Eur. J. Biochem. **270**, 4555–4563 (2003) © FEBS 2003 doi:10.1046/j.1432-1033.2003.03858.x

Construction of hybrid peptide synthetases for the production of α-L-aspartyl-L-phenylalanine, a precursor for the high-intensity sweetener aspartame

Thomas Duerfahrt[1], Sascha Doekel[1,*], Theo Sonke[2], Peter J. L. M. Quaedflieg[2] and Mohamed A. Marahiel[1]

[1]*Philipps-Universität Marburg, Fachbereich Chemie/Biochemie, Marburg, Germany;* [2]*DSM Research, Life Sciences – Advanced Synthesis, Catalysis and Development, Geleen, the Netherlands*

Microorganisms produce a large number of pharmacologically and biotechnologically important peptides by using nonribosomal peptide synthetases (NRPSs). Due to their modular arrangement and their domain organization NRPSs are particularly suitable for engineering recombinant proteins for the production of novel peptides with interesting properties. In order to compare different strategies of domain assembling and module fusions we focused on the selective construction of a set of peptide synthetases that catalyze the formation of the dipeptide α-L-aspartyl-L-phenylalanine (Asp-Phe), the precursor of the high-intensity sweetener α-L-aspartyl-L-phenylalanine methyl ester (aspartame). The *de novo* design of six different Asp-Phe synthetases was achieved by fusion of Asp and Phe activating modules comprising adenylation, peptidyl carrier protein and condensation domains. Product release was ensured by a C-terminally fused thioesterase domains and quantified by HPLC/MS analysis. Significant differences of enzyme activity caused by the fusion strategies were observed. Two forms of the Asp-Phe dipeptide were detected, the expected α-Asp-Phe and the by-product β-Asp-Phe. Dependent on the turnover rates ranging from 0.01–0.7 min^{-1}, the amount of α-Asp-Phe was between 75 and 100% of overall product, indicating a direct correlation between the turnover numbers and the ratios of α-Asp-Phe to β-Asp-Phe. Taken together these results provide useful guidelines for the rational construction of hybrid peptide synthetases.

Keywords: nonribosomal peptide synthesis; rational protein engineering; hybrid enzyme; module exchange; aspartame.

Nonribosomal peptide synthesis is an alternative method for the production of pharmacologically and biotechnologically important products. Many microorganisms produce these as secondary metabolites by using nonribosomal peptide synthetases (NRPSs) [1]. The assembly of nonribosomally synthesized peptides is determined by the sequence of catalytically independent modules of the multifunctional NRPSs [2–4]. Each module consists of a set of catalytic domains essential for the incorporation of one substrate amino acid into the nascent peptide chain through a thiotemplate mechanism [5]. Within one module, the specific recognition and ATP-dependent activation of an amino acid is managed by the adenylation (A) domain that forms an amino acyl adenylate intermediate [6]. Subsequently, this intermediate is covalently bound to the 4′-phosphopantetheinyl (Ppant) cofactor of the corresponding peptidyl carrier protein (PCP) [7]. The activation of the PCP domain by post-translational modification with the Ppant cofactor is catalyzed by a 4′-Ppant transferase [8,9]. The condensation of the PCP-bound aminoacyl or peptidyl moieties with the thioester-linked aminoacyl of two adjacent modules is catalyzed by the condensation (C) domain [10]. A module comprising these three domains, C, A and PCP, can be described as a minimal elongation module which is responsible for recognition, activation and condensation of one specific amino acid. By contrast initiation of nonribosomal peptide synthesis is catalyzed by an initiation module which lacks the N-terminal C domain. Structural diversity of NRPS products is enlarged by optional domains catalyzing substrate modifications such as epimerization [11], heterocyclization [12] or N-methylation [13,14]. After the last elongation step, the synthesized peptide is usually released by the action of a thioesterase (Te)-like domain which is appended to the last module [15,16].

Due to the assembly line-like structure and the architecture of catalytically independent domains, NRPSs are especially suitable for the construction of hybrid enzymes in order to obtain new nonribosomally synthesized compounds [2]. It has previously been shown that such hybrid synthetases can be generated by splicing together gene fragments from different origins. Fusions of domains or modules have led to catalytically active hybrid NRPSs

Correspondence to M. A. Marahiel, Philipps-Universität Marburg, Fachbereich Chemie/Biochemie, Hans-Meerwein-Strasse, D-35032 Marburg, Germany.
Fax: + 49 6421 2822191, Tel.: + 49 6421 2825722,
E-mail: marahiel@chemie.uni-marburg.de

Abbreviations: A, adenylation domain; APM, α-L-aspartyl-L-phenylalanine methyl ester; Asp-Phe, α-L-aspartyl-L-phenylalanine; C, condensation domain; NRPS, nonribosomal peptide synthetase; PCP, peptidyl carrier protein; Ppant, 4′-phosphopantetheine; Te, thioesterase domain.

Present address: Cubist Pharmaceuticals Inc., Lexington, MA, USA.

(Received 20 August 2003, revised 23 September 2003, accepted 29 September 2003)

[14,17,18]. To further enlarge the potential of constructing different hybrid NRPSs, modules can be used for initiation of peptide synthesis by removing the inherent C domain [19]. This allows nonribosomal peptide synthesis to start with an even larger number of substrates. As in native NRPS-systems, product release in recombinant NRPSs can be realized by a C-terminal Te domain which catalyzes the cleavage of the assembled peptides with high efficiency [20].

The increasing number of NRPS systems known, the knowledge about mechanisms of synthetases, modules and individual domains and the experience in constructing hybrid synthetases should, in principle, allow rational design of completely new hybrid peptide synthetases. In this paper, we describe the convenient design of dimodular peptide synthetases enabling the enzymatic production of the dipeptide α-L-aspartyl-L-phenylalanine (Asp-Phe). Asp-Phe is an industrially interesting dipeptide, which is used as a precursor in the production of α-L-aspartyl-L-phenylalanine methyl ester (APM) known as aspartame, a high intensity artificial sweetener, which has a 200-fold higher sweetness than sucrose. APM is currently commercially produced on a scale of many thousands of tons per year by two different routes. In the DSM/Tosoh process *N*-benzyloxycarbonyl-L-Asp is enzymatically coupled to the methyl ester of L-phenylalanine using thermolysin to give *N*-benzyloxycarbonyl-APM which is subsequently deprotected to APM by hydrogenolysis [21]. In the process developed by Nutra-Sweet/Monsanto the internal anhydride of *N*-formyl-L-Asp is chemically condensed with L-Phe giving an approximately 8 : 2 mixture of α- and β-*N*-formyl-Asp-Phe which is subsequently deformylated and simultaneously esterified by treatment with HCl/water/methanol giving α-APM·HCl precipitate in low yield [22].

The aim of this work is to design by different fusion strategies several dimodular peptide synthetases and to explore their potential for efficient Asp-Phe formation (Fig. 1).

