Systematic and Quantitative Analysis of Protein−Protein Recognition between Nonribosomal Peptide Synthetases Investigated in the Tyrocidine Biosynthetic Template†

Uwe Linne, Daniel B. Stein, Henning D. Mootz,‡ and Mohamed A. Marahiel*

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

Received February 7, 2003; Revised Manuscript Received March 12, 2003

ABSTRACT: We present a systematic and quantitative study of the protein−protein recognition between the three tyrocidine synthetases TycA, TycB, and TycC investigated with two artificial *in trans* assay systems, which had been previously developed: the "DKP assay system" for the interaction of TycA with TycB and the "L/D-Phe-L-Asn assay system" for the interaction of TycB with TycC. TycA-A_{Phe}TE and $TycB_3$-A_{Phe}TE, which are used as donor enzymes, both provide D-Phe-S-Ppant, so that no substrate specificities interfered with the quantification of protein−protein recognition. We tested all donor/acceptor enzyme combinations between the two artificial assay systems for product formation activities as well as two hybrid enzymes, where the E-domains of TycA and $TycB_3$ had been exchanged against each other. Furthermore, four donor/acceptor protein fusions were constructed on gene level, resulting in dimodular proteins. We were able to show that the E-domains mediate protein−protein recognition *in trans*. Product formation of the different donor assayed with the two acceptor enzymes $TycB_1$-CA_{Pro}T and $TycC_1$-CA_{Asn}T/Te *in trans* was only obtained if the donor enzyme harbored the cognate E-domain. Interestingly, all *in cis* fusions (dimodular proteins) were active, giving strong evidence that unnatural protein−protein interactions can be "forced" by fusion of the distinct enzymes. Finally, we were able to detect product formation in the "DKP system" with engineered hybrid proteins where the A-domain of TycA had been exchanged against the isoleucine-activating A-domain of $BacA_1$ and the valine-activating A-domain of $TycC_4$, respectively. All of these findings are of high relevance for future NRPS engineering approaches.

Microorganisms produce many pharmacologically active secondary metabolites, among them antibiotics (like penicillin, erythromycin, vancomycin, and tyrocidine), immunosuppresiva (e.g., cyclosporin), and cytostatica (e.g., epothilone). Some of these interesting compounds belong to the groups of polypeptides and polyketides that reveal unusual structures. They are biosynthesized by so-called nonribosomal peptide synthetases (NRPSs)[1] and polyketide synthases (PKSs), respectively (*1*). These enzymes, irrespective of origin, are composed of modules, whereby each module is responsible for the recognition, activation, and incorporation of one specific building block into the growing product chain. Such modules can be further subdivided into domains, each harboring one special catalytic function.

In NRPSs, the adenylation (A) domains are responsible for the substrate recognition and activation as aminoacyl-*O*-AMP under ATP consumption (*2*). Subsequently, the activated amino acid is transferred and bound to a 4′-phosphopantetheine (4′-Ppant) moiety as a thioester under release of AMP. The 4′-Ppant cofactor itself is covalently tethered to an invariant serine residue of a peptidyl carrier protein (PCP) (*3*). The modification of inactive apo-PCPs to functional holo-PCPs is carried out by 4′-phosphopantetheine transferases (*4*, *5*) under consumption of CoA. After activation and covalent binding of the substrates to the enzymes, the condensation (C) domains catalyze the nucleophilic attack of the α-amino group of the new building block onto the thio-esterified carboxy group of the aminoacyl- or peptidyl-S-Ppant moiety bound to the PCP of the preceding module (*6*). By the action of these essential domains (A-domain, PCP, and C-domain) the growing peptide chain is handed over from one module to the next until the termination module is reached. There, a termination domain, in bacterial systems most often a thioesterase (Te) domain, is responsible for the release of the product from the enzyme template (*7*). Besides these essential domains

† This work has been supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

* To whom correspondence should be addressed. Tel: ++49-6421-2825722. Fax: ++49-6421-2822191. E-mail: marahiel@chemie.uni-marburg.de.

‡ Present address: Laboratory of Synthetic Protein Chemistry, The Rockefeller University, 1230 York Ave., New York, NY 10021.

[1] Abbreviations: A, adenylation domain; aminoacyl- or peptidyl-S-Ppant, aminoacylated thioester form of cofactor Ppant bound to the strictly conserved serine residue of PCPs; Abu, α-aminobutyric acid; ATP, adenosine 5′-triphosphate; C, condensation domain; CoASH, coenzyme A; DKP, diketopiperazine; DTE, 1,4-dithioerythritol; E, epimerization domain; EDTA, ethylenediaminetetraacetic acid; HPLC, high-performance liquid chromatography; LSC, liquid scintillation counting; NRPS, nonribosomal peptide synthetase; PCR, polymerase chain reaction; PCP, peptidyl carrier protein, referring to the same thing as "T" but used in the text to stress the function of the protein (in analogy to the well-known ACPs of fatty acid synthases and polyketide synthases); PCP^C, PCP normally localized in front of a C-domain; PCP^E, PCP naturally connected to an E-domain; Ppant, 4′-phosphopantetheine; PP_i, inorganic pyrophosphate; SIM, single ion mode; T, thiolation domain, referring to the same thing as "PCP" but used for the description of proteins ("one letter−one domain" nomenclature of NRPSs); TCA, trichloroacetic acid; TLC, thin-layer chromatography.

10.1021/bi034223o CCC: $25.00
Published on Web 04/09/2003

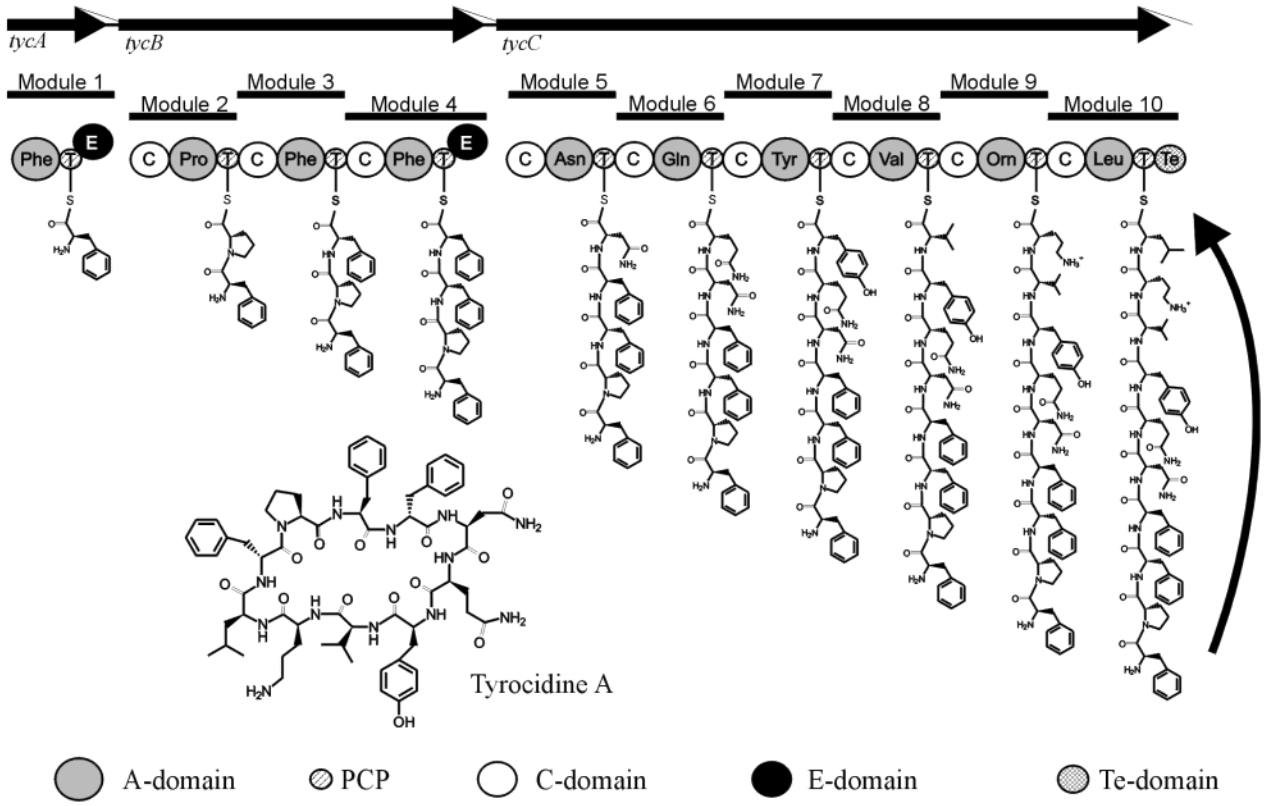


FIGURE 1: Tyrocidine synthetases. The three genes *tycA*, *tycB*, and *tycC* are encoding the tyrocidine synthetases TycA (one module), TycB (three modules), and TycC (six modules). At positions 1 and 4 a D-amino acid is built into the product, and at the C-terminal end of TycC, a Te-domain is localized, generating product turnover by intramolecular cyclization and release of tyrocidine A.

the modules can contain different optional domains. Among them are epimerization (E) (*8*, *9*), heterocyclization (Cy) (*10*, *11*), and N-methylation (M) (*12*) domains, significantly contributing to the chemical diversity and therefore to the bioactivity of the corresponding secondary metabolites.

Whereas in all known fungal NRPS systems the modules are organized on one single large enzyme, for example, cyclosprin synthetase with a mass of 1.4 MDa is the largest single polypeptide chain currently known (*13*), in bacteria there are often several synthetases interacting with each other. Very often, at the C-terminal end of such synthetases interacting with others of the same system, E-domains are localized. In the tyrocidine biosynthesis system, there are three synthetases, TycA (one module), TycB (three modules), and TycC (six modules) (compare Figure 1) (*14*), and at the C-terminal ends of TycA and TycB there are E-domains. So the question arises, by which mechanisms is the formation of only one product with a distinct length and order secured?

