Dipeptide formation on engineered hybrid peptide synthetases

Sascha Doekel and Mohamed A Marahiel

Background: Nonribosomal peptide synthetases (NRPSs) are modular 'megaenzymes' that catalyze the assembly of a large number of bioactive peptides using the multiple carrier thiotemplate mechanism. The modules comprise specific domains that act as distinct units to catalyze specific reactions associated with substrate activation, modification and condensation. Such an arrangement of biosynthetic templates has evoked interest in engineering novel NRPSs.

Results: We describe the design and construction of a set of dimodular hybrid NRPSs. By introducing domain fusions between adenylation and thiolation (PCP) domains we designed synthetic templates for dipeptide formation. The predicted dipeptides, as defined by the specificity and arrangement of the adenylation domains of the constructed templates, were synthesized *in vitro*. The effect of the intramolecular fusion was investigated by determining kinetic parameters for substrate adenylation and thiolation. The rate of dipeptide formation on the artificial NRPSs is similar to that of natural templates.

Conclusions: Several new aspects concerning the tolerance of NRPSs to domain swaps can be deduced. By choosing the fusion site in the border region of adenylation and PCP domains we showed that the PCP domain exhibits no general substrate selectivity. There was no suggestion that selectivity of the condensation reaction was biased towards the donor amino acid, whereas at the acceptor position there was a size-determined selection. In addition, we demonstrated that a native elongation module can be converted to an initiation module for peptide-bond formation. These results represent the first example of rational *de novo* synthesis of small peptides on engineered NRPSs.

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

Correspondence: Mohamed A Marahiel
E-mail: marahiel@chemie.uni-marburg.de

Key words: dipeptide formation, hybrid enzymes, nonribosomal peptide synthetase

Received: **2 February 2000**
Revisions requested: **28 February 2000**
Revisions received: **6 March 2000**
Accepted: **15 March 2000**

Published: **15 May 2000**

Chemistry & Biology 2000, **7**:373–384

1074-5521/00/$ – see front matter
© 2000 Elsevier Science Ltd. All rights reserved.

Introduction

Nonribosomal peptide synthesis is an alternative route of peptide synthesis carried out by large microbial multifunctional enzymes termed nonribosomal peptide synthetases (NRPSs) [1,2]. Molecular characterization of NRPS genes has revealed a modular organization in which each module includes a full complement of catalytic sites (domains) required for a single step of chain elongation [3]. Size and sequence of the nonribosomal peptide are determined by the number and colinear arrangement of modules. Peptides are synthesized through amino acyl adenylate intermediates (recognized and activated by the adenylation or A domain) [4] that are subsequently tethered to a 4′-phosphopantetheinyl cofactor of the adjacent thiolation (T) domain (or peptidyl-carrier protein, PCP) [5]. Thioesterified substrates are concatenated via peptide-bond formation in a stepwise amino→carboxy-terminal elongation reaction catalyzed by the condensation (C) domain, the third essential domain of a minimal module [6]. Structural diversity in this family of bioactive natural products can be further enhanced by substrate modifications such as epimerization [7], N-methylation [8] or heterocyclization [9]. These modifications are catalyzed by auxiliary domains. Release of the nascent peptide is accomplished by a thioesterase (Te)-like domain, found at the carboxyl terminus of most NRPSs [10]. The primary sequence, configuration of each α carbon center and the extent of modification of these products are therefore controlled by the linear sequence of catalytic domains [11].

PCP domains are activated by post-translational transfer of a 4′-phosphopantetheine cofactor to a conserved serine residue, a reaction catalyzed by specialized coenzyme A (CoA)-dependent 4′-phosphopantetheinyl transferases [12,13].

Many nonribosomal peptides have important agricultural or medical uses [14–16]. Examining the catalytic flexibility of NRPSs towards noncognate substrates is of interest because these results could pave the way to new drugs. Insight into the rules that govern the interaction of the multidomain arrangements should allow the rational design of novel peptides by module and domain swapping. Some attempts have been made to re-design NRPSs *in vivo* [17,18]. Biochemical characterization of distinct domain types *in vitro* should provide a deeper mechanistic understanding of NRPSs [5,19,20]. Designing a model NRPS system to evaluate the feasibility of such manipulations should test our understanding of NRPS architecture.

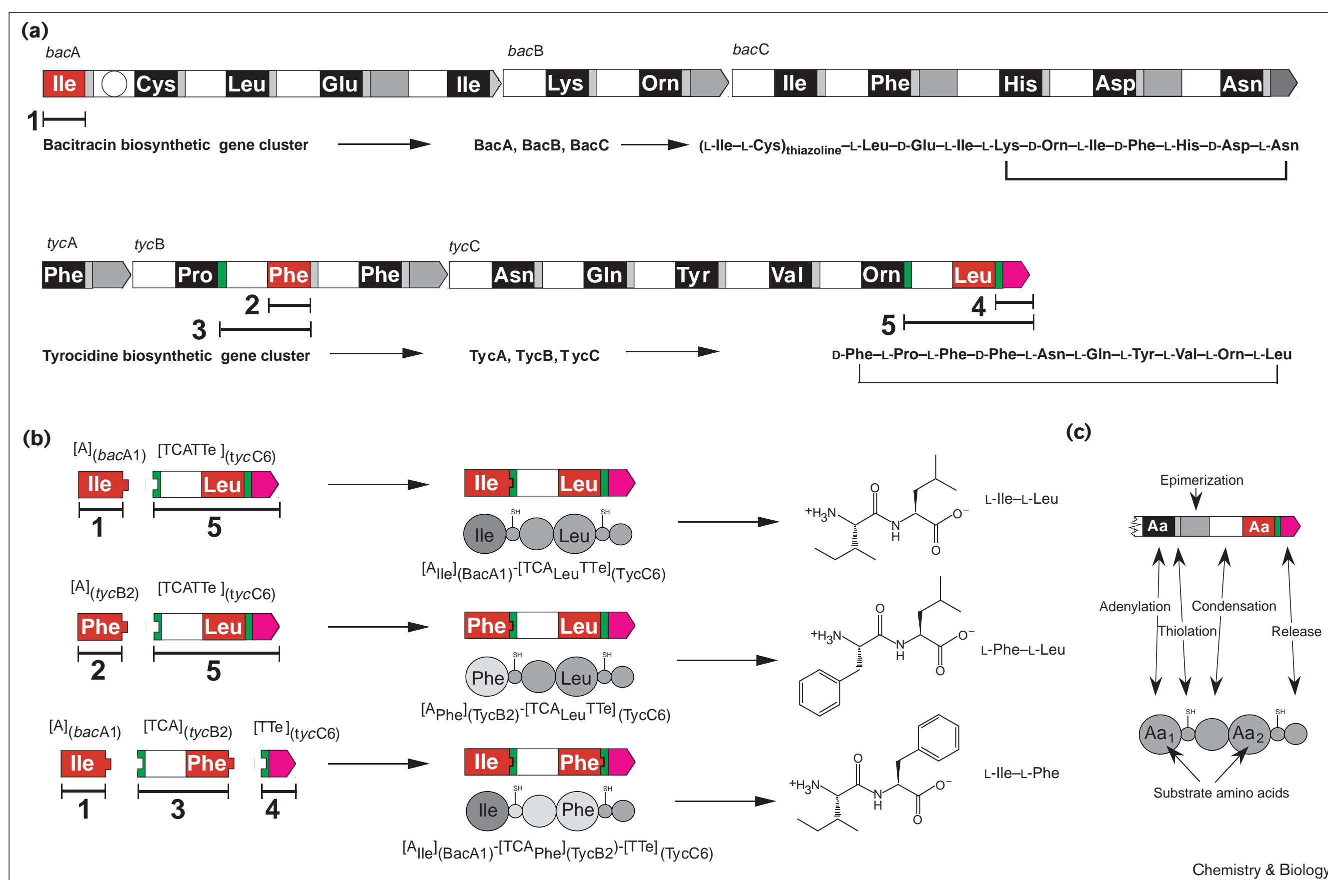
In this paper we examine how NRPSs can be used to design novel artificial templates to produce small peptides *de novo*. To this end, a set of dimodular hybrid NRPSs was constructed on the basis of fusing A and PCP domains. These hybrid enzymes were shown to catalyze specific adenylation and thiolation reactions of substrate amino acids, a condensation reaction and a terminating release that yielded dipeptides of the predicted sequence. Biochemical studies investigating individual steps in dipeptide formation revealed that the domain types (apart from A domains) are quite tolerant with respect to processing non-native substrates. This approach should encourage the rational design of small peptides on engineered NRPS templates [21,22].