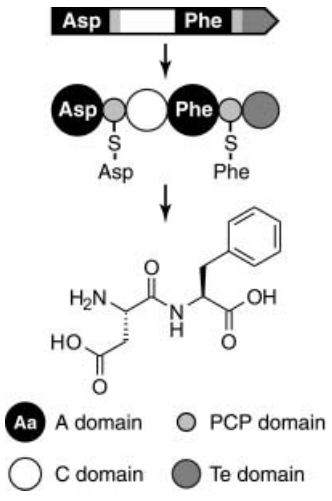

Fig. 1. Nonribosomal Asp-Phe formation. Hybrid dimodular peptide synthetases were generated by fusion of native gene fragments, to contain the domain sequence A(Asp)-PCP-C-A(Phe)-PCP-Te, and used for *in vitro* production of the dipeptide Asp-Phe.

Experimental procedures

Bacterial strains

Recombinant plasmids were prepared using *Escherichia coli* XL1 Blue and *E. coli* XL10 Gold, respectively (Stratagene, Heidelberg, Germany). Overproduction of recombinant proteins was performed as described before using *E. coli* BL21(λDE3)-gsp [23]. Therefore, cells were grown in 2xYT medium supplemented with 25 $\mu g \cdot mL^{-1}$ kanamycin and 100 $\mu g \cdot mL^{-1}$ ampicillin (final concentrations).

Cloning of hybrid dimodular *NRPS* genes

DNA manipulations and *E. coli* transformations were carried out using standard methods [24].

All recombinant gene fragments were amplified from chromosomal DNA of *Bacillus subtilis* ATCC 21332 or *Bacillus brevis* ATCC 8185 using Expand long-range PCR-kit (Roche, Mannheim, Germany). For introducing restriction sites, modified oligonucleotides (MWG Biotech, Ebersberg, Germany) were used. All plasmids are based on the pQE system (Qiagen, Hilden Germany). The plasmids containing the gene fragments encoding the hybrid synthetases are numbered according to the numbering of the enzymes I–VI.

The plasmid pAspPhe-1 is based on the pQE70-vector. A DNA fragment encoding the domains A-PCP-C-A-PCP of surfactin synthetase modules SrfB2-3 was amplified from chromosomal DNA from *B. subtilis* ATCC 21332 using the primers 5′-TAA**GCATGC**TGCTTTCATCTGCAGAA AC-3′ and 5′-AAT**GGATCC**TTCGGCACGCTCTAC-3′ (restriction sites in bold). After digestion with *Sph*I and *Bam*HI the 4903 bp DNA fragment was cloned into the *Sph*I and *Bam*HI sites of pQE70 resulting in the plasmid p[A-PCP-C-A-PCP]$_{srfB2}$. The gene fragment encoding the A domain of tyrocidine synthetase TycA was obtained using chromosomal DNA from *B. brevis* ATCC 8185 and the oligonucleotides 5′-ATTT**GGTCACC**AATCTCATCGA CAA-3′ and 5′-ATA**GGATCC**TGTATTCGTAAAGTT TTTC-3′. The resulting PCR fragment was digested with *Bst*EII and *Bam*HI. The 1888 bp fragment was ligated into p[A-PCP-C-A-PCP]$_{srfB2}$ digested with the same enzymes, yielding p[A-PCP-C]$_{srfB2}$-[A-PCP]$_{tycA}$. This plasmid was digested with *Cla*I and *Bgl*II to remove the 3′-terminal *tycA* DNA and ligated with a 897 bp [PCP-Te]$_{srfC}$ DNA fragment, which was amplified from chromosomal DNA from *B. subtilis* ATCC 21332 by using the primers 5′-ATA**ATCGAT**AATCGCACAAATATGGTC-3′ and 5′-ATA**AGATCT**AACCGTTACGGTTTGTGT-3′ and digested with *Cla*I and *Bgl*II, resulting in the plasmid pAspPhe-1 containing the gene fragment [A-PCP-C]$_{srfB2}$-[A]$_{tycA}$-[PCP-Te]$_{srfC}$ encoding the construct enzyme I.

The plasmid pAspPhe-3 is a derivative of p[A]$_{bacA1}$-[PCP-C-A]$_{tycB2}$-[PCP-Te]$_{tycC6}$ based on the pQE60-vector [18]. A DNA fragment encoding the domains A-PCP of surfactin synthetase module SrfB2 was amplified from chromosomal DNA from *B. subtilis* ATCC 21332 using the primers 5′-TAA**CCATGG**TGCTTTCATCTGCAGAAAC-3′ and 5′-TAT**GATATC**CTCCATATAAGCCGC-3′. After digestion with *Nco*I and *Eco*RV the 1816 bp DNA fragment was ligated into the *Nco*I and *Eco*RV digested

p[A]$_{bacA1}$-[PCP-C-A]$_{tycB2}$-[PCP-Te]$_{tycC6}$ to get the plasmid pAspPhe-3 containing the gene fragment [A-PCP]$_{srfB2}$-[C-A]$_{tycB2}$-[PCP-Te]$_{tycC6}$ encoding the enzyme II.

The plasmids pAspPhe-2 and pAspPhe-4 are derivatives of pAspPhe-3. The DNA fragment encoding the domains PCP-Te of surfactin synthetase SrfC was amplified from chromosomal DNA from *B. subtilis* ATCC 21332 using the primers 5′-TAT**GTTAAC**TGGATTGGACCGCG GAAC-3′ and 5′-TAT**GGATCC**TGAAACCGTTACG GTTTGTG-3′. The primers 5′-TAT**GTTAAC**GAATAC GTGGCCCCGAG-3′ and 5′-TAT**GGATCC**GAAATC GGCCACCTTTTCG-3′ were used to amplify the DNA fragment encoding the PCP-domain of tyrocidine synthetase module TycC6 from chromosomal DNA from *B. brevis* ATCC 8185. The PCR-fragments were digested with *Hpa*I and *Bam*HI, yielding the 931 bp [PCP-Te]$_{srfC}$ and the 226 bp [PCP]$_{tycC}$ fragment, respectively. These were cloned into *Hpa*I and *Bam*HI-digested p[A-PCP]$_{srfB2}$-[C-A]$_{tycB2}$-[PCP-Te]$_{tycC6}$, resulting in the plasmids pAspPhe-2 ([A-PCP]$_{srfB2}$-[C-A]$_{tycB2}$-[PCP-Te]$_{srfC}$) and pAspPhe-4 ([A-PCP]$_{srfB2}$-[C-A]$_{tycB2}$-[PCP]$_{tycC6}$) for overproduction of the respective synthetases enzyme II and enzyme IV.