To understand this ordered protein interaction, we here present a systematic and quantitative study of the protein−protein recognition in the tyrocidine biosynthesis system. We tested all donor/acceptor enzyme combinations between the two described artificial systems for product formation activity as well as two hybrid enzymes, where the E-domains of TycA and $TycB_3$ had been exchanged against each other. Furthermore, four donor/acceptor protein fusions were constructed on the gene level, resulting in dimodular proteins. We were able to show that the E-domains mediate protein−protein interaction *in trans*. Interestingly, all *in cis* fusions (dimodular proteins) were active, giving strong evidence that unnatural protein−protein interactions can be "forced" by fusion of the distinct enzymes. Finally, we were able to detect product formation in the DKP system with engineered hybrid proteins where the A-domain of TycA had been exchanged against the isoleucine-activating A-domain of $BacA_1$ and the valine-activating A-domain of $TycC_4$, respectively (*9*).

EXPERIMENTAL PROCEDURES

Cloning of the Recombinant Peptide Synthetase Genes and Protein Overproduction. All plasmids used are based on the pQE vector system from Qiagen (Hilden, Germany). Standard procedures were applied for PCR amplification, purification, and cloning of recombinant DNA (*15*). For cloning, *Escherichia coli* XL1 Blue (Stratagene, Heidelberg, Germany) was used.

The cloning of the recombinant peptide synthetase fragments encoding the proteins TycA-A_{Phe}TE (*16*), $TycB_1$-CA_{Pro}T (*16*), TycA-A_{Phe}T/E^{TycB3} (*9*), $TycB_3$-A_{Phe}TE (*17*), $TycB_3$-A_{Phe}T/E^{TycA} (*9*), $TycC_1$-CA_{Asn}T/Te (*17*), $TycC_4$-A_{Val}/TE^{TycA} (*9*), and $BacA_1$-A_{Ile}/TE^{TycA} (*9*) was described elsewhere.

The plasmids encoding the dimodular fusion proteins originate from plasmids pQE60-tycA-ATE (*16*), pQE70-$tycB_3$-ATE (*17*), pMS-$tycB_1$-CAT (cloning; see below), and pQE60-$tycC_1$-CAT/Te (*17*). Plasmid pMS is a derivative of the pQE60 (Qiagen). Phosphorylated primers 5′-GATCGGGTACCGCAGGAACTAGTC-3′ and 5′-GATCGACTAGTTCCTGCGGTACCC-3′ were annealed with each other and subsequently ligated with the *Bgl*II linearized pQE60 plasmid resulting in plasmid pMS. Gene fragment $tycB_1$-CAT was PCR-amplified using chromosomal DNA of *Bacillus brevis* ATCC 8185 with the primers 5′-TATCCATGGGAAGATCTAGTGTATTTAGCAAAGAACA-3′ and 5′-TATG-

GATCCTTCCACATACGCTGCCAG-3′. Plasmid pMS and PCR fragment tycB$_1$-CAT were digested with *Nco*I and *Bam*HI and ligated with each other, resulting in pMS-tycB$_1$-CAT. Plasmid pMS-tycB$_1$-CAT was then digested with *Eco*RI and *Bgl*II and ligated with the tycA-ATE-encoding fragment obtained from pQE60-tycA-ATE (*16*) using the same restriction enzymes to give plasmid pMS-tycA-ATE/tycB$_1$-CAT. For generation of plasmid pQE-tycB$_3$-ATE/tycB$_1$-CAT, pQE-tycB$_3$-ATE was linearized with *Bam*HI and dephosphorylated. Plasmid pMS-tycB$_1$-CAT was digested with *Bgl*II and *Bam*HI. Subsequently, the purified fragments containing tycB$_3$-ATE and tycB$_1$-CAT were ligated. Plasmid pQE60-$^{Bam\text{HI}}$tycC$_1$-CAT/Te was generated from plasmid pQE60-tycC$_1$-CAT/Te by utilizing the QuickChange site-directed mutagenesis kit (Stratagene) as recommended by the manufacturer's protocols with the primers 5′-GAGGAGAAATTAGGATCCGAAAAGCAGGAAAAC-ATCGCAAAAATTTACCCGC-3′ and 5′-GCGGGTAAA-TTTTTGCGATTGTTTCCTGCTTTTCGGATCCTAATT-TCTCCTC-3′ (exchanged bases are underlined). Subsequently, the new plasmid was digested with *Bam*HI and *Xba*I. With the same restriction enzymes, fragments of plasmids pQE60-tycA-ATE and pQE70-tycB$_3$-ATE were obtained. The purified digestion products containing the gene fragments of tycA-ATE and tycC$_1$-CAT/Te as well as tycB$_3$-ATE and tycC$_1$-CAT/Te were ligated with each other, resulting in the two fusion plasmids pQE-tycA-ATE/tycC$_1$-CAT/Te and pQE-tycB$_3$/tycC$_1$-CAT/Te, respectively.

E. coli M15/pREP4 was transformed with all plasmids described. Expression and purification of the His$_6$-tagged apoproteins were performed as previously described (*6*). Overproduction and purification after single-step Ni^{2+} affinity chromatography were confirmed by SDS−PAGE (*18*). The protein concentrations were assigned using the calculated extinction coefficients for the A_{280} of the enzymes (TycA-A$_{Phe}$TE, 136580 M^{-1} cm^{-1}; TycA-A$_{Phe}$T/E^{TycB3}, 132620 M^{-1} cm^{-1}; TycB$_1$-CA$_{Pro}$T, 92230 M^{-1} cm^{-1}; TycB$_3$-A$_{Phe}$TE, 135870 M^{-1} cm^{-1}; TycB$_3$-A$_{Phe}$T/E^{TycA}, 139830 M^{-1} cm^{-1}; TycC$_1$-CA$_{Asn}$T/Te, 150800 M^{-1} cm^{-1}; TycA-A$_{Phe}$TE/TycB$_1$-CA$_{Pro}$T, 228810 M^{-1} cm^{-1}; TycB$_3$-A$_{Phe}$TE/TycB$_1$-CA$_{Pro}$T, 228100 M^{-1} cm^{-1}; TycA-A$_{Phe}$TE/TycC$_1$-CA$_{Asn}$T/Te, 287380 M^{-1} cm^{-1}; TycB$_3$-A$_{Phe}$TE/TycC$_1$-CA$_{Asn}$T/Te, 286670 M^{-1} cm^{-1}; TycC$_4$-A$_{Val}$/TE^{TycA}, 108820 M^{-1} cm^{-1}; BacA$_1$-A$_{Ile}$/TE^{TycA}, 121480 M^{-1} cm^{-1}; also see Method for the Calculation of Active Enzyme Concentration in Experimental Procedures).

The dimodular fusion proteins TycA-A$_{Phe}$TE/TycB$_1$-CA$_{Pro}$T, TycB$_3$-A$_{Phe}$TE/TycB$_1$-CA$_{Pro}$T, TycA-A$_{Phe}$TE/TycC$_1$-CA$_{Asn}$T/Te, and TycB$_3$-A$_{Phe}$TE/TycC$_1$-CA$_{Asn}$T/Te were further purified by ion-exchange chromatography on an ÄKTA purifier FPLC system (Amersham Biosciences, Freiburg, Germany). Dialysis against ion-exchange buffer A (see below) was carried out utilizing HiTrap desalting columns (Amersham Biosciences). For subsequent ion-exchange chromatography an Source 15Q HR 16/5 anion-exchange column (Amersham Biosciences) was used with the following gradient at a flow rate of 1 mL/min [buffer A, 50 mM HEPES (pH 8.0), 100 mM NaCl, 1 mM EDTA, 2 mM DTE, 10 mM $MgCl_2$; buffer B, 50 mM HEPES (pH 8.0), 1000 mM NaCl, 1 mM EDTA, 2 mM DTE, 10 mM $MgCl_2$]: loading 100% bufferA, followed by washing with 5 column volumes of 100% buffer A and a linear gradient up to 100% buffer B in 60 min. The proteins eluting were identified by UV detection at 280 nm and SDS−PAGE analysis of the fractions.

After dialysis against assay buffer [50 mM HEPES (pH 7.6), 100 mM NaCl, 1 mM EDTA, 2 mM DTE, 10 mM $MgCl_2$] all of the purified proteins were shock-frozen in liquid nitrogen and could be stored at −80 °C over several weeks without significant loss of activity.

Posttranslational Modification of the Enzymes by Sfp and CoASH. Priming of heterologously expressed proteins was carried out in vitro by addition of 200 μM CoASH and 15 nM recombinant *Bacillus subtilis* Ppant-transferase Sfp (*5*) to all assays that required holoenzymes. After preincubation of the samples for 10 min at 37 °C for priming, the assays were started by addition of the amino acids.

ATP−PP$_i$ Exchange Reaction. Adenylation domain activity, specificity, and correct folding of all proteins utilized were examined by applying the ATP−PP$_i$ exchange assay (*6*). We tested the cognate substrate amino acids relevant for the subsequent assays.

Aminoacylation Assay. To a reaction mixture containing 500 nM enzyme, 50 mM HEPES (pH 7.6), 200 mM NaCl, 1 mM EDTA, 10 mM $MgCl_2$, and 1 mM ATP, equilibrated (10 min) at 37 °C, ^{14}C-labeled substrate amino acids were rapidly added to a final concentration of 2.5 μM. At various time points, 100 μL aliquots were taken and immediately quenched by addition of 1 mL of ice-cold TCA (10%). After 15 min of incubation on ice, samples were centrifuged (4 °C, 13000 rpm) for 20 min, washed twice with 1 mL of ice-cold TCA, redissolved in 150 μL of formic acid, and quantified by LSC.