Results and discussion

Engineering dimodular hybrid NRPSs

We designed and constructed three dimodular hybrid NRPSs (Figure 1b). These constructs serve as a model system for addressing important questions about nonribosomal peptide synthesis. The main determinant of NRPS substrate specificity is the A domain, which selectively recognizes and adenylates a cognate acyl amino acid [23]. In contrast, little is known about the substrate selectivity of downstream domains such as the PCP and C domains and whether (and to what extent) domain organization can be exploited to obtain A-domain-swapped NRPSs that form peptides with the predicted sequences. To answer these questions we started with a dimodular NRPS with a native termination module (CATTe) and introduced a domain fusion between the A and PCP domains of module 1 (A-T; Figure 1b; see the Materials and methods section for module definitions). This operation yielded a hybrid template $[A_{Ile}]_{(BacA1)}$–$[TCA_{Leu}TTe]_{(TycC6)}$ (enzyme **I**; Table 1) that connects the native initiation module of the bacitracin synthetase 1 BacA1 [9] with the carboxy-terminal module of the tyrocidine synthetase TycC [19] (for details of

Figure 1



(a) Organization of the *bac* operon of *Bacillus licheniformis* ATCC10716 (top) and the *tyc* operon of *Bacillus brevis* ATCC8185 (bottom). Domain organization is depicted by boxes in different shadings. Three genes, *bac*A, *bac*B and *bac*C, encode the synthetases BacA, BacB and BacC that assemble the branched cyclic dodecapeptide bacitracin. Tyrocidine is synthesized by three peptide synthetases TycA, TycB and TycC, encoded by *tyc*A, *tyc*B and *tyc*C. The domains used for the engineering of hybrid NRPSs are highlighted in color. **(b)** Engineering of dimodular hybrid NRPSs by domain fusion. Left: gene fragments corresponding to NRPS domains were amplified. Middle: fusion leads to artificial hybrid genes and their corresponding hybrid dimodular NRPSs. Right: potential dipeptides formed by the hybrid templates. **(c)** Color/shading code of NRPS domain types. A circle in the condensation domain of BacA2 indicates a cyclization domain.

cloning, see the Materials and methods section). We predicted that the dipeptide Ile–Leu would be formed on the basis of specificity of the A domains used (Figure 1a). A second construct $[A_{Phe}]_{(TycB2)}$–$[TCA_{Leu}TTe]_{(TycC6)}$ (enzyme **II**; Table 1) was designed in an analogous manner (Figure 1b). We also wanted to address the question of whether a native elongation module can be transformed into an initiation module. Here the native elongation module TycB2 was repositioned to serve as an initiation module. Phe–Leu was predicted to be the product on the basis of the arrangement and specificity of the A domains. In a third construct (Figure 1b), a second fusion site was introduced between the A and PCP domains of module 2. The resulting hybrid $[A_{Ile}]_{(BacA1)}$–$[TCA_{Phe}]_{(TycB2)}$–$[TTe]_{(TycC6)}$ (enzyme **III**; Table 1), a template for the production of the dipeptide Ile–Phe, converts a native elongation module into a termination module. The latter hybrid construct is a model for an engineered nonribosomal template for any dipeptide A–B that can be generated simply by fusing two modules with the specificities for A and B, respectively.

We engineered the hybrid sites within the linker regions of the A and PCP domains. To define the optimal fusion site, we aligned the potential boundary region between the A and PCP domains, and identified a relatively unconserved stretch of nine amino acids embedded within highly conserved residues (Figure 2). The same site has been previously shown to be susceptible to proteolytic digestion [24]. Moreover, the recently solved solution structure of the PCP domain of TycC3 [25] also supports this definition of a linker region, because the secondary fold of the PCP domain starts from a highly conserved tyrosine next to this linker region (Figure 2). We identified no sequence conservation among the linker regions. It has been reported, however, that linker regions have a propensity to form defined secondary folds with a conformationally sensitive turn [26]. This turn is often represented by a highly conserved proline residue that acts as a switch to coordinate or de-coordinate catalytic domains. Single residues among linker regions have, therefore,

Table 1

Hybrid dimodular NRPSs made in this study.

	Construct	Modules
I	$[A_{Ile}]_{(BacA1)}$–$[TCA_{Leu}TTe]_{(TycC6)}$	Ile, Leu
II	$[A_{Phe}]_{(TycB2)}$–$[TCA_{Leu}TTe]_{(TycC6)}$	Phe, Leu
III	$[A_{Ile}]_{(BacA1)}$–$[TCA_{Phe}]_{(TycB2)}$–$[TTe]_{(TycC6)}$	Ile, Phe

Domains used as segments are indicated with square brackets, and their names related to the genes from which they have been extracted (e.g. Tyc6). We include the proposed amino acid substrate of the A domain (e.g. Ile). Hybrid fusion sites are indicated with a dash.

been shown to be invariant for the function of multidomain proteins [27].