To assemble pAspPhe-5 first the [A-PCP] fragment of the *srfB2*-gene was amplified from chromosomal DNA from *B. subtilis* ATCC 21332 by using the primers 5′-TAA**CCATGG**TGCTTTCATCTGCAGAAAC-3′ and 5′-TAT**GGATCC**CTCCATATAAGCCGC-3′. After digestion with *Nco*I and *Bam*HI the 1814 bp fragment was cloned into the *Nco*I and *Bam*HI sites of pQE60 to give p[A-PCP]$_{srfB2}$. The gene fragment encoding the module [C-A-PCP]$_{tycB2}$ was amplified from chromosomal DNA from *B. brevis* ATCC 8185 using oligonucleotides 5′-ATT**AGATCT**GAGGAAAGCGCGTATCTCG-3′ and 5′-AAT**AGATCT**TTCGATCAAGCGGGCCAAG-3′. The PCR-fragment was digested with *Bgl*II and the resulting 3105 bp fragment was ligated into p[A-PCP]$_{srfB2}$ linearized with *Bam*HI, yielding p[A-PCP]$_{srfB2}$[C-A-PCP]$_{tycB2}$. The orientation of the insert was identified by restriction analysis. The last step was the cloning of the gene fragment encoding the Te domain of tyrocidine synthetase TycC. The [Te]$_{tycC6}$ gene fragment was amplified from chromosomal DNA from *B. brevis* ATCC 8185 by using the primers 5′-TAA**AGATCT**GCCATTTTGTTAAATCAG-3′ and 5′-TAT**GGATCC**TTTCAGGATGAACAGTTCTTG-3′ and digested with *Bgl*II and *Bam*HI. To obtain pAspPhe-5 ([A-PCP]$_{srfB2}$-[C-A-PCP]$_{tycB2}$-[Te]$_{tycC6}$) the resulting 732 bp DNA fragment was ligated into the *Bgl*II site of p[A-PCP]$_{srfB2}$[C-A-PCP]$_{tycB2}$. This plasmid pAspPhe-5 was used for overproduction of enzyme V.

The plasmid pAspPhe-6 was derived form the plasmid p[A-PCP-C-A-PCP]$_{srfB2}$ (see construction of pAspPhe-1). The gene fragment encoding the A and PCP domain of tyrocidin synthetase TycB2 was amplified from chromosomal DNA from *B. brevis* ATCC 8185 using the primers 5′-ATA**GGTCACC**GCGCATGAGAAGCAG-3′ and 5′-AAT**GGATCC**TTCGATCAAGCGGGCC-3′. The PCR fragment was digested with *Bst*EII and *Bam*HI. The 1789 bp fragment was ligated into p[A-PCP-C-A-PCP]$_{srfB2}$ which was digested with the same enzymes yielding p[A-PCP-C]$_{srfB2}$-[A-PCP]$_{tycB2}$. The [Te]$_{tycC6}$ gene fragment was amplified from chromosomal DNA from *B. brevis* ATCC 8185 by using the primers 5′-TAA**AGATCT**GCCATT TTGTTAAATCAG-3′ and 5′-TAT**GGATCC**TTTCAG GATGAACAGTTCTTG-3′ and digested with *Bgl*II and *Bam*HI. The resulting 732 bp DNA fragment was ligated into the *Bam*HI and *Bgl*II digested p[A-PCP-C]$_{srfB2}$-[A-PCP]$_{tycB2}$, yielding the plasmid pAspPhe-6 containing the gene fragment [A-PCP-C]$_{srfB2}$-[A-PCP]$_{tycB2}$-[Te]$_{tycC6}$ encoding the construct enzyme VI.

Overproduction and purification of recombinant enzymes

Escherichia coli BL21(λDE3)-gsp was transformed with the plasmids mentioned above. This *E. coli* strain carries the pREP4-gsp plasmid supplying the gene *gsp* coding for the 4′-Ppant transferase Gsp [23]. The cells were induced at D_{600} = 0.6–0.7 with 0.2 mM isopropyl thio-β-D-galactoside and incubated for an additional 1.5 h at 30 °C before being harvested. Purification using Ni^{2+}-NTA chromatography was performed as described before [20]. Overproduction and purification were analyzed using SDS/PAGE analysis (7.5% Laemmli gels). Pooled fractions were dialyzed against assay buffer (50 mM Hepes, 300 mM NaCl, 1 mM EDTA, pH 8.0) using HiTrap™ desalting columns (Amersham Biosciences). Protein concentrations were determined using the Bradford method [24].

Enzyme assays

To examine A-domain activity the ATP/PP_i exchange reaction was carried out as described previously [23]. Enzyme (20 pmol) in 100 μL reaction volume were incubated with substrate amino acids at a final concentration of 1 mM. The thioester formation assay was performed as described before applying radiolabeled [^{14}C]aspartic acid (200 mCi·mmol^{-1}) and [^{14}C]phenylalanine (450 mCi·mmol^{-1}) [23]. For this purpose 50 pmol enzyme were incubated for 15 min at 37 °C with 150 pmol of radiolabeled amino acid.

Product formation assay

Dipeptide formation by the hybrid synthetases was investigated by using nonlabeled substrate amino acids and subsequent HPLC/MS analysis. The product assay was performed as described before [18]. HPLC/MS analysis was carried out using 1100 MSD HPLC-Systems (Hewlett Packard) and a Nucleosil C18 3/120–250 column (Macherey-Nagel, Dueren, Germany). Amount of formed dipeptide was determined by comparison to solutions of α-Asp-Phe and β-Asp-Phe reference compounds with different concentrations. The amount of Asp-Phe was calculated by integration of the HPLC/MS signals after certain time of incubation. Product formation rates were determined from initial linear turnover.

Results

In order to develop dimodular NRPSs which would enable the nonribosomal production of the dipeptide Asp-Phe, we designed a set of constructs, using different modules and various fusion sites for combination of L-Asp and L-Phe activating domains. Module and domain fusions were realized by combination of the corresponding gene fragments.