Method for the Calculation of Active Enzyme Concentration. A common margin of error is made when determining a protein concentration by the calculated extinction coefficient or by the method of Bradford. According to our experience, aminoacylation activities may vary significantly (up to 3-fold) depending on the batch of protein utilized in the assays. However, expression and purification conditions were the same, and expression rates are not varying significantly from one preparation to another one (as estimated by SDS−PAGE; data not shown). So the reasons for the observed differences in enzyme activities remain obscure at the moment. However, in the present work we tried to take this behavior into consideration for all product formation assays by calculating the enzyme concentration on the basis of the aminoacylation activity as an internal standard for the amount of active protein as described previously (*19*).

L/D-Phe-L-Asn Formation Assay. For the detection of peptide products, TycA-A$_{Phe}$TE, TycA-A$_{Phe}$T/E^{TycB3}, TycB$_3$-A$_{Phe}$TE or TycB$_3$-A$_{Phe}$T/E^{TycA}, and TycC$_1$-CA$_{Asn}$T/Te (100 nM active protein each) were assayed at 37 °C with 40 μM L- or D-Phe and 40 μM L-Asn for up to 6 h in a final volume of 500 μL. In the case of the dimodular fusion proteins, 100 nM TycA-A$_{Phe}$TE/TycC$_1$-CA$_{Asn}$T/Te or TycB$_3$-A$_{Phe}$TE/TycC$_1$-CA$_{Asn}$T/Te was utilized under the same reaction conditions. The reaction was stopped by the addition of a 3-fold excess of methanol. After centrifugation (13000 rpm, 15 min) and transferring the supernatant to fresh tubes, the solvent was removed under vacuum, and the residue was dissolved in 100 μL of 10% methanol (v/v) and applied (90 μL) to HPLC-MS. Separation of the reaction products was achieved on a

250/3 Nucleosil C18 reversed-phase column (Machery and Nagel) by applying the following gradient at a flow rate of 0.3 mL min^{-1} [buffer A, 0.05% formic acid/water (v/v); buffer B, 0.045% formic acid/methanol (v/v)]: loading 10% buffer B, linear gradient up to 60% buffer B in 25 min, followed by a linear gradient up to 100% buffer B in 5 min, and then holding 100% buffer B for 5 min. The products were identified by comparison with standards (UV detection at 215 nm) and by on-line ESI-MS analysis with an Agilent 1100 MSD.

DKP Formation Assay. For comparison of DKP production rates (*6*), TycA-A_{Phe}TE, TycA-A_{Phe}T/E^{TycB3}, $TycB_3$-A_{Phe}TE or $TycB_3$-A_{Phe}T/E^{TycA}, and $TycB_1$-CA_{Pro}T (50 nM active protein each) were assayed at 37 °C with 100 μM L- or D-Phe and 100 μM L-Pro for up to 3 h in a final volume of 200 μL. In the case of the dimodular fusion proteins, 50 nM TycA-A_{Phe}TE/$TycB_1$-CA_{Pro}T or $TycB_3$-A_{Phe}TE/$TycB_1$-CA_{Pro}T was utilized under the same reaction conditions. $TycC_4$-A_{Val}/TE^{TycA} and $BacA_1$-A_{Ile}/TE^{TycA} were assayed together with 250 nM $TycB_1$-CA_{Pro}T under the same conditions except that 100 μM L-Val and 100 μM L-Ile were used as donor substrates, respectively. The reaction was stopped by the addition of 1 mL of methanol. After centrifugation (13000 rpm, 15 min) and transferring the supernatant to a fresh tube, the solvent was removed under vacuum, and the residue was dissolved in 100 μL of 30% methanol (v/v) and applied (50 μL) to HPLC. Separation of the reaction products was achieved on a 70/4 Nucleosil 100-5 C18 reversed-phase column (Machery-Nagel) by applying an isocratic method at a flow rate of 0.6 mL min^{-1} with 30% buffer B [buffer A, 0.05% formic acid/water (v/v); buffer B, 0.045% formic acid/methanol (v/v)]. The products were identified by comparison with standards (UV detection at 215 nm) and by on-line ESI-MS analysis with an Agilent 1100 MSD.

In case of the $TycC_4$-A_{Val}/TE^{TycA} and $BacA_1$-A_{Ile}/TE^{TycA} assays, separation of reaction products was achieved on a 250/3 Nucleodur 100-3 C18ec reversed-phase column (Machery and Nagel) by applying the following gradient at a flow rate of 0.3 mL min^{-1} and a column temperature of 40 °C [buffer A, 0.05% formic acid/water (v/v); buffer B, 0.045% formic acid/methanol (v/v)]: loading 10% buffer B, linear gradient up to 60% buffer B in 55 min, followed by a linear gradient up to 95% buffer B in 5 min, and then holding 100% buffer B for 10 min. On-line ESI-SIM (single ion mode)-MS analysis (Agilent 1100 MSD) was used for highly sensitive product detection.

Alternatively, a 125/2 Nucleodur C18 gravity (3 μm) reversed-phase column (Machery and Nagel) was used with the following gradient at a flow rate of 0.2 mL min^{-1} and a column temperature of 40 °C [buffer A, 0.05% formic acid/water (v/v); buffer B, 0.045% formic acid/methanol (v/v)]: loading 5% buffer B, linear gradient up to 61.3% buffer B in 52 min, followed by a linear gradient up to 95% buffer B in 1 min, and then holding 100% buffer B for 2 min. On-line MS analysis was performed by ESI-TOF measurement with a Qstar mass spectrometer (Applied Biosystems) in positive ion mode.

Radio Assay for the Detection of Elongation and Product Release. $TycC_4$-A_{Val}/TE^{TycA} or $BacA_1$-A_{Ile}/TE^{TycA} (500 nM) was preincubated at 37 °C with their substrate amino acids [2 μM [^{14}C]-L-Val or L-Ile (450 mCi/mmol), respectively]. $TycB_1$-CA_{Pro}T was separately preincubated with 1 mM L-Abu for 30 min at 37 °C. Product formation was initiated by mixing equal volumes of preincubated reaction mixtures of the hybrid proteins and $TycB_1$-CA_{Pro}T, respectively. At various time points, 100 μL aliquots were taken and immediately quenched by addition of 1 mL of ice-cold TCA (10%). After 15 min of incubation on ice, samples were centrifuged (4 °C, 13000 rpm) for 20 min, washed two times with 1 mL of ice-cold TCA, redissolved in 150 μL of formic acid, and quantified by LSC.

RESULTS

Nomenclature of the Recombinant Enzymes and Artificial Assay Systems Utilized. In this study we used several recombinant enzymes. For demonstrating the origin and the domain composition of each domain in an easy matter, we utilize a special nomenclature. For example, the hybrid enzyme $TycB_3$-A_{Phe}T/E^{TycA} originates from the third module of tyrocidine synthetase TycB (compare Figure 1) and consists of an A-, a T-, and an E-domain; thereby the covalently fused E-domain is the E-domain of TycA (the protein fusion is indicated by a "/"), and the A-domain specificity is mainly phenylalanine.

In the tyrocidine biosynthesis system product−intermediate transfer between TycA/TycB and TycB/TycC was previously studied in artificial systems. The well-studied "DKP formation system" consists of the one modular synthetase TycA-A_{Phe}TE and the first module of TycB ($TycB_1$-CA_{Pro}T). D-Phe is transferred from TycA to $TycB_1$, resulting in the formation of linear D-Phe-L-Pro bound as a thioester to the 4′-Ppant moiety of the enzyme. Cyclic D-Phe-L-Pro-DKP is subsequently released by the nucleophilic attack of the N-terminal amino group of D-Phe on the carboxy group of the D-Phe-L-Pro-S-Ppant thioester, resulting in the regeneration of the enzymes for a new round of catalysis (see Figure 2a) (*6*). Furthermore, an artificial system had also been developed and characterized for the second transition (TycB/TycC). The terminal module of TycB ($TycB_3$-A_{Phe}TE) transfers Phe to the first module of TycC ($TycC_1$-CA_{Asn}T/Te), which harbors the Te-domain of $TycC_6$ directly fused to its C-terminus ensuring product turnover and release of linear L/D-Phe-L-Asn (compare Figure 2b) (*17*).

The E-domain of $TycB_3$ naturally epimerizes a tetrapeptide, but it has been demonstrated that aminoacyl-S-Ppant is also accepted as a substrate, although the epimerization efficiency is strongly decreased compared to peptidyl-S-Ppant substrates (*17*). Therefore, when L-Phe is used as substrate, the reaction is slow and a 2:1 mixture of L- and D-Phe-L-Asn is formed, while with D-Phe exclusively D-Phe-L-Asn is observed at much higher product formation rates (*17*, *20*). Therefore, the C-domain of $TycC_1$ is capable of accepting and elongating L- or D-Phe-S-Ppant instead of its natural substrate D-Phe-L-Pro-L-Phe-D-Phe-S-Ppant. D-Phe is also transferred from TycA to TycB. However, up to now, in vivo no shortened peptides resulting from a "misinteraction" between TycA and TycC have been observed, giving rise for a correct and selectively ordered recognition of the different synthetases. Moreover, when deleting the E-domain of TycA in the DKP system, no product is formed, neither with L-Phe nor with D-Phe as substrate (*6*), but when only the function of the E-domain was knocked out by a point mutation, D-Phe-L-Pro-DKP formation had been observed exclusively with D-Phe as substrate (*21*).