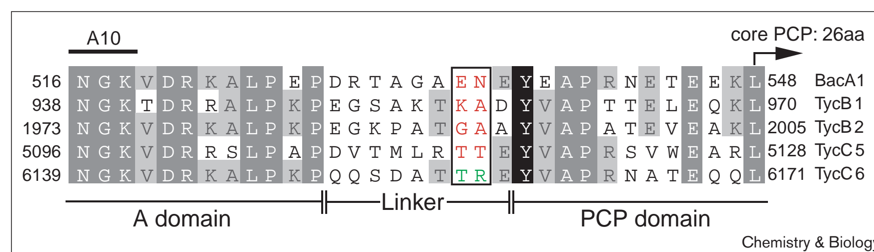
Production of the hybrid proteins

Hybrid enzymes (with a predicted mass of 215,000 Da) were overproduced in *Escherichia coli* BL21(p*gsp*) as (His_6)-tagged fusion proteins and purified using a Ni-agarose column. Typically 6 mg of protein was obtained from 1 l culture (see the Supplementary material section). The amount of **III** was considerably lower than that of **I** and **II**. Overproduction in *E. coli* BL21(p*gsp*) yielded holoenzymes, resulting from co-expression with the 4′-phosphopantetheinyl transferase gene *gsp* [28].

Adenylation reaction

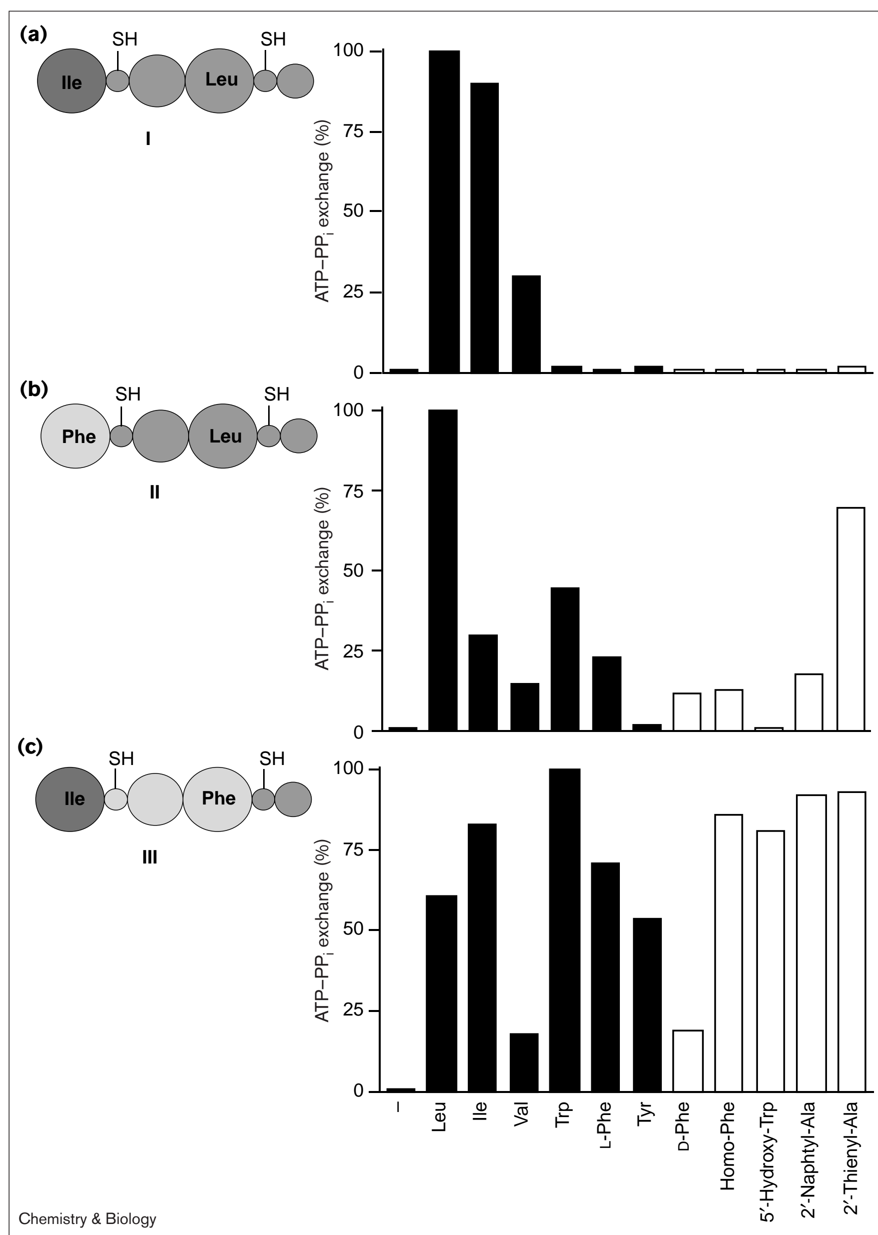
To test the substrate specificity of the engineered NRPS, ATP–PP_i exchange reactions were performed [6]. As shown in Figure 3a, recombinant enzyme **I** activated the cognate amino acids isoleucine and leucine to the same level (compared with valine, which was only activated to 30% the level of isoleucine or leucine; Figure 3a). The A domain of BacA1 was shown to activate only isoleucine [9]. The level of isoleucine activation on the second module TycC6 is anticipated to be ~30% (see the activation pattern of **II**).

Figure 2



Amino acid sequence alignment (using ClustalW and single-letter code) of the potential boundary region of the A and PCP domains used in this study. A poorly conserved stretch of nine amino acids was identified. This region is believed to represent a linker connecting the A and PCP domains. Boxed, the fusion site that alters the native sequence to LQ (introduction of a *Pst*I site, in red) or VN (introduction of a *Hpa*I site, in green), respectively. Numbers at the left and right indicate the position within the polypeptide chain. A10 is a highly conserved motif of A domains.

Figure 3



Relative representation of ATP–PP_i exchange reaction with the purified hybrid NRPSs. **(a)** I, **(b)** II and **(c)** III. The highest exchange rate was defined as 100%. Maximum labeling was 300,000 cpm in all three panels, and the highest activation rates reached ~110,000–130,000 cpm each. Black columns represent proteinogenic amino acids, white columns nonproteinogenic amino acids.

Enzyme **II** activates tryptophan and phenylalanine (by TycB2), as well as leucine and isoleucine (mainly by TycC6; Figure 3b). The activation of the donor amino acids tryptophan and phenylalanine was reduced compared with that of the acceptor amino acids. Activation of tryptophan by the A domain of TycB2 was predicted previously [19] and explains the occurrence of tyrocidine analogs that have a Phe→Trp substitution in the third position (Figure 1a).