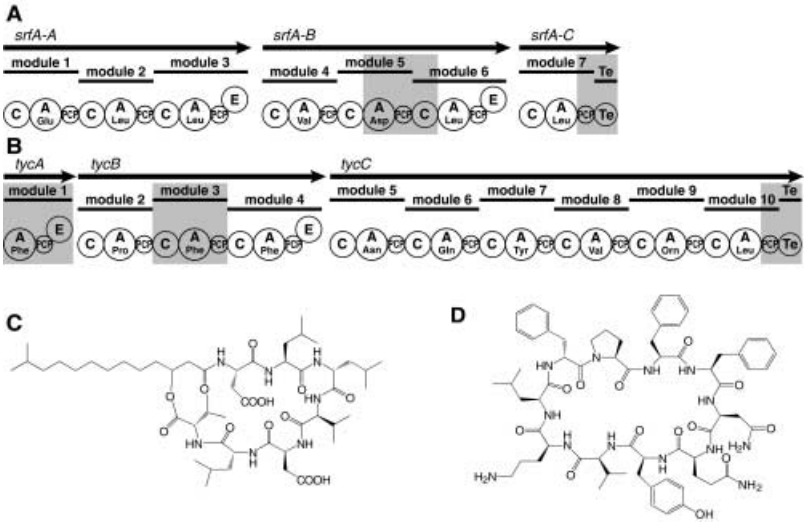

Fig. 2. Organization of native NRPS operons. (A) Organization of the surfactin (*srf*) operon of *Bacillus subtilis* ATCC 21332. The three genes *srfA-A*, *srfA-B*, *srfA-C* and the corresponding synthetases are shown. SrfA and SrfB consists of three modules each, SrfC represents the last module of the surfactin synthetase assembly line. (B) Organization of the tyrocidine (*tyc*) operon of *Bacillus brevis* ATCC 8185. The three genes *tycA, tycB* and *tycC* encode the NRPSs TycA, TycB and TycC comprising one, three and six modules, respectively. Modules and domains used to generate the hybrid Asp-Phe-NRPSs are underscored with gray fields. (C) Surfactin is a cyclic heptapeptide with an N-terminal condensed fatty acid (FA-Glu-Leu-D-Leu-Val-Asp-D-Leu-Leu). (D) Tyrocidine is a cyclic decapeptide (D-Phe-Pro-Phe-D-Phe-Asn-Gln-Tyr-Val-Orn-Leu).

All recombinant enzymes described here were assembled using domains from the well-known biosynthetic operons of surfactin (Srf) and tyrocidine (Tyc) of *B. subtilis* ATCC 21332 and *B. brevis* ATCC 8185, respectively (Fig. 2) [25,26]. All constructs contained the A and PCP domains of the Asp activating module of surfactin synthetase SrfB2. As the second module for dimodule construction, the Phe activating modules of either tyrocidine synthetases TycA or TycB2 were selected. In order to obtain catalyzed product release, termination domains (Te or PCP-Te) of Srf and Tyc operons were used. Recent studies have shown that these Te domains are able to catalyze the cleavage of peptide products in hybrid NRPSs with high efficiency [20]. We tried various arrangements to develop a hybrid synthetase with best domain communication. The domains were combined utilizing artificial fusion sites within the linker regions between the respective domains [18,27,28]. The interdomain linkers of about 7–15 amino acids were identified by sequence alignments. Furthermore, the definition of the boundaries of PCP domains is supported by the NMR-structure of the PCP domain of TycC3 [29]. Domain fusions between C and A were performed on the genetic level at position 91 amino acids downstream of the C7 core motif and 40 amino acids upstream of the core A1. Fusions with PCP domains were located 38/39 residues upstream or downstream of the invariant serine of the PCP. The dimodular synthetases constructed in this work are shown in Fig. 3.

The described enzymes I–VI lack the N-terminal C domain of the Asp processing module. Therefore, synthesis of Asp-Phe is initiated by the selective activation of aspartic acid. The hybrid $[A\text{-}PCP\text{-}C]_{SrfB2}\text{-}[A]_{TycA}\text{-}[PCP\text{-}Te]_{SrfC}$ (enzyme I; Fig. 3) is a particular construct, because it is the only hybrid equipped with the Phe activating A domain of TycA synthetase. TycA is a single module

I $[A\text{-}PCP\text{-}C]_{SrfB2}\text{-}[A]_{TycA}\text{-}[PCP\text{-}Te]_{SrfC}$ Asp Phe

II $[A\text{-}PCP]_{SrfB2}\text{-}[C\text{-}A]_{TycB2}\text{-}[PCP\text{-}Te]_{SrfC}$ Asp Phe

III $[A\text{-}PCP]_{SrfB2}\text{-}[C\text{-}A]_{TycB2}\text{-}[PCP\text{-}Te]_{TycC6}$ Asp Phe

IV $[A\text{-}PCP]_{SrfB2}\text{-}[C\text{-}A]_{TycB2}\text{-}[PCP]_{TycC6}$ Asp Phe

V $[A\text{-}PCP]_{SrfB2}\text{-}[C\text{-}A\text{-}PCP]_{TycB2}\text{-}[Te]_{TycC6}$ Asp Phe

VI $[A\text{-}PCP\text{-}C]_{SrfB2}\text{-}[A\text{-}PCP]_{TycB2}\text{-}[Te]_{TycC6}$ Asp Phe

Fig. 3. Hybrid dimodular Asp-Phe-NRPSs described in this work. Fragments containing native domains are specified with square brackets and are named according to the original synthetases they were taken from (e.g. SrfB2). Fusion sites are indicated by dashes between the brackets. In the schematic illustration domains from different origins are indicated by different shades of gray.

synthetase responsible for initiation of tyrocidine synthesis. It is still uncertain whether domains of a native initiation module can be converted into a part of an elongation module. Furthermore, it has been shown that the C domain exhibits acceptor site specificity [30]. In case of enzyme I the C domains has to recognize phenylalanine at its acceptor site instead of leucine in the native connection. Additionally the hybrid enzyme $[A\text{-}PCP\text{-}C]_{SrfB2}\text{-}[A\text{-}PCP]_{TycB2}\text{-}[Te]_{TycC6}$ (enzyme VI) was made to clarify the effect of fusions between C and A domains, and to investigate whether the preference of the C domain for leucine at the acceptor site might slow down the formation of Asp-Phe. To avoid possible negative effects of such specificity on the condensation reaction, hybrids II, III, IV and V were constructed to have the same C and A domains from the native elongation

module TycB2. The hybrid synthetases $[A\text{-}PCP]_{SrfB2}$-$[C\text{-}A]_{TycB2}$-$[PCP\text{-}Te]_{SrfC}$ (enzyme II) and $[A\text{-}PCP]_{SrfB2}$-$[C\text{-}A]_{TycB2}$-$[PCP\text{-}Te]_{TycC6}$ (enzyme III) differ in the fused Te domain. Enzyme II was provided with the PCP-Te fragment of SrfC, whereas enzyme III carries the PCP-Te fragment of TycC6. To investigate the importance of the Te domain for product release the dimodular synthetase $[A\text{-}PCP]_{SrfB2}$-$[C\text{-}A]_{TycB2}$-$[PCP]_{TycC6}$ (enzyme IV) which lacks the Te domain, was constructed. Hybrid V was engineered with a fusion site within the linker region between PCP and Te domain $[A\text{-}PCP]_{SrfB2}$-$[C\text{-}A\text{-}PCP]_{TycB2}$-$[Te]_{TycC6}$ (enzyme V). In contrast to this, enzymes II and III were obtained by fusions between A (Phe) and PCP domains.

Genes encoding the recombinant proteins were cloned into pQE60 and pQE70, respectively, and expressed in *E. coli* BL21. From the pREP4-*gsp* plasmid the *gsp* gene encoding the 4′-Ppant transferase Gsp was coexpressed to ensure the post-translational modification of enzymes into their active holo-form [23]. Recombinant proteins with predicted masses of 209–212 kDa (enzymes I, II, III, V, VI) and 183 kDa (enzyme IV) were obtained with a yield of 2.5–5 mg protein per 1 L culture. Proteins were purified using Ni^{2+}-NTA affinity chromatography.