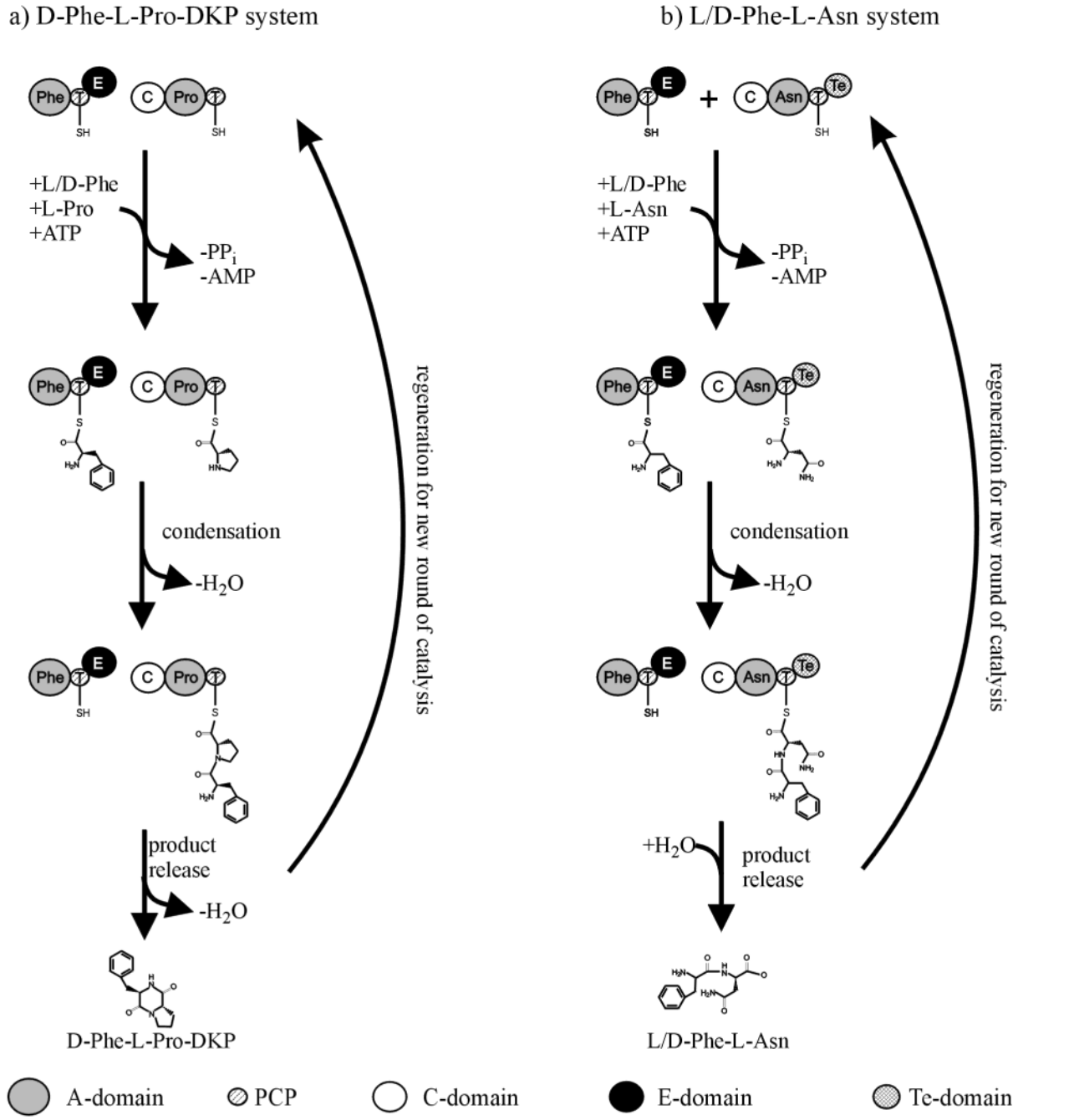
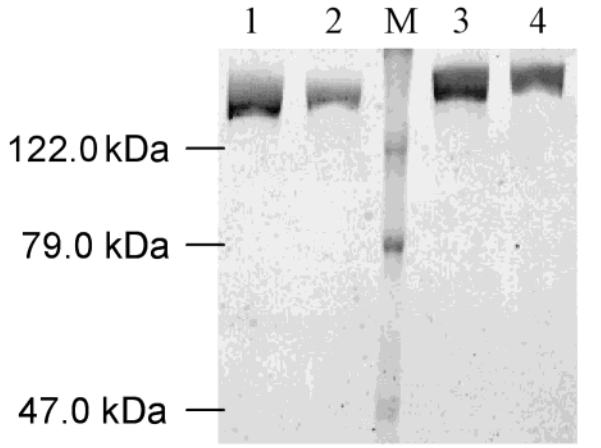


FIGURE 2: The D-Phe-L-Pro-DKP and L/D-Phe-L-Asn artificial NRPS systems are shown schematically. (a) D-Phe-L-Pro-DKP system: TycA-A_{Phe}TE and $TycB_1$-CA_{Pro}T activate and covalently bind L/D-Phe and L-Pro to their 4′-phosphopantheteine cofactors, respectively. Thereby, ATP is hydrolyzed to PP_i and AMP. Subsequently, D-Phe is transferred to $TycB_1$-CA_{Pro}T upon the action of the C-domain. Water is released, and the linear D-Phe-L-Pro is now covalently attached to the cofactor of the acceptor enzyme. Because L-Pro is a secondary amino acid, the nucleophilic attack of the N-terminal amino group of D-Phe on the carboxy group of the thioester becomes favored, so that finally cyclic D-Phe-L-Pro is set free under release of one molecule of water, resulting in the regeneration of the enzymes for a new round of catalysis. (b) L/D-Phe-L-Asn-system: $TycB_3$-A_{Phe}TE and $TycC_1$-CA_{Asn}T/Te activate and covalently bind their substrate amino acids under consumption of ATP. The L/D-Phe is transferred to the acceptor enzyme, and L/D-Phe-L-Asn-S-Ppant is formed under release of water. The Te-domain directly fused at the C-terminus of $TycC_1$ ensures release of linear L/D-Phe-L-Asn and thereby recycles the two enzymes for a new round of catalysis. In contrast to the D-Phe-L-Pro-DKP system, the epimerization reaction is very slow, because the peptidyl-S-Ppant E-domain of $TycB_3$ is optimized to epimerize a peptidyl-S-Ppant substrate. Furthermore, the C-domain of $TycC_1$ normally accepts a D-Phe-L-Pro-L-Phe-D-Phe-S-Ppant and is obviously not capable of discriminating against L-Phe-S-Ppant. In conclusion, when L-Phe is utilized as substrate, a 2:1 mixture of L/D-Phe-L-Asn is formed, while with D-Phe as substrate, the reaction is much more efficient, and exclusively D-Phe-L-Asn is produced.

Generation and Purification of the Recombinant Enzymes. In this study, we tested TycA-A_{Phe}TE (*16*), TycA-A_{Phe}T/E^{TycB3} (*9*) (in which the natural E-domain was exchanged against the E-domain of $TycB_3$), $TycB_3$-A_{Phe}TE (*17*), and $TycB_3$-A_{Phe}T/E^{TycA} (*9*) (in which the natural E-domain was exchanged against the E-domain of TycA) as donor enzymes in the D-Phe-L-Pro-DKP assay (*6*) with $TycB_1$-CA_{Pro}T (*16*) as acceptor enzyme as well as in the L/D-Phe-L-Asn formation assay (*17*) with $TycC_1$-CA_{Asn}T/Te (*17*) as acceptor enzyme (compare Figure 4a). Additionally, the two hybrid proteins $TycC_4$-A_{Val}/TE^{TycA} (the Phe-activating A-domain of TycA-A_{Phe}TE was exchanged against the valine-activating A-



M: Protein-size standard
1: TycA-A_{Phe}TE-$TycB_1$-CA_{Pro}T (241 kDa)
2: $TycB_3$-A_{Phe}TE-$TycB_1$-CA_{Pro}T (241 kDa)
3: TycA-A_{Phe}TE-$TycC_1$-CA_{Asn}T/Te (269 kDa)
4: $TycB_3$-A_{Phe}TE-$TycC_1$-CA_{Asn}T/Te (267 kDa)

FIGURE 3: SDS−PAGE of the dimodular fusion enzymes after ion-exchange chromatography.

domain of $TycC_4$) (*9*) and $BacA_1$-A_{Ile}/TE^{TycA} (the Phe-activating A-domain of TycA-A_{Phe}TE was exchanged against the Val-activating A-domain of $BacA_1$) (*9*) were tested for product formation with $TycB_1$-CA_{Pro}T (see Figure 4a). The overproduction and purification of these enzymes were performed as previously described (for literature, see above). They were purified to homogeneity by single-step Ni^{2+} affinity chromatography (confirmed by SDS−PAGE; data not shown).

We constructed additionally a set of four dimodular fusion enzymes. In their natural context, the modules utilized for the fusions are interacting with other ones *in trans*. The four fusion proteins TycA-A_{Phe}TE/$TycB_1$-CA_{Pro}T (241 kDa), TycA-A_{Phe}TE/$TycC_1$-CA_{Asn}T/Te (269 kDa), $TycB_3$-A_{Phe}TE/$TycB_1$-CA_{Pro}T (241 kDa), and $TycB_3$-A_{Phe}TE/$TycC_1$-CA_{Asn}T/Te (267 kDa) were individually expressed as C-terminal His_6-tagged fusions in the heterologous host *E. coli* (compare Figure 4b). After the purification by Ni^{2+} affinity chromatography, all of them showed the expected band size on SDS−PAGE, but additionally, a second band was observed at ~120 kDa in all cases. Because the size of the smaller band (~120 kDa) correlated with putative single-modular degradation fragments generated by a cleavage of the dimodular enzymes near the fusion site, we applied the Ni-NTA-purified recombinant proteins to ion-exchange chromatography to make sure that no enzyme fragments can interfere with the assays. After ion-exchange chromatography, the dimodular proteins were purified to homogeneity (see Figure 3). For all product formation assays the amount of active protein used was calculated from the results of the aminoacylation assays (data not shown), as described in Experimental Procedures, to guarantee the comparability of product formation activities among the different enzyme pairs (see Experimental Procedures).