As shown in Figure 3c, enzyme **III** also activated the four amino acids isoleucine and leucine (mainly by BacA1) and tryptophan and phenylalanine (by TycB2). In addition, considerable activation of tyrosine was observed. The activation of tryptophan and phenylalanine by the TycB2 module was enhanced compared with **II**. Also remarkable was the observed high level of tyrosine activation in **III** in contrast to the low activation by **II**. Both hybrid enzymes contain the same module TycB2 (Figure 3b,c). This finding prompted us to investigate the substrate specificity of **III** and **II** in more detail. We found that the activation patterns were clearly different for the nonproteinogenic amino acids 5′-Hydroxy-Trp, Homo-Phe and 2′-Napthyl-Ala (Figure 3b,c). In general, **III** was less specific than **II**. The differences in substrate specificity are surprising because it has been proposed

Table 2

Substrate affinity of the TycB2 A domain.

Enzyme	Substrate amino acid	K_m (mM)	K_{cat}/K_m (mM^{-1} s^{-1})
II	Tryptophan	0.04	7.2×10^7
	Homo-Phe	0.4	3.9×10^6
	2′-Thienyl-Ala	0.8	4.5×10^7
III	Tryptophan	<0.01	$>10^8$
	Homo-Phe	0.01	6.9×10^8
	2′-Thienyl-Ala	1.2	1.4×10^7

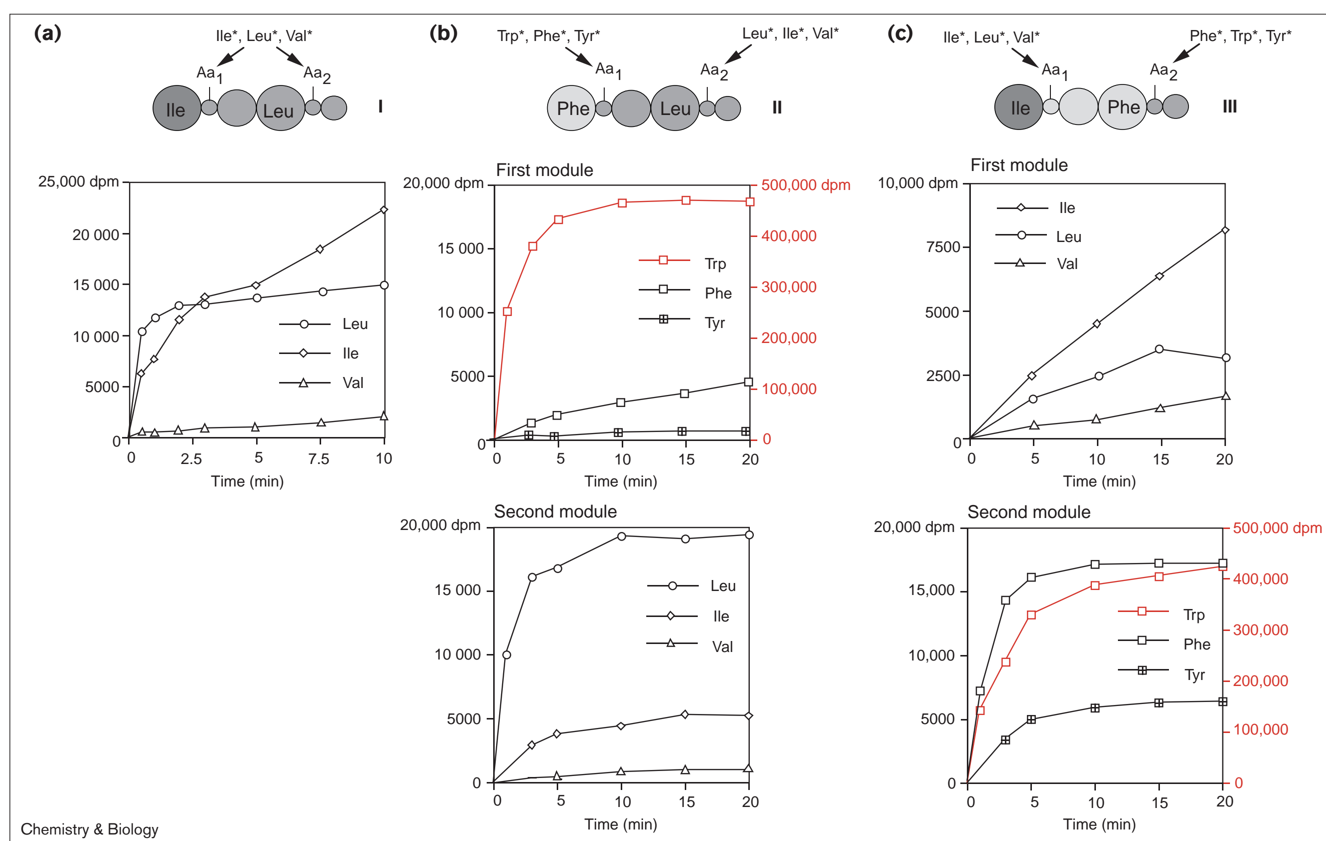
that A domains are autonomous units that retain their catalytic potential even when overproduced as distinct proteins [19]. Our results, however, provide evidence that the exact activation pattern of an A domain is a function of an active-site configuration involving adjacent domain interaction [29]. The importance of conformational changes in the process of adenylation and substrate release has been discussed recently [24]. The kinetic parameters for the adenylation reaction demonstrate that **II** and **III** have different K_m and k_{cat}/K_m values for tryptophan and homo-Phe, whereas the values for 2′-Thienyl-Ala were comparable for both proteins (Table 2).

Substrate thiolation

To monitor transfer of the adenylated amino acids to the PCP domain, the trichloracetic acid (TCA)-precipitation assay [6] using [^{14}C]- or [^{3}H]-labeled amino acids was employed. The formation of thioesters is a crucial step in the investigation of the hybrid NRPSs used in this study, because this step directly involves substrate transfer between domains of different origin. Only a proper interaction between the A and PCP domains ensures significant loading of substrate amino acids [29]. Previous heterologous loading experiments *in trans* with dissected A and PCP domains were obviously hampered by poor domain–domain interactions, which revealed an impaired substrate thiolation [30].

As shown in Figure 4, in the case of the three investigated proteins **I**, **II** and **III**, a significant level of substrate transfer to the heterologous PCP domain was feasible.

Figure 4



Substrate thiolation on PCP domains monitored by trichloroacetic acid (TCA)-precipitation assay. The schematic presentation of the hybrid NRPS shows which substrate amino acids are expected to be incorporated by the first and second modules, respectively (deduced from the specificity of adenylation domains). Kinetics for the thiolation of substrate amino acids in **(a) I**, **(b) II** and **(c) III**.

Table 3

Dipeptides formed by the hybrid NRPSs used in this study.