Catalytic activities

The A domains are responsible for recognition and activation of the substrates as acyl adenylates. To investigate the activity of the A domains in the different hybrid dimodular NRPSs the ATP/PP_i exchange assay was used. All recombinant enzymes recognized and activated the cognate amino acids L-aspartic acid and L-phenylalanine as expected. For easier comparison of the quality of the amino acid activation the highest $^{32}PP_i$ incorporation into ATP was taken as 100% (ATP/PP_i-exchange reaction, see materials and methods). Enzyme I forms aminoacyl adenylates from L-Asp (45%) and L-Phe (100%). The Phe A domain in this case also activates the stereoisomeric form D-Phe (98%). Lower rates were observed for Thr (8%), Trp (6%), Tyr (4%) and Met (3%). Other amino acids gave rates of less than 1%. All enzymes constructed with the A domain of TycB2(II-VI) exhibited nearly the same ratio of specificities for L-Asp (40%) and L-Phe (100%). In these constructs the activation of D-Phe (27%) was decreased compared to enzyme I. Aside from the D-Phe site specificity all recombinant enzymes showed similar activation of the cognate amino acids L-Asp and L-Phe.

Activated amino acids are covalently bound to the Ppant-arm of the PCP domain adjacent to the activating A domain. The efficiency of communication between A and PCP and the functionality of the PCP domains were tested using the thioester formation assay. Due to the coexpression of the Ppant transferase Gsp, all enzymes were assumed to be present in the modified holo-form. Substrate specificities exhibited in the ATP/PP_i exchange assay were confirmed with the thioester formation assay. The recombinant proteins were incubated with radiolabeled L-Asp and L-Phe, respectively. Both L-Asp and L-Phe became covalently tethered to the enzymes regardless of the constructs tested. Approximately 10–15% of total protein could be loaded with L-Asp and 40–50% could be loaded with L-Phe. Additional treatment of the purified hybrid proteins with Ppant-transferases Sfp or Gsp *in vitro* did not lead to higher enzyme activity. In summary, ATP/PP_i exchange and thioester formation assay demonstrate that the catalytic activity of the combined A and PCP domains is guaranteed by the way the fusions were realized.

Product formation

Asp-Phe product formation catalyzed by the different hybrid dimodular NRPSs was investigated using HPLC/MS analysis. The recombinant proteins were incubated with nonlabeled substrate amino acids and reaction products were identified using reference materials and by the $[M + H]^+$ mass of Asp-Phe (calculated $[M + H]^+ = 281$ Da). We observed Asp-Phe formation for all constructs, depending on the presence of ATP and both substrate amino acids. In Fig. 4, a typical HPLC/MS diagram and a mass spectrum of the detected α-Asp-Phe

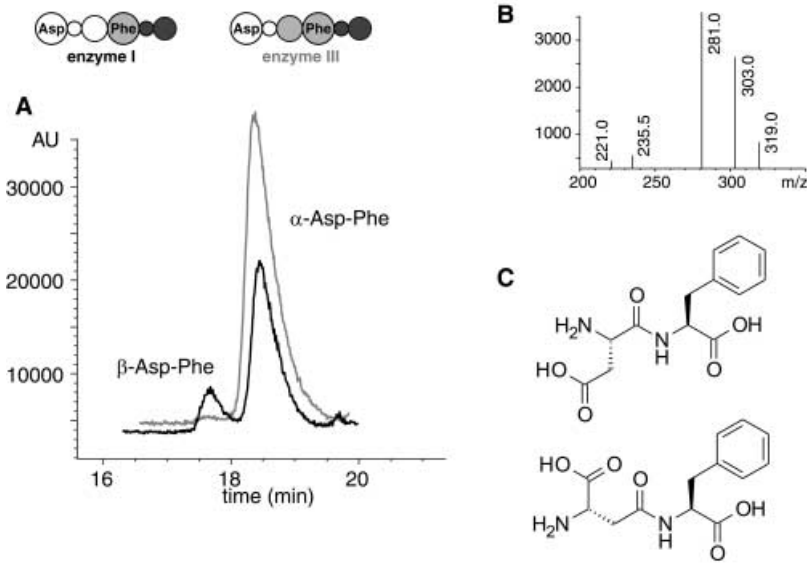

Fig. 4. HPLC/MS analysis of Asp-Phe formation. (A) Comparative HPLC/MS diagram of Asp-Phe produced by enzymes I (black) and enzyme III (gray) after 60 min; enzyme I produces α-Asp-Phe and β-Asp-Phe in a ratio of 83 : 17, whereas no β-Asp-Phe formation was detected for enzyme III. (B) Mass spectrum of the α-Asp-Phe product (calculated $[M + H]^+$-mass of Asp-Phe: 281.1). (C) Structures of α-Asp-Phe (top) and β-Asp-Phe (bottom).

Table 1. Catalytic activity of hybrid Asp-Phe synthetases.

Enzyme	Domain arrangement	Turnover (min^{-1})	α/β ratio
I	$[A\text{-}PCP\text{-}C]_{SrfB2}\text{-}[A]_{TycA}\text{-}[PCP\text{-}Te]_{SrfC}$	0.12	83 : 17
II	$[A\text{-}PCP]_{SrfB2}\text{-}[C\text{-}A]_{TycB2}\text{-}[PCP\text{-}Te]_{SrfC}$	0.27	92 : 8
III	$[A\text{-}PCP]_{SrfB2}\text{-}[C\text{-}A]_{TycB2}\text{-}[PCP\text{-}Te]_{TycC6}$	0.70	100 : 0
IV	$[A\text{-}PCP]_{SrfB2}\text{-}[C\text{-}A]_{TycB2}\text{-}[PCP]_{TycC6}$	0.01	81 : 19
V	$[A\text{-}PCP]_{SrfB2}\text{-}[C\text{-}A\text{-}PCP]_{TycB2}\text{-}[Te]_{TycC6}$	0.08	75 : 25
VI	$[A\text{-}PCP\text{-}C]_{SrfB2}\text{-}[A\text{-}PCP]_{TycB2}\text{-}[Te]_{TycC6}$	0.07	75 : 25