Investigation of Protein−Protein Recognition. We assayed all TycA- and $TycB_3$-donor enzymes (TycA-A_{Phe}TE, TycA-A_{Phe}T/E^{TycB3}, $TycB_3$-A_{Phe}TE, and $TycB_3$-A_{Phe}T/E^{Tyc}) described in this study in the DKP assay system (compare Figure 2a) (*6*) and in the L/D-Phe-L-Asn assay system (see Figure 2b) (*17*) for product formation activities. $TycB_1$-CA_{Pro}T and $TycC_1$-CA_{Asn}T/Te were utilized as acceptor enzymes, respectively. Moreover, the set of four dimodular fusion enzymes harboring modules which are naturally interacting with other synthetases *in trans* (TycA-A_{Phe}TE/$TycB_1$-CA_{Pro}T, TycA-A_{Phe}TE/$TycC_1$-CA_{Asn}T/Te, $TycB_3$-A_{Phe}TE/$TycC_1$-CA_{Asn}T/Te, and $TycB_3$-A_{Phe}TE/$TycB_1$-CA_{Pro}T) was also tested in these assays (see Figure 4b). All of the assays were separately done with L-Phe and D-Phe as donor substrates, because the peptidyl-S-Ppant E-domain of $TycB_3$ is known to be less efficient when epimerizing aminoacyl-S-Ppant substrates (*17*, *20*). It had been demonstrated previously that this effect can be overcome in product formation assays by utilizing D-Phe instead of L-Phe as substrate amino acid (*17*). All enzymes are illustrated in Figure 4.

Product Formation in the D*-Phe/*L*-Pro-DKP System.* The results of the D-Phe-L-Pro-DKP assays are shown in Figure 5a. In all cases, relative values were calculated, whereby the value for the wild-type system (TycA-A_{Phe}TE combined with $TycB_1$-CA_{Pro}T *in trans*) with D-Phe as substrate was used as reference (set to 100%). Previously, the D-Phe-L-Pro-DKP formation rate was estimated to be 0.74−0.8 min^{-1} (*16*). All products formed were exclusively D-Phe-L-Pro-DKP. In no case was L-Phe-L-Pro-DKP observed. Wild-type TycA-A_{Phe}TE and the hybrid enzyme $TycB_3$-A_{Phe}T/E^{TycA}, both harboring the E-domain of TycA, showed comparable product formation activities with $TycB_1$-CA_{Pro}T. The activities are in the same range for both substrates, L-Phe and D-Phe (D-Phe, TycA-A_{Phe}TE, 100%; $TycB_3$-A_{Phe}T/E^{TycA}, 93%; L-Phe, TycA-A_{Phe}TE, 141%; $TycB_3$-A_{Phe}T/E^{TycA}, 104%). In contrast, wild-type $TycB_3$-A_{Phe}TE, normally interacting *in trans* with $TycC_1$, was not capable of transferring any Phe to $TycB_1$-CA_{Pro}T *in trans*. Surprisingly, TycA-A_{Phe}T/E^{TycB3}, harboring the same E-domain, showed a residual D-Phe-L-Pro-DKP formation activity (D-Phe, 11%; L-Phe, 7%). Very interestingly, the two dimodular fusion enzymes TycA-A_{Phe}TE/$TycB_1$-CA_{Pro}T and $TycB_3$-A_{Phe}TE/$TycB_1$-CA_{Pro}T, the latter one imitating a system showing no product formation activity *in trans*, both produced comparable amounts of product when D-Phe was used as substrate amino acid (42% and 44%, respectively). In the case of L-Phe as substrate, the fusion enzyme imitating the natural *in trans* system (TycA-A_{Phe}TE/$TycB_1$-CA_{Pro}T) showed a marginally higher amount of product synthesized (64%), while with $TycB_3$-A_{Phe}TE/$TycB_1$-CA_{Pro}T only 8% activity was observed. This can be explained by the poor efficiency of the $TycB_3$ E-domain, which is a peptidyl-S-Ppant E-domain, to epimerize aminoacyl-S-Ppant substrates (*17*). In principle, all dimodular fusion enzymes showed comparable high-product formation activities, indicating an efficient interaction between artificially linked modules.

Product Formation in the L/D*-Phe-*L*-Asn System.* The results of the L/D-Phe-L-Asn assays are shown in Figure 5b. In all cases, relative values were calculated, whereby the value for the wild-type system ($TycB_3$-A_{Phe}TE combined with $TycC_1$-CA_{Asn}T/Te *in trans*) with D-Phe as substrate was used as reference (set to 100%). Previously, the product turnover rate for this system with D-Phe as $TycB_3$ substrate was

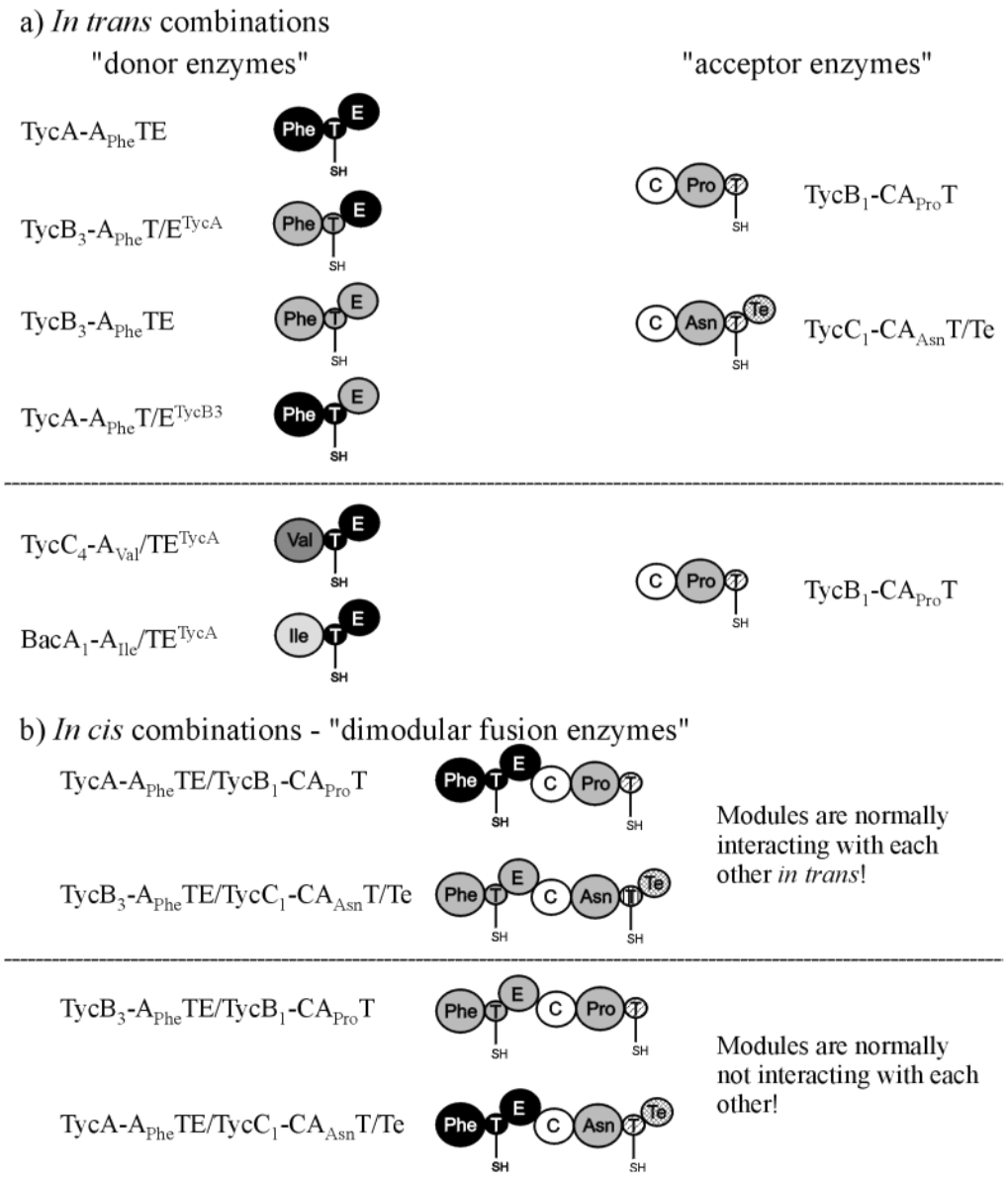


FIGURE 4: Schematic overview of enzyme combinations used in this study. (a) For the *in trans* combinations four enzymes were used as donor enzymes (TycA-A$_{Phe}$TE, TycB$_3$-A$_{Phe}$T/E^{TycA}, TycB$_3$-A$_{Phe}$TE, and TycA-A$_{Phe}$T/E^{TycB3}) and assayed in all combinations with TycB$_1$-CA$_{Pro}$T and TycC$_1$-CA$_{Asn}$T/Te as acceptor enzymes for product formation activities. Black was used for the illustration of TycA domains and gray for TycB$_3$ domains. The acceptor enzymes are shown in the same color codes as in Figures 1 and 2. Additionally, two TycA derivatives containing A-domains for the activation of L-Val (TycC$_4$-A) and L-Ile (BacA$_1$-A) instead of the Phe-activating A-domain of TycA were assayed with TycB$_1$-CA$_{Pro}$T for product formation *in trans*. (b) The four dimodular fusion enzymes used in this study are shown. The upper two *in cis* combinations (TycA-A$_{Phe}$TE/TycB$_1$-CA$_{Pro}$T and TycB$_3$-A$_{Phe}$-TE/TycC$_1$-CA$_{Asn}$T/Te) are imitating the natural *in trans* protein−protein combinations, while the lower ones (TycB$_3$-A$_{Phe}$-TE/TycB$_1$-CA$_{Pro}$T and TycA-A$_{Phe}$TE/TycC$_1$-CA$_{Asn}$T/Te) represent *in cis* combinations of modules normally not interacting with each other.

determined to be 1.16 min^{-1} (*17*). The results of these assays are consistent with the D-Phe-L-Pro-DKP system. Enzymes harboring the E-domain of TycB$_3$ were capable of product formation *in trans* while those with the TycA E-domain were not. For wild-type TycA-A$_{Phe}$TE and the hybrid enzyme TycB$_3$-A$_{Phe}$T/E^{TycA}, absolutely no product formation activity was observed. Interestingly, the hybrid enzyme TycA-A$_{Phe}$T/E^{TycB3} (D-Phe, 204%; L-Phe, 93%) showed higher product formation activity than TycB$_3$-A$_{Phe}$TE (D-Phe, 100%; L-Phe, 10%) itself, which is naturally interacting with the acceptor module.