Enzyme	Product*	Detection	Velocity†
I	Ile–Leu	Yes	0.3‡
	Ile–Ile	Yes	0.1
	Leu–Leu	Yes	0.3
	Leu–Ile	–	
II	Trp–Leu	Yes	2.1
	Phe–Leu	Yes	0.9
	Trp–Ile	Yes	0.1#
	Phe–Ile	–§	
	2′-Thienyl-Ala–Leu¶	Yes	n.d.
	2′-Napthyl-Ala–Leu¶	Yes	n.d.
	5′-Hydroxy-Trp–Leu¶	Yes	n.d.
	Homo-Phe–Leu¶	Yes	n.d.
	D-Phe–Leu	–	
III	Ile–Phe	Yes	0.5
	Ile–Trp	Yes	<0.02
	Leu–Phe	Yes	<0.02
	Leu–Trp	Yes	n.d.
	Ile–2′-Thienyl-Ala¥	Yes	n.d.
	Ile–2′-Napthyl-Ala¥	Yes	n.d.
	Ile–5′-Hydroxy-Trp¥	Yes	n.d.
	Ile–Homo-Phe¥	Yes	n.d.
	Ile–D-Phe¥	–	

'Yes' indicates that the dipeptide product was observed using HPLC MS and radioactive TLC. *If not indicated otherwise, then both amino acids are in the L-configuration. †In mol per min and mol of enzyme. n.d., not determined. ‡In parallel Ile–Ile but not Leu–Leu is formed. #Dipeptide products resulting from an inspecific enzyme reaction are also observed: Ile–Ile, Ile–Trp and Trp–Trp. §Trace amounts of Ile–Ile are formed. ¶Analogous dipeptides with Ile in the acceptor position cannot be formed. ¥Analogous dipeptides with Leu in the donor position cannot be formed.

Analysis of protein **I** revealed fast incorporation of leucine and delayed incorporation of isoleucine (Figure 4a). Moreover, the kinetics of isoleucine incorporation might be a result of loading on both modules at the same time. Incorporation of valine (whose adenylation rate was 30% of the isoleucine and leucine rates) was very slow (Figure 4a) and may be caused by an uncatalyzed thioester formation.

On the second module of **II** the incorporation rate of leucine was fast compared with isoleucine and valine (see Figure 4b, lower panel). The ratio of leucine, isoleucine and valine incorporation reflects the ATP–PP_i activation pattern for these amino acids. Substrate thiolation on the first module was fast with tryptophan (Figure 4b, upper panel), whereas the incorporation of phenylalanine was slow and not completed within 20 min of incubation. The low incorporation of tyrosine (see Figure 4b) again may result from an uncatalyzed transfer. In both proteins **I** and **II**, transfer of the noncognate amino acids isoleucine and tryptophan to a hybrid PCP domain that normally serves as a carrier for ornithine (TycC5; Figure 1a) was feasible.

In the dimodular construct **III**, PCP domains in both modules represent hybrid junctions: a hydrophobic amino acid (isoleucine) is to be loaded on a PCP domain naturally engaged in proline processing (first module) and a large aromatic amino acid (tryptophan) is to be loaded on a PCP domain originally associated with a leucine-activating module (second module; Figure 1a). As shown in Figure 4c, the incorporation rates of putative donor amino acids such as isoleucine, leucine and valine were slow and not completed within 20 minutes (upper panel). This slow incorporation on the first module had been observed with protein **I** (Figure 4a). Incorporation of the acceptor amino acids tryptophan, phenylalanine and tyrosine was relatively fast (Figure 4c), but tryptophan and phenylalanine reached high levels of incorporation, those of tyrosine stayed at a lower but significant level. The difference in loading of aromatic amino acids in **III** and **II** can be attributed to the altered substrate affinity of the TycB2 A domain.

Using intramolecular fusions of A and PCP domains we have been able to demonstrate a proper thiolation reaction. The rules that govern domain–domain interactions between A and PCP domains are probably conserved irrespective of the cognate substrate. A fixed arrangement and proximity of the two domains, as provided by a short linker region, seems to be sufficient for productive communication [31]. We found no evidence to suggest that substrate specificity is influenced at the level of thioester formation.

Product formation

Peptide formation on dimodular NRPSs implies an elongation reaction catalyzed by a C domain and subsequent product release catalyzed by a Te domain. The essential role of the C domain for peptide-bond formation has been demonstrated previously using a truncated model system [6]. The role of the Te domain in releasing nonribosomal peptides from the enzyme has been underlined by *in vivo* studies [18]. Te domains may also catalyze cyclization or multimerization of products [32]. A thioesterase with inherent substrate specificity was a topic of investigation for the Te domain in the erythromycin-producing polyketide synthase [33]. It has been shown recently that the thioesterase domain of NRPSs is sufficient for product release from artificial templates *in vitro* [34].

There is little data available about the influence of the C domain on substrate specificity [35]. Recent studies suggest the C domain exerts more influence on the nature of the acceptor amino acid than the donor amino acid [20].

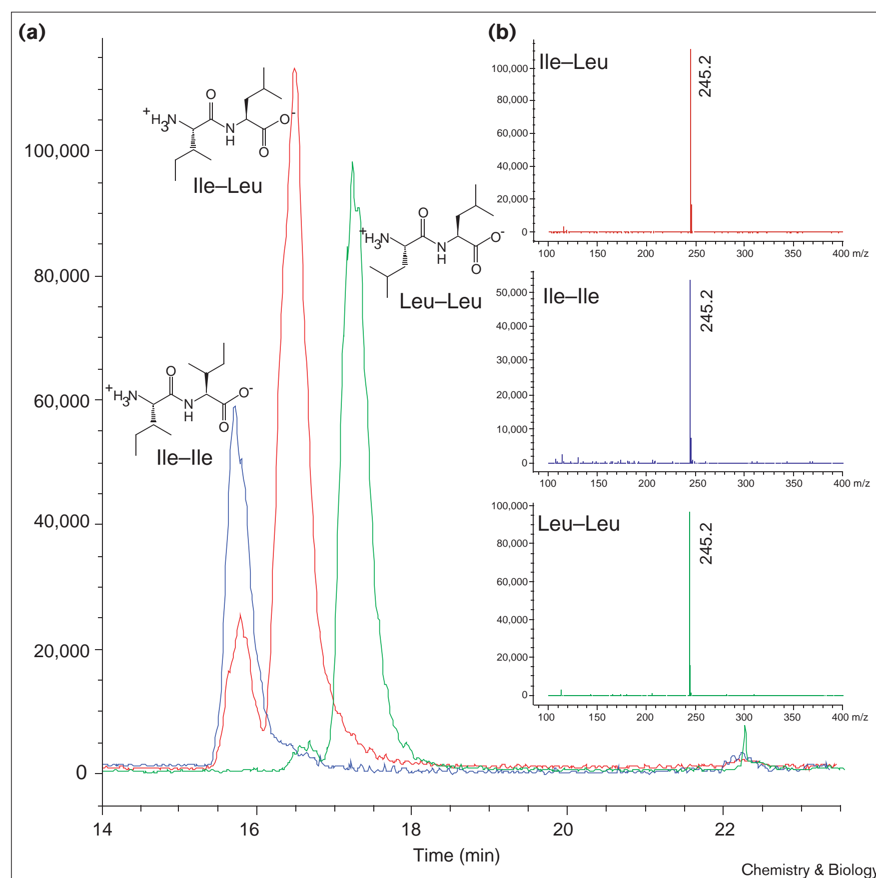
The dipeptide products of the three hybrid NRPSs prepared in this study were investigated. The structure of the predicted dipeptides should be dependant on the order and specificity of the two modules used. For product detection we used thin-layer chromatography (TLC) with