product are presented. In addition to the expected α-Asp-Phe a by-product was found and identified as the isomeric form, β-Asp-Phe (Fig. 4C). In Fig. 4 analysis of product formation is exemplified by the HPLC/MS diagram of the constructs I and III. The results for the other constructs comprising a Te domain differed in the turnover rate and the ratio of α-Asp-Phe to β-Asp-Phe (α/β-ratio). Turnover numbers between 0.08 min^{-1} and 0.7 min^{-1} were observed for constructs having the C-terminal Te domain (Table 1). In relation to the turnover number, the α/β ratio varies for the different constructs. As expected enzyme IV, which lacks the product releasing Te domain, showed marginal Asp-Phe formation. Only a negligible turnover number of less than 0.01 min^{-1} was obtained. Although the product formation is insignificant we were able to calculate a α/β ratio of 81 : 19. The lowest efficiency of enzymes containing a Te domain was found for enzyme VI. It showed a k_{cat} of 0.07 min^{-1} and an α/β ratio of 75 : 25. Similar data were determined for enzyme V (k_{cat} = 0.08 min^{-1}, α/β ratio 75 : 25). Higher turnover numbers were obtained for enzymes I (k_{cat} = 0.12 min^{-1}; α/β ratio 83 : 17) and II (k_{cat} = 0.27 min^{-1}; α/β ratio 92 : 8). Enzyme III achieved the highest turnover rate of 0.7 min^{-1}. Interestingly, when analyzing enzyme III no β-Asp-Phe formation was detectable (Fig. 4). Obviously, the production of the by-product is suppressed by the high turnover rate of this particular construct III.

Comparison of the amount of Asp-Phe generated by different recombinant proteins illustrates that enzyme III is more efficient than the others. Analyzing Asp-Phe formation after incubation for a prolonged period of time confirms the advantage of enzyme III. All enzymes showed time-dependent increase of Asp-Phe formation, but the productivity slowed down over time, but this effect is least pronounced for enzyme III. Most likely this loss of function is caused by instability of the enzymes during incubation because adding of additional substrate amino acids or ATP to the reaction mixture does not lead to changes in the productivity of the enzymes. Comparing the different constructs and their catalytic activities it can be concluded, that the formation of the β-form by-product is directly related to the turnover number. It appears from this correlation, that β-product formation is strongly reduced at high turnover rates (Fig. 5).

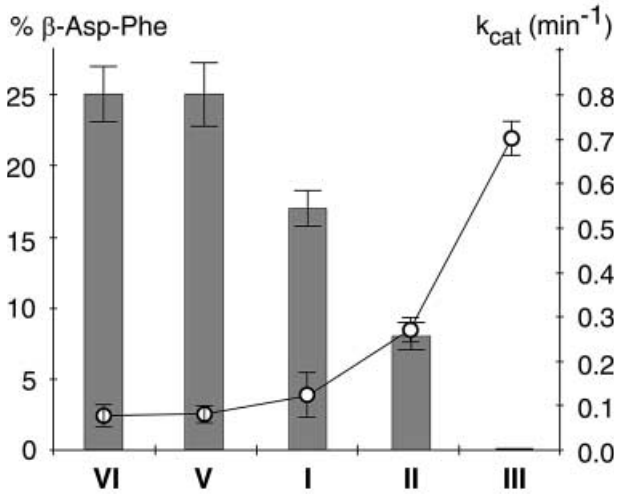

Fig. 5. Correlation of β-Asp-Phe formation and turnover. Product formation by the enzymes I, II, III and V is compared. The amount of produced β-Asp-Phe is illustrated with bars. Turnover numbers are shown as open circles (○).

Discussion

This study was to compare several recombinant Asp-Phe NRPSs created by different fusion strategies, to gain more information on important features for rational engineering of NRPSs. Earlier studies dealing with the development of hybrid NRPSs demonstrated that different strategies can be utilized for successful construction of new enzymes [17,18,20]. Modules were fused for example between A and PCP domain or PCP and C domain. The product release was obtained by fusion of Te domains or PCP-Te units. However these examinations were made using different systems. The advantage of this study is the comparison of various hybrid synthetases for production of the same peptide.

All engineered hybrid Asp-Phe synthetases showed product formation indicating that different strategies can be used for the construction of synthetases for production of novel peptides. The most important result is the effect of different fusion sites used for combining the respective domains. The individual catalytic activity of single domains is certainly not affected, but the results demonstrate that the arrangement of modules and domains has important effects on domain interactions.

In addition to the expected dipeptide α-Asp-Phe we observed the production of β-Asp-Phe. It has been reported that dipeptide products can be rapidly released from the enzyme template by formation of diketopiperazine (DKP), resulting in lower yields of expected product [20]. No such Asp-Phe-DKP-formation was observed in case of the Asp-Phe-hybrids, indicating that turnover rates are not affected by a nonenzymatic product release. The amount of β-Asp-Phe by-product formed, however, is directly correlated to the turnover rate of the enzymes. Different reasons for β-Asp-Phe formation are conceivable. For instance L-aspartic acid could be activated by the A domain in such a way that a β-L-aspartyl adenylate is formed. However, as the α/β ratio depends on the catalytic rate this explanation is not likely. The absence of β-Asp-Phe production by enzyme III particularly rebuts this explanation. Supposing the A domain activates L-aspartic acid as β-L-aspartyl adenylate at a basic level, the β-Asp-Phe amount should be independent of the turnover rate. Additionally, the Asp A domain of the SrfB2 module was shown to be highly specific for the

activation of aspartic acid. The structurally similar amino acids asparagine, glutamate and D-aspartic acid are not activated. Given this, and the known substrate recognition of A domains [6,31,32] direct activation of the β-carboxyl group is improbable.

Therefore, a more probable explanation for β-Asp-Phe formation is by conversion of enzyme-bound aspartic acid. We suggest a rearrangement mechanism that can explain the particular behavior of these hybrid synthetases (Fig. 6). According to this model α-aspartyl-*S*-Ppant covalently tethered to the PCP domain can be cleaved by an intramolecular nucleophilic attack of the β-carboxylate group on the thioester carbonyl function, forming the five-membered 2-amino-succinic anhydride. This anhydride can be attacked on its α- or β-position by the thiolate of the Ppant-moiety, resulting in α-L-aspartyl-*S*-Ppant and β-L-aspartyl-*S*-Ppant, respectively. These two isomeric aspartyl-thioesters are then processed into PCP-bound α- and β-Asp-Phe, which are subsequently released from the protein template, giving liberated α-Asp-Phe and β-Asp-Phe.

Our results show that increasing speed of product formation leads to reduction of α-aspartyl-*S*-Ppant isomerization. In terms of the mechanism of Fig. 6 this means if $k_4 \gg k_{1'}$ most of the product will be α-Asp-Phe. If, on the other hand, $k_4, k_5 \ll k_1, k_2, k_3$ the equilibrium between α-aspartyl-*S*-Ppant and β-aspartyl-*S*-Ppant will be fully established. Such a situation is most likely to occur for enzymes V and VI. Furthermore, if the condensation reaction and anhydride formation occur with nearly the same rate ($k_1, k_2, k_3 = k_4, k_5$) α/β ratios between 75 : 25 and 100 : 0 are expected. Rapid hydrolysis of 2-amino-succinic anhydride can explain the low thiolation rate ($k_6 > k_4, k_5$). The α to β-Asp rearrangement is a potential liability in any NRPS assembly line where an Asp-thioester would accumulate. Similar rearrangement mechanisms are also possible for other amino acids having nucleophilic sidechains, such as ornithine or glutamate.