Again, both dimodular fusion enzymes TycB$_3$-A$_{Phe}$TE/TycC$_1$-CA$_{Asn}$T/Te (D-Phe, 295%; L-Phe, 187%) and TycA-A$_{Phe}$TE/TycC$_1$-CA$_{Asn}$T/Te (D-Phe ,93%; L-Phe, 85%) were active with L-Phe and D-Phe as substrates, regardless of the E-domain they harbored. However, in contrast to the D-Phe-L-Pro-DKP assay system, the latter enzyme, harboring the TycA E-domain naturally not interacting with TycC$_1$ *in trans*, showed a marginally reduced activity compared to TycB$_3$-A$_{Phe}$TE/TycC$_1$-CA$_{Asn}$T/Te (L-Phe, $^1/_2$; D-Phe, $^1/_3$). Furthermore, the fusion of TycB$_3$-A$_{Phe}$TE to the acceptor module led to an increased amount of product synthesis when compared to the *in trans* system.

Val-Pro-DKP and Ile-Pro-DKP Formation. Our results clearly indicated that the E-domains are involved in mediating the protein−protein recognition between the tyrocidine synthetases. It was also demonstrated previously that the electrophilic donor site of the C-domain of TycB$_1$-CA$_{Pro}$T

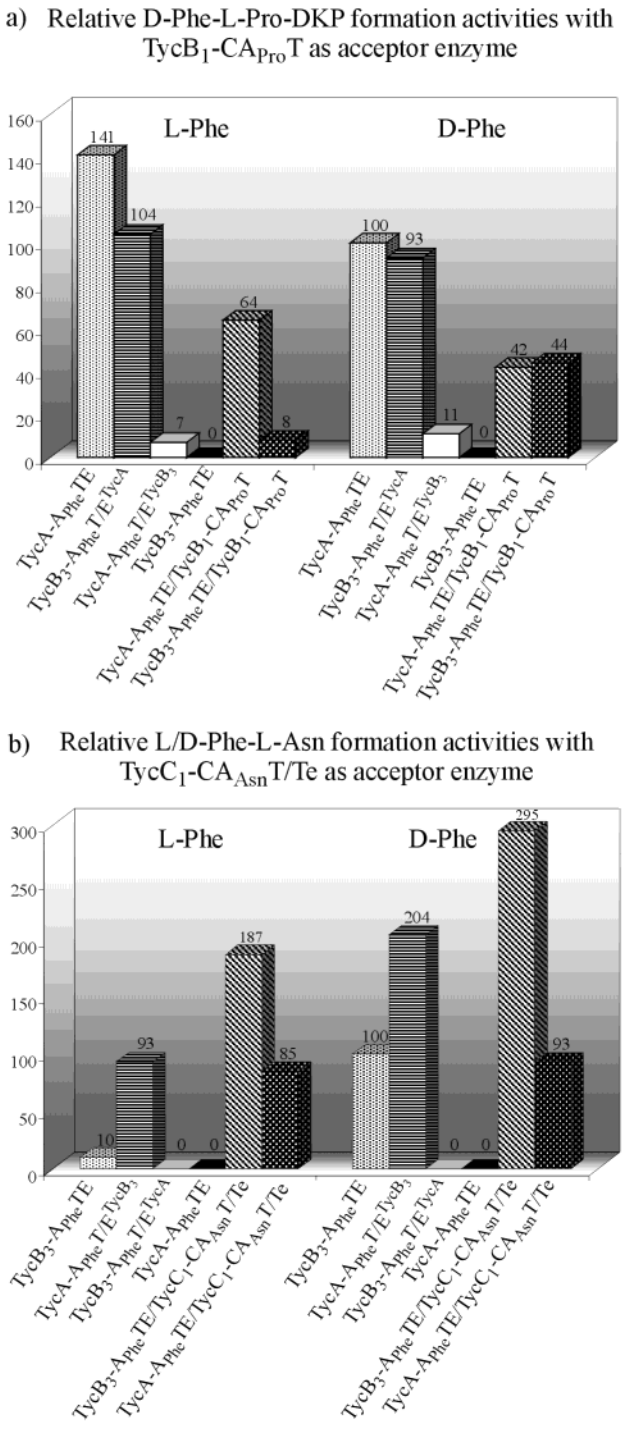


FIGURE 5: Comparison of product formation activities. (a) The relative D-Phe-L-Pro-DKP formation activities with $TycB_1$-$CA_{Pro}T$ as acceptor enzyme are shown for the different donor enzymes with L-Phe (left side) and D-Phe (right side) as substrate amino acids. (b) The relative L/D-Phe-L-Asn formation activities with $TycC_1$-$CA_{Asn}T$/Te as acceptor enzyme are shown for the different donor enzymes with L-Phe (left side) and D-Phe (right side) as substrate amino acids. In all cases, the first two bars of each block represent donor enzymes carrying an E-domain naturally interacting with the corresponding acceptor enzyme, followed by two enzymes with E-domains normally not interacting with it and finally by two dimodular fusion enzymes harboring the donor and acceptor modules on one single polypeptide chain (one with the corresponding donor module, the other with the donor module naturally not interacting with the acceptor). All data represent average values obtained out of at least three independent assays, whereby the margin of error was in all cases less than 5%.

is, at least to some extent, unspecific (*22*). Therefore, we decided to investigate the capability of $TycC_4$-A_{Val}/TE^{TycA} and $BacA_1$-A_{Ile}/TE^{TycA} to transfer Val or Ile to $TycB_1$-$CA_{Pro}T$. In both cases, we were able to detect products by HPLC-ESI-TOF analysis with the masses of Ile-Pro-DKP ($M + H^+$ measured, 211.153; $M + H^+$ calculated, 211.145) and Val-Pro-DKP ($M + H^+$ measured, 197.133; $M + H^+$ calculated, 197.129). The products were formed in a time-dependent manner (compare Figure 6), and no products were detected when ATP or $TycB_1$-$CA_{Pro}T$ was omitted in the assays. In both cases two peaks were obtained (compare Figure 6). With a different column, it was possible to separate these peaks by baseline (data not shown), giving rise to the conclusion that both L,L- and D,L-dipeptide products were formed. However, because no chemical standards were available, we were not able to prove this by co-injection and to quantify the product formation activities exactly. The amount of products formed seemed to be very low when compared to mass and UV signal intensity of D-Phe-Pro-DKP formation under comparable conditions (in the range of ~1%). However, albeit the physical−chemical properties of D-Phe-Pro-DKP with respect to UV absorption and ionization efficiencies should be only slightly different from those of Val-Pro-DKP and Ile-Pro-DKP, they are not directly comparable with each other.

In another set of experiments, it was not possible to detect any linear products (Ile-Pro or Val-Pro) if two $TycB_1$-$CA_{Pro}T$/Te hybrids (*23*) were used as acceptor enzymes, where the Te-domains of tycrocidine or surfactin synthetases were covalently fused to and which were previously shown to cleave the linear dipeptides in another context (*23*) (data not shown).

To demonstrate that Ile and Val were transferred to the acceptor enzyme before product release, we exploited the property of $TycC_1$-$CA_{Pro}T$ to activate primary proline analogons by side specificities (*23*). With primary nucleophilic acceptor substrates such as L-Abu, no product release takes place and the putative dipeptides formed remain covalently bound to the enzymes. Therefore, ^{14}C-labeled L-Ile and L-Val were used to measure transfer to the acceptor enzyme. The donor enzymes $TycC_4$-A_{Val}/TE^{TycA} and $BacA_1$-A_{Ile}/TE^{TycA} were first allowed to covalently load their ^{14}C-labeled substrate amino acids. Without acceptor enzyme, after approximately 20−30 min a stable amount of radioactivity bound to the enzymes was reached (compare Figure 7). When $TycB_1$-$CA_{Pro}T$, which had been separately preincubated with ATP and L-Abu, was added after 1 h to this reaction mixture, the amount of enzyme-bound radioactivity increased significantly (see Figure 7). This effect is interpreted as the transfer of ^{14}C-labeled L-Ile or L-Val to the acceptor enzyme ($TycB_1$-$CA_{Pro}T$) by dipeptide formation, subsequently followed by a reaminoacylation of the donor enzymes with ^{14}C-labeled substrates. In summary, we were able to show that the engineered hybrid enzymes were capable of transferring Ile or Val (instead of D-Phe) to $TycB_1$. However, the efficiencies were very low, and a mixture of diasteriomers seemed to be formed, while in the natural system exclusively D-Phe is transferred to $TycB_1$-$CA_{Pro}T$.