radiolabeled substrate amino acids (not shown), reverse-phase high-performance liquid chromatography (HPLC) and coupled HPLC-MS methods. Relevant compounds were compared with standard dipeptides. The results of dipeptide formation are summarized in Table 3.

In the presence of both substrate amino acids, isoleucine and leucine, **I** produced the dipeptides Ile–Leu and Ile–Ile, but no Leu–Leu or Leu–Ile were formed (Figure 5). The ratio between Ile–Leu and Ile–Ile was about 5:1. When the enzyme was incubated with just leucine, Leu–Leu was detected. Larger amounts of Ile–Ile were produced in the absence of leucine. As predicted from the substrate specificity of the first and second modules, Ile–Leu is the major product. The parallel formation of Ile–Ile may reflect the reduced isoleucine activation of the second module. Surprising is the high level of Leu–Leu formed in the absence of isoleucine — when isoleucine is present, no Leu–Leu is formed. The processing of leucine is probably affected when the more favored donor amino acid isoleucine is present. Regarding the absence of Leu–Leu formation and the reduced level of Ile–Ile in the presence of both substrate amino acids isoleucine and leucine, it is clear that the dipeptide Leu–Ile is not formed. Kinetic parameters for dipeptide formation revealed a turnover of 0.3 min^{-1} for Ile–Leu and Leu–Leu, in contrast to Ile–Ile, which was formed with a turnover of 0.1 min^{-1} (Table 3).

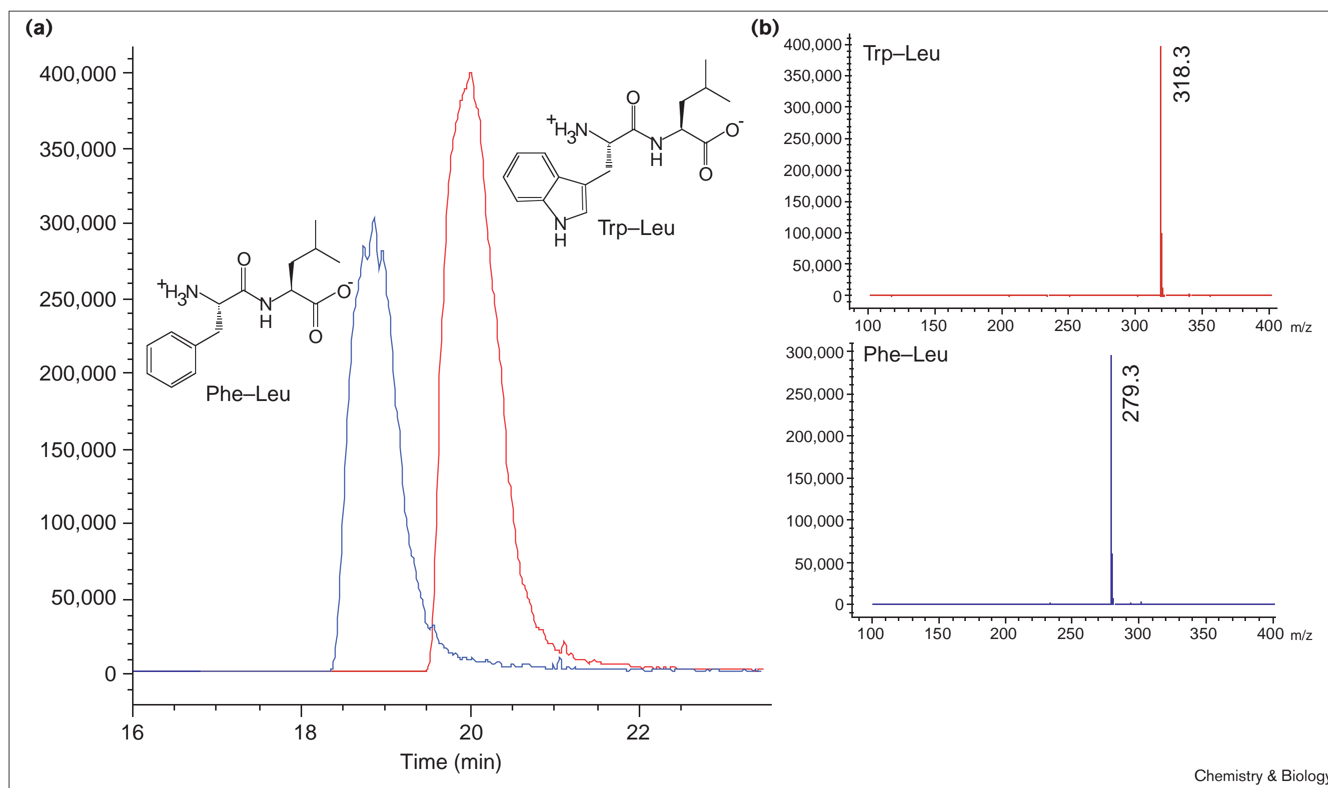
Hybrid enzyme **II** was shown to produce the dipeptides Trp–Leu, Phe–Leu and Trp–Ile when incubated with the relevant substrate amino acids (Figure 6). Phe–Ile was not detected. The formation of Trp–Leu was preferred and reached a turnover rate of 2.1 min^{-1}. Substitution of the donor amino acid tryptophan with phenylalanine and the acceptor amino acid leucine with isoleucine reduced the turnover rates to 0.9 min^{-1} and 0.1 min^{-1}, respectively (see Table 3). The reduced level of Phe–Leu production is consistent with the lower adenylation and thiolation rates for phenylalanine compared with tryptophan. When the enzyme was incubated with tryptophan and isoleucine, other dipeptides resulting from unspecific reactions such as Ile–Ile, Trp–Trp and Ile–Trp were observed to a smaller degree than Trp–Ile (not shown). This might be explained in the following way: when elongation of the thioesterified amino acid precursors runs too slowly, competitive capturing of other amino acids in solution becomes more prominent [36]. A similar observation was made when phenylalanine and isoleucine were incubated: instead of

Figure 5



Dipeptides formed by the hybrid enzyme **I** detected by reverse-phase HPLC MS analysis. **(a)** HPLC MS diagrams: in the presence of substrate amino acids isoleucine and leucine (red) the dipeptides Ile–Leu and (to a lesser extent) Ile–Ile are formed. Incubation with isoleucine (blue) or leucine (green) leads to formation of Ile–Ile and Leu–Leu, respectively. **(b)** Mass spectra of detected products: Ile–Ile, Ile–Leu and Leu–Leu.