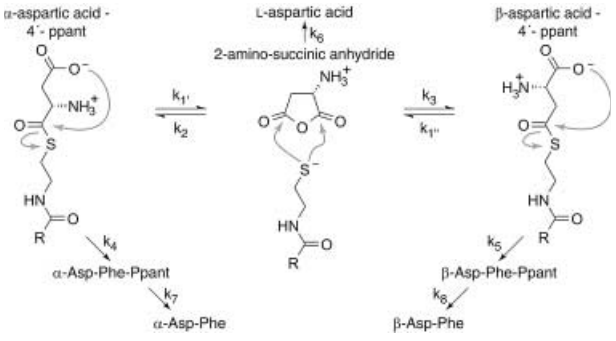

Fig. 6. Proposed mechanism of β-Asp-Phe formation. The synthesis of β-Asp-Phe is proposed to be a consequence of conversion of the enzyme-bound aspartic acid into a cyclic anhydride. Due to the higher intrinsic reactivity of the α-position in the anhydride the back reaction forming α-Asp-Ppant is preferred ($k_2 > k_3$). The proportion of k_1, k_2, k_3 to k_4, k_5 is responsible for the ratio of produced α- and β-form. Faster processing of the enzyme-bound aspartic acid leads to lower amounts of β-Asp-Phe. A slower condensation reaction (k_4) results in greater rearrangement of enzyme-bound aspartic acid. Condensation of the amino acids is also be affected by the velocity of the termination (k_7, k_8). Rapid hydrolysis of the anhydride ($k_6 > k_2, k_3$) may lead to lower thiolation rates.

As described above, different rates of product formation were observed for the recombinant proteins. The most important effect on enzyme activity was exerted by changes of Te fusions. The hybrid enzyme yielding the highest turnover number is enzyme III (0.7 min^{-1}) in with the second module contains the PCP of TycC6 attached to its cognate Te. Replacing these to domains by the domains PCP-Te of SrfC leads to a decrease in product formation (enzyme II; 0.27 min^{-1}), demonstrating the importance of an optimal choice of Te domain. Obviously, Asp-Phe release is more efficiently catalyzed by the tyrocidine Te. In addition the results demonstrate that the position of fusion is important. Enzyme V which contains the complete elongation module TycB2 and the Te domain fused without the native PCP forms Asp-Phe with a 10-fold reduced rate compared to enzyme III. In hybrid Asp-Phe synthetases Te domains cleave the Asp-Phe dipeptide with higher efficiency if they are connected to their cognate PCP domain. This result broadens the basic knowledge concerning the fusion of Te domains in hybrid NRPSs [20]. PCP domains naturally connected to Te domains (PCP^{Te}) exhibit only slight differences in their primary sequence compared to PCPs located at the N-terminus of a C domain (PCP^{C}). The entire sequence of PCP^{Te} domains shows about 30–50% identity to PCP^{C} domains, which is almost the same value as calculated for identities of different PCP^{C} domains to each other. As there are only slight changes in sequence between PCP^{C} and PCP^{Te}, it is probable that structural changes caused by the artificial domain combinations affected the activity. We showed that the intrinsic activities of single domains are independent of changes within the interdomain linker. However, it is still not known exactly how modifications within the interdomain linkers influence the communication of domains in a hybrid NRPS [17].

Increase of β-Asp-Phe formation is an indication for slow transmission of the activated amino acids to the next processing domain, indicating that the enzyme bound dipeptide is not cleaved off rapidly enough. Referring to the proposed mechanism (Fig. 6) it can be concluded that the condensation reaction (k_4) is slowed down if product release is the rate limiting step, resulting in the rearrangement of enzyme bound α-Asp-Phe. If k_4 is much higher than k_1, L-aspartic acid is processed into α-Asp-Phe only, as seen for enzyme III. Here the higher catalytic activity of the Te domain diminishes the extent to which termination limits overall turnover.

In case of enzyme I the Phe activating A domain of the native initiation module TycA was used and fused at its C-terminus to the domains A-PCP-C of SrfB2-3. This assembly probably limits the catalytic rate independently of the Te reaction. If so the rate limiting step in these constructs is the condensation reaction. It is known that the C domain possesses an acceptor site specificity [30,33]. On account of the native connection, the SrfB3 C domain, which is a component of the enzymes I and VI, exhibits an intrinsic substrate specificity for leucine and the processing of phenylalanine may be impeded. In this case an improvement of the Te domain reaction would not accelerate product formation.

This comparative study of hybrid Asp-Phe synthetases confirms the idea that the C and A domains of the same module represent a catalytic unit. Therefore fusions of NRPS modules should preferably be chosen between PCP and C domain in order to leave the C-A junction intact. However, the activity of the hybrid enzyme I demonstrate that in principle it is possible to use domains of native initiation modules for the construction of hybrid elongation modules utilizing fusions between C and A domain, even though the catalytic efficiency is lower compared to constructs with native elongation modules like enzymes II and III. On the other hand A-PCP units of native elongation modules can be used for initiation of hybrid synthetases, as it is realized in the described Asp-Phe synthetases. In summary several different strategies have proved successful for rational engineering of an NRPS-based dipeptide synthetase, but some proved better than others.

Acknowledgements

This work was supported by Holland Sweetener Company (Geleen, the Netherlands). We thank Robert Finking, Philipps-Universität Marburg, Germany and Marco Krämer, DSM Biotech GmbH, Jülich, Germany for discussion and critical comments on the manuscript. Finally, we thank Thei Verstappen, DSM Research, Geleen, the Netherlands for assistance in the HPLC/MS measurements.