DISCUSSION

Despite the existence of many bacterial NRPS biosynthesis systems consisting of multiple synthetases, in all cases only

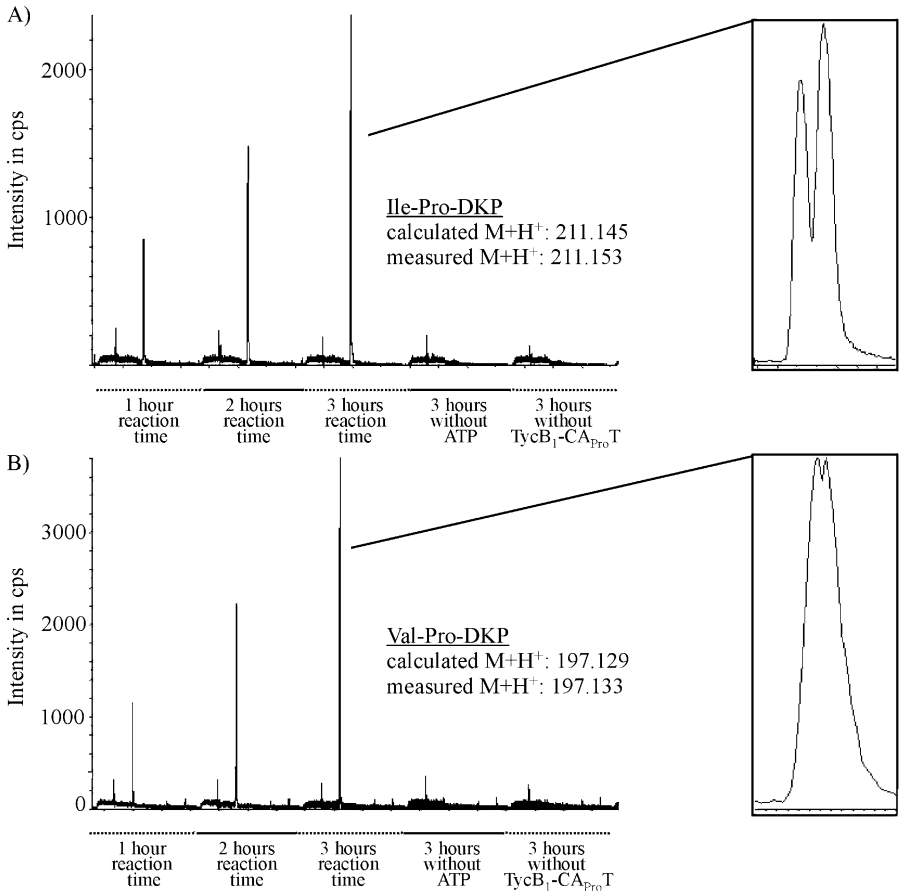


FIGURE 6: LC-ESI-TOF data (extracted ion) of Ile-Pro- and Val-Pro-DKP formation. (A) An assembly of five LC-MS runs is shown. $BacA_1$-A_{Ile}/TE^{TycA} and $TycB_1$-$CA_{Pro}T$ were assayed with each other in the presence of their substrates and in the dependence of the time (1, 2 and 3 h reaction time). Additionally, as negative controls ATP or the acceptor enzyme $TycB_1$-$CA_{Pro}T$ was omitted, respectively. On the right side, an enlargement of one peak (after 3 h reaction time) is shown as exemplary. (B) An assembly of five LC-MS runs is shown. $TycC_4$-A_{Val}/TE^{TycA} and $TycB_1$-$CA_{Pro}T$ were assayed with each other in the presence of their substrates and in the dependence of the time (1, 2, and 3 h reaction time). Additionally, as negative controls ATP or the acceptor enzyme $TycB_1$-$CA_{Pro}T$ was omitted, respectively. On the right side, an enlargement of one peak (after 3 h reaction time) is shown as exemplary.

products of a distinct length and order are found. In principle, one can imagine two mechanisms ensuring this correct order of interaction between the different subunits of a synthetase complex: substrate specificity or selective protein−protein recognition. Previous studies gave some evidence for the latter possibility. It was reported that GrsA-$A_{Phe}T$, where the E-domain was deleted, was not capable of transferring Phe to $TycB_1$-$CA_{Pro}T$, even when D-Phe was utilized as substrate (*6*). Furthermore, mutational analysis of the GrsA E-domain revealed that point mutants abolishing E-domain activity are still capable of transferring D-Phe to $TycB_1$-$CA_{Pro}T$ (*21*). Additionally, it was shown that the C-domain of $TycC_1$, normally elongating a tetrapeptidyl-S-Ppant, can also catalyze peptide bond formation with L/D-Phe-S-Ppant (*17*). The latter substrate is the same as would be provided by TycA. However, a shortened tyrocidine derivative resulting from a mistransfer from TycA to TycC has never been reported so far. These observations suggested selective protein−protein recognitions to govern correct peptide chain transfer between the TycA and TycB, as well as between TycB and TycC synthetase subunits.

In this study, we sytematically and quantitatively investigated the impact of TycA and $TycB_3$ E-domains on selective protein−protein recognition between the tyrocidine synthetases TycA, TycB, and TycC (see Figure 1). Thereby, we took advantage of the similarities between TycA-A_{Phe}TE and $TycB_3$-A_{Phe}TE. Both enzymes show nearly identical substrate specificities. The most significant difference is that the E-domain of TycA is an aminoacyl epimerase, whereas that of $TycB_3$ is a peptidyl epimerase exhibiting low efficiency when epimerizing aminoacyl-S-Ppant substrates (*17*, *20*). However, this effect can be compensated for by utilizing D-Phe as substrate amino acid. For both enzymes, product formation assay systems were developed previously, the DKP (*6*) and the L/D-Phe-L-Asn (*17*) assay systems, to measure transfer between TycA and TycB as well as between TycB and TycC, respectively. Additionally, two hybrid enzymes were tested, in which the E-domains were swapped against each other (TycA-$A_{Phe}T$/E^{TycB3} and $TycB_3$-$A_{Phe}T$/E^{TycA}) (*9*).

Both assay systems used in this study are artificial systems. Therefore, it cannot be totally excluded that the data presented could be marginally skewed. Anyhow, our results clearly indicate that the origin of the E-domain controls protein−protein recognition *in trans*. Although all donor enzymes provided D-Phe, an efficient product formation activity was only detected when the E-domain naturally interacting with the corresponding acceptor enzymes was present in the donor proteins. With noncognate E-domains, no product formation activities were observed, except in the case of TycA-$A_{Phe}T$/E^{TycB3} when assayed with $TycB_1$-$CA_{Pro}T$. For this noncognate E-domain/acceptor enzyme combination,

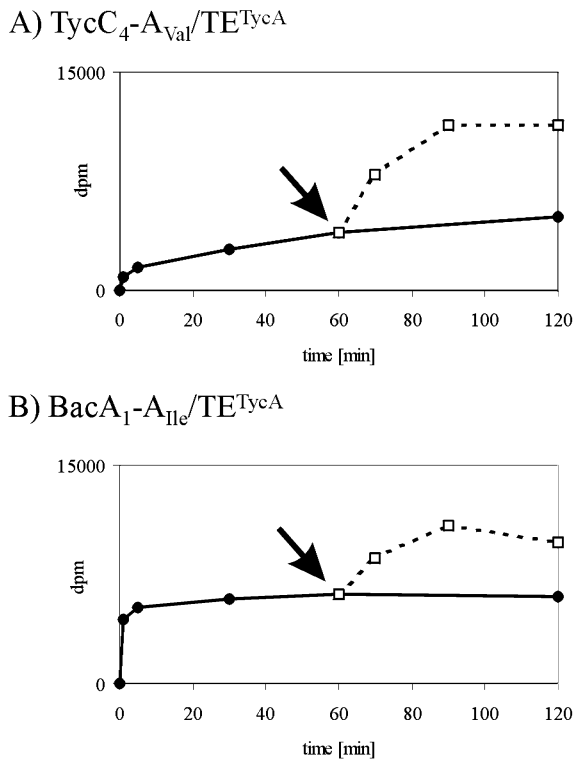


FIGURE 7: Aminoacylation and transfer kinetics. The aminoacylation reactions of $TycC_4$-A_{Val}/TE^{TycA} (A) and $BacA_1$-A_{Ile}/TE^{TycA} (B) with their cognate ^{14}C-labeled substrate amino acids (Val and Ile) are shown in the dependence of the time. At defined time points, aliquots were taken and subsequently quenched by the addition of 1 mL of ice-cold 10% TCA. The acid-stable amount of radioactivity bound to the enzymes was quantified by scintillation counting. At the position marked by an arrow (60 min reaction time), the assays were split. To one half, $TycB_1$-$CA_{Pro}T$ was added, which had been preincubated for 60 min with nonlabeled L-Abu as substrate amino acid. If Val or Ile is transferred to the acceptor enzyme, one would expect an increase of the total amount of radioactivity bound to the enzymes, because after transfer of the labeled amino acids to $TycB_1$-$CA_{Pro}T$, the donor enzymes are capable of covalently loading a new labeled substrate amino acid.

a residual product formation activity (7% with L-Phe and 11% with D-Phe) was detected. One possible explanation for this effect could be that to a minor extent A- and PCP domains also participate as protein−protein recognition elements. However, a similar effect has not been observed for $TycB_3$-$A_{Phe}T$/E^{TycA} when assayed with $TycC_1$-$CA_{Asn}T$/Te. Probably, the C-domain of $TycC_1$ is somehow more efficient for peptidyl-S-Ppant substrates than for aminoacyl-S-Ppant substrates, suppressing such a "residual activity" in the latter enzyme system, while the C-domain of $TycB_1$ is naturally elongating D-Phe-S-Ppant. A surprising result was that the hybrid enzyme TycA-$A_{Phe}T$/E^{TycB3} was 10 times more efficient in product formation with $TycC_1$-$CA_{Asn}T$/Te than $TycB_3$-$A_{Phe}TE$ itself when L-Phe was used as substrate and still two times more efficient with D-Phe. One possible explanation is that the aminoacylation velocity is faster in the case of the hybrid enzyme. However, with L-Phe as substrate, the rate-limiting step should be the epimerization reaction (*20*). On the other hand, a more speculative interpretation would consider a stimulating effect of the PCP^E domain of TycA, normally interacting with an aminoacyl-S-Ppant epimerase, on the peptidyl-S-Ppant E-domain of $TycB_3$ when epimerizing aminoacyl-S-Ppant substrates. This result remains somehow obscure at the moment and will require further investigation.