Figure 6



Dipeptides formed by the hybrid enzyme **II** detected by reverse-phase HPLC MS analysis. **(a)** HPLC MS diagrams of Phe–Leu (blue) and Trp–Leu (red) formation. **(b)** Mass spectra of detected products: Phe–Leu and Trp–Leu.

the predicted dipeptide Phe–Ile being synthesized, Ile–Ile was formed (data not shown).

Nonproteinogenic amino acids resembling tryptophan and phenylalanine were also processed into dipeptides, but D-Phe was not a substrate (Table 3). As shown in Figure 7 the only dipeptide formed to a considerable degree by hybrid enzyme **III** was Ile–Phe, with a velocity of 0.5 min^{-1} (Table 3). Hybrid enzyme **III** therefore displayed the highest selectivity. Substitution of the donor amino acid isoleucine with leucine and the acceptor amino acid phenylalanine with tryptophan strongly lowered dipeptide formation (less than 0.02 min^{-1}; Table 3). No Leu–Trp was detectable. The system also did not tolerate a D-configured phenylalanine.

Synthetic templates

The formation of dipeptides by a dimodular NRPS is a multistep process [1] that involves activation of substrate amino acids to adenylates and subsequent thiolation of the activated precursors to enzyme-bound intermediates, followed by peptide-bond formation. The Te domain completes the process by releasing the product. To monitor single reactions in the progress of dipeptide formation we used an *in vitro* system.

Activation: ATP–PP_i exchange reaction

The A domain is the specificity-conferring domain in each NRPS module [23], and therefore dictates the primary structure of nonribosomal peptides. It is for this reason A domains have been studied in the greatest detail. A domains were the first domain type of NRPSs to be investigated as recombinant proteins [37]. They are autonomous domains that retain their catalytic activity when overproduced as distinct proteins [38]. The amino-acid specificity of these recombinant domains was found to accurately match the predictions of *in vivo* studies [19]. By employing A domains artificially linked to PCP domains we were able to investigate the influence of neighboring domains on substrate specificity and affinity. We found that the selectivity and affinity of the TycB2 A domain was affected by the surrounding domain arrangement in the dimodular hybrid enzymes **II** and **III**. The latter was found to have a reduced selectivity towards homologous nonproteinogenic substrates (Figure 3b,c; Table 2).

The importance of conformational changes in multi-domain modules during the process of amino-acid adenylation and product release has been discussed previously [29]. It is possible that the introduction of fusion sites could impair these conformational changes.

Because enzyme **II** is able to catalyze dipeptide formation, we have essentially demonstrated that the internal elongation module TycB2 can be converted into an initiation module simply by repositioning (and deleting of the C domain). No other information appears to be required.

Thiolation: TCA–precipitation assay

In the dimodular NRPSs examined in this study, the transfer of adenylated amino acid precursors from the A domain to the PCP domain involved an engineered junction. Heterologous substrate loading on fused A and PCP domains was shown to be feasible. Two conclusions can be drawn from this finding. First, the domain–domain interaction between the A and PCP domains is maintained sufficiently when fused together via a short linker region. Second, PCP domains exhibit no obvious specificity towards the amino acid to be loaded. In principle, thiolation was found to be dictated by the specificity of the preceding A domain. A possible function of PCP domain specificity could have been interpreted as a sort of proof-reading function of the NRPS.

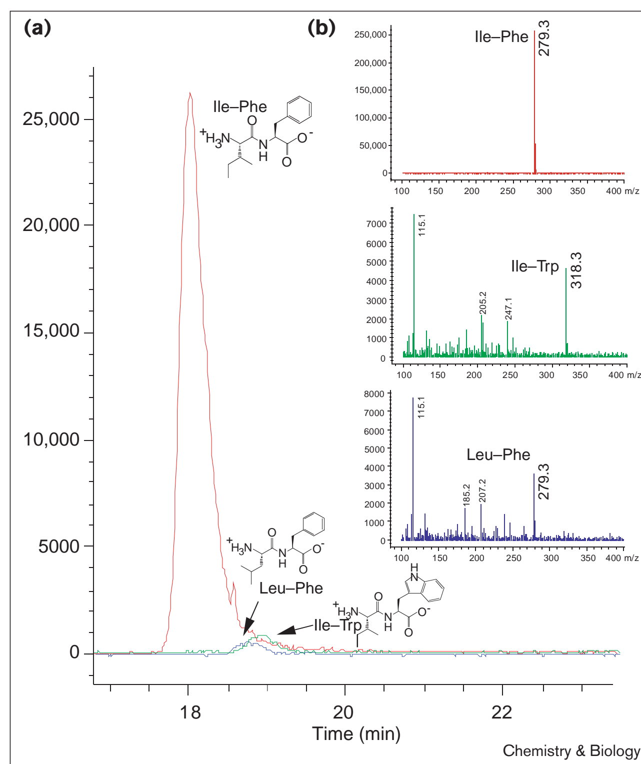
Because the ATP–PP_i exchange reaction monitors the reverse reaction of substrate activation (the breakdown of adenylates to PP_i and amino acid) this assay is not a direct measurement of adenylate formation. It has been proposed that less-stable trapped adenylates are more susceptible to hydrolysis, and that the stability of adenylate intermediates may therefore contribute to substrate selectivity of downstream processes including substrate thiolation [24]. Following on from this idea, differences in the thiolation of aromatic amino acids between the hybrid enzymes **II** and **III** (Figure 4b,c; Table 2) can be explained by an altered kinetic pattern of the preceding A domain rather than by selectivity of the PCP domain. *Vice versa*, the downstream process of the thiolation reaction could influence the upstream A domain reaction by inducing the capturing of substrate adenylates.

Elongation and product release

There is no direct assay available to follow the elongation reaction. The formation of dipeptides and their subsequent release from the hybrid enzymes used in this study can be monitored using TCA-precipitation kinetics as described in the Materials and methods section.

TCA-precipitation kinetics were performed **III**, which is a highly specific template that allows only Ile–Phe to be formed in significant amounts (Table 3). Substrate adenylation and thiolation also predicted the formation of dipeptides with leucine in the donor position and tryptophan in the acceptor position (Figures 3c and 4c). The kinetics monitoring the incorporation level of [^{14}C]-labeled isoleucine are shown in Figure 8. The acceptor amino acids in this reaction can be categorized in three groups. First, phenylalanine, 2′-Thienyl-Ala and Homo-Phe induce an immediate decline in isoleucine incorporation.