References

1. Schwarzer, D. & Marahiel, M.A. (2001) Multimodular biocatalysts for natural product assembly. *Naturwissenschaften* **88**, 93–101.
2. Cane, D.E., Walsh, C.T. & Khosla, C. (1998) Harnessing the biosynthetic code: combinations, permutations, and mutations. *Science* **282**, 63–68.
3. Walsh, C.T., Chen, H., Keating, T.A., Hubbard, B.K., Losey, H.C., Luo, L., Marshall, C.G., Miller, D.A. & Patel, H.M. (2001) Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. *Curr. Opin. Chem. Biol.* **5**, 525–534.
4. Keating, T.A. & Walsh, C.T. (1999) Initiation, elongation, and termination strategies in polyketide and polypeptide antibiotic biosynthesis. *Curr. Opin. Chem. Biol.* **3**, 598–606.
5. Stein, T., Vater, J., Kruft, V., Otto, A., Wittmann-Liebold, B., Franke, P., Panico, M., McDowell, R. & Morris, H.R. (1996) The multiple carrier model of nonribosomal peptide biosynthesis at modular multienzymatic templates. *J. Biol. Chem.* **271**, 15428–15435.
6. Conti, E., Stachelhaus, T., Marahiel, M.A. & Brick, P. (1997) Structural basis for the activation of phenylalanine in the nonribosomal biosynthesis of gramicidin S. *EMBO J.* **16**, 4174–4183.
7. Stachelhaus, T., Huser, A. & Marahiel, M.A. (1996) Biochemical characterization of peptidyl carrier protein (PCP), the thiolation domain of multifunctional peptide synthetases. *Chem. Biol.* **3**, 913–921.
8. Lambalot, R.H., Gehring, A.M., Flugel, R.S., Zuber, P., LaCelle, M., Marahiel, M.A., Reid, R., Khosla, C. & Walsh, C.T. (1996) A new enzyme superfamily – the phosphopantetheinyl transferases. *Chem. Biol.* **3**, 923–936.
9. Mofid, M.R., Finking, R. & Marahiel, M.A. (2002) Recognition of hybrid PCP/ACP acyl carrier proteins in nonribosomal peptide synthetase modules by the 4′-phophopantetheinyl transferases AcpS and Sfp. *J. Biol. Chem.* **19**, 17023–17031.
10. Bergendahl, V., Linne, U. & Marahiel, M.A. (2002) Mutational analysis of the C-domain in nonribosomal peptide synthesis. *Eur. J. Biochem.* **269**, 620–629.
11. Luo, L., Kohli, R.M., Onishi, M., Linne, U., Marahiel, M.A. & Walsh, C.T. (2002) Timing of epimerization and condensation reactions in nonribosomal peptide assembly lines: kinetic analysis of phenylalanine activating elongation modules of tyrocidine synthetase B. *Biochemistry* **41**, 9184–9196.
12. Miller, D.A. & Walsh, C.T. (2001) Yersiniabactin synthetase: probing the recognition of carrier protein domains by the catalytic heterocyclization domains, Cy1 and Cy2, in the chain-initiating HWMP2 subunit. *Biochemistry* **40**, 5313–5321.
13. Haese, A., Schubert, M., Herrmann, M. & Zocher, R. (1993) Molecular characterization of the enniatin synthetase gene encoding a multifunctional enzyme catalysing N-methyldepsipeptide formation in Fusarium scirpi. *Mol. Microbiol.* **7**, 905–914.
14. Schauwecker, F., Pfennig, F., Grammel, N. & Keller, U. (2000) Construction and *in vitro* analysis of a new bi-modular polypeptide synthetase for synthesis of *N*-methylated acyl peptides. *Chem. Biol.* **7**, 287–297.
15. Trauger, J.W., Kohli, R.M., Mootz, H.D., Marahiel, M.A. & Walsh, C.T. (2000) Peptide cyclization catalysed by the thioesterase domain of tyrocidine synthetase. *Nature* **407**, 215–218.
16. Tseng, C.C., Bruner, S.D., Kohli, R.M., Marahiel, M.A., Walsh, C.T. & Sieber, S.A. (2002) Characterization of the surfactin synthetase C-terminal thioesterase domain as a cyclic depsipeptide synthase. *Biochemistry* **41**, 13350–13359.
17. Mootz, H.D., Schwarzer, D. & Marahiel, M.A. (2000) Construction of hybrid peptide synthetases by module and domain fusions. *Proc. Natl Acad. Sci. USA* **97**, 5848–5853.
18. Doekel, S. & Marahiel, M.A. (2000) Dipeptide formation on engineered hybrid peptide synthetases. *Chem. Biol.* **7**, 373–384.
19. Linne, U. & Marahiel, M.A. (2000) Control of directionality in nonribosomal peptide synthesis: role of the condensation domain in preventing misinitiation and timing of epimerization. *Biochemistry* **39**, 10439–10447.
20. Schwarzer, D., Mootz, H.D. & Marahiel, M.A. (2001) Exploring the impact of different thioesterase domains for the design of hybrid peptide synthetases. *Chem. Biol.* **8**, 997–1010.
21. Irino, S., Oyama, K., Nakamura, S.-I., Van Dooren, T.J. & Quaedflieg, P.J.L. (1994) Enzymatic coupling of L-phenylalanine methyl ester and *N*-benzyloxycarbonyl-aspartic acid. European Patent, EP0664338.
22. Hill, J.B. & Gelman, Y. (1990) Process for the preparation of α- L-aspartyl-L-phenylalanine methyl ester hydrochloride by use of isolated *N*-formyl L-apartic anhydride. US Patent, US-A-4,946,988.
23. Stachelhaus, T., Mootz, H.D., Bergendahl, V. & Marahiel, M.A. (1998) Peptide bond formation in nonribosomal peptide biosynthesis. Catalytic role of the condensation domain. *J. Biol. Chem.* **273**, 22773–22781.
24. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
25. Nakano, M.M., Corbell, N., Besson, J. & Zuber, P. (1992) Isolation and characterization of sfp: a gene that functions in the production of the lipopeptide biosurfactant, surfactin. *Bacillus subtilis*. *Mol. Gen. Genet.* **232**, 313–321.
26. Mootz, H.D. & Marahiel, M.A. (1997) The tyrocidine biosynthesis operon of *Bacillus brevis*: complete nucleotide sequence and biochemical characterization of functional internal adenylation domains. *J. Bacteriol.* **179**, 6843–6850.
27. Dieckmann, R., Pavela-Vrancic, M., von Dohren, H. & Kleinkauf, H. (1999) Probing the domain structure and ligand-induced conformational changes by limited proteolysis of tyrocidine synthetase 1. *J. Mol. Biol.* **288**, 129–140.

28. Gokhale, R.S., Lau, J., Cane, D.E. & Khosla, C. (1998) Functional orientation of the acyltransferase domain in a module of the erythromycin polyketide synthase. *Biochemistry* **37**, 2524–2528.
29. Weber, T., Baumgartner, R., Renner, C., Marahiel, M.A. & Holak, T.A. (2000) Solution structure of PCP, a prototype for the peptidyl carrier domains of modular peptide synthetases. *Structure Fold. Des.* **8**, 407–418.
30. Belshaw, P.J., Walsh, C.T. & Stachelhaus, T. (1999) Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis. *Science* **284**, 486–489.
31. May, J.J., Kessler, N., Marahiel, M.A. & Stubbs, M.T. (2002) Crystal structure of DhbE, an archetype for aryl acid activating domains of modular nonribosomal peptide synthetases. *Proc. Natl Acad. Sci. USA* **99**, 12120–12125.
32. Stachelhaus, T., Mootz, H.D. & Marahiel, M.A. (1999) The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol.* **6**, 493–505.
33. Ehmann, D.E., Trauger, J.W., Stachelhaus, T. & Walsh, C.T. (2000) Aminoacyl-SNACs as small-molecule substrates for the condensation domains of nonribosomal peptide synthetases. *Chem. Biol.* **7**, 765–772.