In the past few years many efforts were undertaken to reprogram and to engineer NRPS templates with the aim to generate enzymatic templates biosynthesizing new valuable products (*16*, *24*−*28*). In this study we did the proof of principle, that swapping PCP-E-didomains represents an additional possibility to obtain new peptide products by directed engineering approaches. In the case of the two hybrid enzymes $TycC_4$-A_{Val}/TE^{TycA} and $BacA_1$-A_{Ile}/TE^{TycA}, we were able to demonstrate that dipeptide-DKP products are formed in the presence of the acceptor enzyme $TycB_1$-$CA_{Pro}T$ and substrates. However, product yields were very low. There are some effects that can be responsible for these observations. Earlier studies revealed that the C-domain of $TycB_1$ is enantioselective for the D-isomer at its electrophilic donor site but at least to some extent unspecific to the nature of the amino acid (*22*). However, no steady-state kinetics were done, so that possibly only the catalytic efficiency toward unnatural substrates is lowered. On the other hand, many examples can be found in the literature for C-domains processing unnatural substrates with acceptable efficiencies (*16*, *24*, *26*). Therefore, the C-domain of $TycB_1$ should accept unnatural substrates in their D-configuration at least with a moderate activity. Surprisingly, in the case of the two hybrid enzymes tested, always two substances which had the same masses were observed, giving strong evidence for being a mixture of L- and D-Val/Ile-Pro-DKP diastereoisomers, respectively. Amino acid transfer studies clearly demonstrated that product formation takes place by transfer of Val and Ile to the acceptor enzyme and not by an unspecific reaction of reactive compounds in the reaction mixture. Thus, the $TycB_1$ C-domain does not seem to be strictly enantioselective toward altered substrates. Another effect which probably contributes to the low efficiencies of the reactions is that the E-domain of TycA is selective for Phe-S-Ppant as substrate and the epimerization efficiency toward unnatural substrates is strongly decreased (*9*, *29*). It was not possible to bypass the E-domain function by utilizing D-Val or D-allo-Ile, because both A-domains ($TycC_4$ and $BacA_1$) exclusively activate the L-isomers (data not shown).

Besides using such hybrid enzymes with PCP E-didomains, which at the moment seem to be only less efficient, we describe a second engineering approach in this study, which could be of much higher relevance for practical approaches in the future. As demonstrated, the E-domains play an important role in mediating protein−protein interactions between synthetases interacting with each other *in trans*. The question arose, if such *in trans* systems would be still active *in cis*. Therefore, the two fusion enzymes TycA-$A_{Phe}TE$/$TycB_1$-$CA_{Pro}T$ and $TycB_3$-$A_{Phe}TE$/$TycC_1$-$CA_{Asn}T$/Te were designed. Our results revealed that it is possible to fuse modules *in cis*, which normally interact with each other *in trans*. The product formation activities *in cis* were slightly reduced for the first dimodular enzyme (∼50%) and surprisingly significantly higher for the latter one (three times higher with D-Phe as substrate). Obviously, the modules and domains are still flexible enough to interact functionally with each other when being fused together on one single polypeptide chain. Very interestingly, the two unnatural *in cis* fusions, TycA-$A_{Phe}TE$/$TycC_1$-$CA_{Asn}T$/Te and $TycB_3$-$A_{Phe}TE$/$TycB_1$-$CA_{Pro}T$, were also active, with comparable activities

to that of TycA-A_{Phe}TE/TycB_1-CA_{Pro}T and TycB_3-A_{Phe}TE/TycC_1-CA_{Asn}T/Te, indicating that the protein−protein recognition mediated by the E-domains is only necessary for the *in trans* systems. When fused together, the enzymes do not have to find each other in solution, and this close proximity seems to be sufficient for a productive interaction. This result also suggests that no "negative" recognition elements are present, which might lead to repelling forces in fusion enzymes.

In summary, we clearly demonstrated in this study that the E-domains which are localized at C-terminal modules of NRPSs are involved in protein−protein recognition. Furthermore, we were able to show that modules normally interacting with each other *in trans* can be fused together *in cis* under preservation of their activities. Most interestingly, modules not interacting with each other *in trans* become active when fused together *in cis*. Especially the latter finding could be of high relevance for future rational engineering approaches. Further experiments should aim at a more detailed mapping of the protein−protein recognition elements within the C-terminal E-domains, possibly with regard to small linkers as was reported for the polyketide synthases (*30*).

ACKNOWLEDGMENT

We thank Antje Schäfer and Inge Schüler for excellent technical assistance. Alexandra Kitta supported this work by doing some cloning. Dirk Schwarzer provided the two engineered enzymes TycB_1-CA_{Pro}T/Te^{Tyc} and TycB_1-CA_{Pro}T/Te^{Srf} (*23*).

REFERENCES

1. Schwarzer, D., and Marahiel, M. A. (2001) *Naturwissenschaften 88*, 93−101.
2. Conti, E., Stachelhaus, T., Marahiel, M. A., and Brick, P. (1997) *EMBO J. 16*, 4174−4183.
3. Stachelhaus, T., Hüser, A., and Marahiel, M. A. (1996) *Chem. Biol. 3*, 913−921.
4. Lambalot, R. H., Gehring, A. M., Flugel, R. S., Zuber, P., LaCelle, M., Marahiel, M. A., Reid, R., Khosla, C., and Walsh, C. T. (1996) *Chem. Biol. 3*, 923−936.
5. Reuter, K., Mofid, M. R., Marahiel, M. A., and Ficner, R. (1999) *EMBO J. 18*, 6823−6831.
6. Stachelhaus, T., Mootz, H. D., Bergendahl, V., and Marahiel, M. A. (1998) *J. Biol. Chem. 273*, 22773−22781.
7. Bruner, S. D., Weber, T., Kohli, R. M., Schwarzer, D., Marahiel, M. A., Walsh, C. T., and Stubbs, M. T. (2002) *Structure 10*, 301−310.
8. Stein, T., Kluge, B., Vater, J., Franke, P., Otto, A., and Wittmann-Liebold, B. (1995) *Biochemistry 34*, 4633−4642.
9. Linne, U., Doekel, S., and Marahiel, M. A. (2001) *Biochemistry 40*, 15824−15834.
10. Gehring, A. M., Mori, I., Perry, R. D., and Walsh, C. T. (1998) *Biochemistry 37*, 11637−11650.
11. Keating, T. A., Miller, D. A., and Walsh, C. T. (2000) *Biochemistry 39*, 4729−4739.
12. Zocher, R., Nihira, T., Paul, E., Madry, N., Peeters, H., Kleinkauf, H., and Keller, U. (1986) *Biochemistry 25*, 550−553.
13. Schmidt, B., Riesner, D., Lawen, A., and Kleinkauf, H. (1992) *FEBS Lett. 307*, 355−360.
14. Mootz, H. D., and Marahiel, M. A. (1997) *J. Bacteriol. 179*, 6843−6850.
15. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
16. Mootz, H. D., Schwarzer, D., and Marahiel, M. A. (2000) *Proc. Natl. Acad. Sci. U. S. A. 97*, 5848−5853.
17. Linne, U., and Marahiel, M. A. (2000) *Biochemistry 39*, 10439−10447.
18. Laemmli, U. K. (1970) *Nature 227*, 491−493.
19. Bergendahl, V., Linne, U., and Marahiel, M. A. (2002) *Eur. J. Biochem. 269*, 620−629.
20. Luo, L., Kohli, R. M., Onishi, M., Linne, U., Marahiel, M. A., and Walsh, C. T. (2002) *Biochemistry 41*, 9184−9196.
21. Stachelhaus, T., and Walsh, C. T. (2000) *Biochemistry 39*, 5775−5787.
22. Belshaw, P. J., Walsh, C. T., and Stachelhaus, T. (1999) *Science 284*, 486−489.
23. Schwarzer, D., Mootz, H. D., and Marahiel, M. A. (2001) *Chem. Biol. 8*, 997−1010.
24. Mootz, H. D., Kessler, N., Linne, U., Eppelmann, K., Schwarzer, D., and Marahiel, M. A. (2002) *J. Am. Chem. Soc. 124*, 10980−10981.
25. Mootz, H. D., and Marahiel, M. A. (1999) *Curr. Opin. Biotechnol. 10*, 341−348.
26. Doekel, S., and Marahiel, M. A. (2000) *Chem. Biol. 7*, 373−384.
27. Schauwecker, F., Pfennig, F., Grammel, N., and Keller, U. (2000) *Chem. Biol. 7*, 287−297.
28. Stachelhaus, T., Schneider, A., and Marahiel, M. A. (1995) *Science 269*, 69−72.
29. Luo, L., Burkart, M. D., Stachelhaus, T., and Walsh, C. T. (2001) *J. Am. Chem. Soc. 123*, 11208−11218.
30. Gokhale, R. S., Tsuji, S. Y., Cane, D. E., and Khosla, C. (1999) *Science 284*, 482−485.

BI034223O