen. Genet.* **257,** 308–318.
8. Conti, E., Stachelhaus, T., Marahiel, M. A. & Brick, P. (1997) *EMBO J.* **16,** 4174–4183.
9. Mootz, H. D. & Marahiel, M. A. (1997) *J. Bacteriol.* **179,** 6843–6850.
10. Dieckmann, R., Pavela-Vrancic, M., von Döhren, H. & Kleinkauf, H. (1999) *J. Mol. Biol.* **288,** 129–140.
11. Suo, Z., Walsh, C. T. & Miller, D. A. (1999) *Biochemistry* **38,** 14023–14035.
12. Belshaw, P. J., Walsh, C. T. & Stachelhaus, T. (1999) *Science* **284,** 486–489.
13. Stachelhaus, T., Mootz, H. D., Bergendahl, V. & Marahiel, M. A. (1998) *J. Biol. Chem.* **273,** 22773–22781.
14. Walzel, B., Riederer, B. & Keller, U. (1997) *Chem. Biol.* **4,** 223–230.
15. Lambalot, R. H., Gehring, A. M., Flugel, R. S., Zuber, P., LaCelle, M., Marahiel, M. A., Reid, R., Khosla, C. & Walsh, C. T. (1996) *Chem. Biol.* **3,** 923–936.
16. Gadow, A., Vater, J., Schlumbohm, W., Palacz, Z., Salnikow, J. & Kleinkauf, H. (1983) *Eur. J. Biochem.* **132,** 229–234.
17. Gocht, M. & Marahiel, M. A. (1994) *J. Bacteriol.* **176,** 2654–2662.
18. Pfeifer, E., Pavela-Vrancic, M., von Döhren, H. & Kleinkauf, H. (1995) *Biochemistry* **34,** 7450–7459.
19. Weber, A. L. & Miller, S. L. (1981) *J. Mol. Evol.* **17,** 273–284.
20. Vater, J., Schlumbohm, W., Palacz, Z., Salnikow, J., Gadow, A. & Kleinkauf, H. (1987) *Eur. J. Biochem.* **163,** 297–302.
21. de Ferra, F., Rodriguez, F., Tortora, O., Tosi, C. & Grandi, G. (1997) *J. Biol. Chem.* **272,** 25304–25309.
22. Schneider, A. & Marahiel, M. A. (1998) *Arch. Microbiol.* **169,** 404–410.
23. Shaw-Reid, C. A., Kelleher, N. L., Losey, H. C., Gehring, A. M., Berg, C. & Walsh, C. T. (1999) *Chem. Biol.* **6,** 385–400.
24. Fujikawa, K., Sakamoto, Y., Suzuki, T. & Kurahashi, K. (1968) *Biochim. Biophys. Acta* **169,** 520–533.
25. Gokhale, R. S., Tsuji, S. Y., Cane, D. E. & Khosla, C. (1999) *Science* **284,** 482–485.
26. Albertini, A. M., Caramori, T., Scoffone, F., Scotti, C. & Galizzi, A. (1995) *Microbiology* **141,** 299–309.
27. Molnar, I., Schupp, T., Ono, M., Zirkle, R., Milnamow, M., Nowak-Thompson, B., Engel, N., Toupet, C., Stratmann, A., Cyr, D. D., *et al*. (2000) *Chem. Biol.* **7,** 97–109.
28. Silakowski, B., Schairer, H. U., Ehret, H., Kunze, B., Weinig, S., Nordsiek, G., Brandt, P., Blocker, H., Hofle, G., Beyer, S., *et al*. (1999) *J. Biol. Chem.* **274,** 37391–37399.