Figure 7

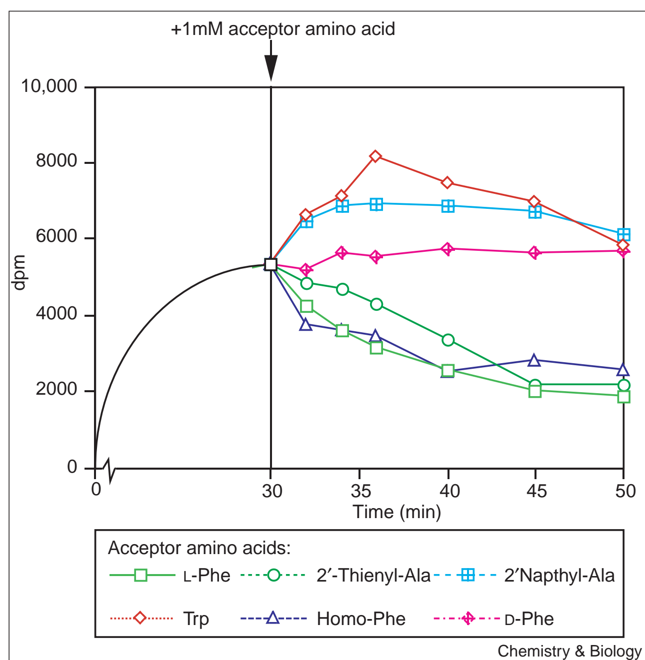


Dipeptides formed by **III** detected using reverse-phase HPLC MS analysis. **(a)** HPLC MS diagrams of Ile–Phe (red), Leu–Phe (blue) and Ile–Trp (green) formation. **(b)** Mass spectra of detected products: Ile–Phe, Leu–Phe and Ile–Trp.

The expectation is that the elongated dipeptides are readily cleaved by the action of the Te domain. The second group includes acceptor amino acids tryptophan and 2′-Napthyl-Ala, which induce a slight increase in TCA precipitate followed by a slow decline. Because no doubling in isoleucine labeling is observed and the incorporation level remains high for more than 20 minutes we conclude that both condensation and termination reactions are slow. The third group is represented by D-Phe, an acceptor amino acid that leads to a constant incorporation level. Here the condensation reaction seems to be blocked. These data are consistent with the observation (Figure 7 and Table 3) that only Ile–Phe, Ile–2′-Thienyl-Ala and Ile–Homo-Phe are formed in significant amounts. In contrast, Ile–Trp and Ile–2′-Napthyl-Ala are formed only very slowly (0.02 min^{-1}; Table 3), and Ile–D-Phe is not detectable.

All three dimodular hybrid NRPSs presented in this study employ C domains fused to noncognate A domains in the donor position. None of these artificial NRPSs were affected in dipeptide formation, which supports the model that the donor position of C domains has little or no influence on substrate selectivity (Table 3).

Figure 8



Kinetics of amino acid thiolation (isoleucine) on the first module of **III** in the presence of excess amounts of amino acids thioesterified on the second module. Phenylalanine, 2′-Thienyl-Ala and Homo-Phe induce an immediate decline in TCA-precipitable isoleucine label, due to formation and release of the dipeptides. Dipeptide formation with tryptophan and 2′-Napthyl-Ala is delayed resulting in a slight increase in TCA-precipitable isoleucine label (reloading of the donor position). With D-Phe as a donor amino acid, a constant level of isoleucine thiolation is observed, indicating that condensation to the Ile–D-Phe dipeptide does not occur. The level of isoleucine thiolation for the first 30 min is represented in an idealized shape.

In contrast, the flexibility of the C domain to elongate noncognate amino acids was restricted to isoleucine — the structurally related amino acid leucine was not used. The C domain of TycB2 in the acceptor position was found to be selective for the cognate amino acid phenylalanine rather than tryptophan (Table 3). This explains the substitution of phenylalanine to tryptophan at position 3 in the cyclic decapeptide tyrocidine [39,40], even though the substrate adenylation favors tryptophan. The observation that C domains exhibit a selectivity towards the acceptor amino acid and not towards the donor amino acid is consistent with recent results [20].

Significance

Many medically and agriculturally important peptides are synthesized nonribosomally by large, multienzyme complexes called nonribosomal peptide synthetases (NRPSs). Their modular nature has attracted attention because of the potential for designing, engineering and generating novel peptides that might be useful as drugs. We have demonstrated here that a set of designed dimodular hybrid NRPSs catalyzes formation of the predicted dipeptides as defined by the specificity of the employed adenylation (A) domains. The *de novo* design of dipeptides with a given sequence was accomplished by fusing various A and thiolation or peptidyl-carrier protein (PCP) domains. The identification of a potential interdomain linker region uncovered the suitable fusion site. The results show that the interaction between foreign A and PCP domains is not impaired, allowing heterologous substrate thiolation. We found that there is a considerable degree of tolerance towards noncognate substrates for the condensation reaction at the donor position, in contrast to the reaction at the acceptor position, where there appears to be a size-exclusion selectivity. In addition, we demonstrated that a native elongation module can be converted into a initiation module by repositioning. The efficiency of the hybrid NRPSs to generate dipeptides *in vitro* was not impaired compared with that of native system. The results characterize NRPSs as a versatile tool for generating bioengineered peptides.

Materials and methods

Bacterial strains and growth conditions

E. coli XL1 Blue (Stratagene, Heidelberg, Germany) was used for preparation of recombinant plasmids. Overproduction of recombinant proteins was carried out in *E. coli* BL21 (λDE3)/p*gsp* using standard protocols [41].

Definition of domains and modules

In this study we have defined an NRPS module as an arrangement of a C, A and PCP (or T) domains (e.g. CAT). Initiation modules do not contain a C domain (i.e. AT). A dimodular NRPS that initiates and terminates dipeptide formation has the domain organization ATCATTe. In this study we use, for the first time, an intramodular A–PCP domain fusion and the domain arrangement to be re-positioned is PCP (or T), C and A domains.

Identification of a potential interdomain linker between A and PCP domains

Sequence alignments of the potential boundary region of A and PCP domains of about 50 NRPS modules from *Bacilli* were performed using the ClustalW program. These alignments revealed a stretch of nine amino acids not highly conserved between 38 and 46 residues amino-terminal of the PCP domain core motif LGGH**S**. This nine-residue site is believed to function as a linker region connecting the NRPS's A and PCP domains. For hybrid fusions at positions 7 and 8 of the nine amino acid stretch, the sequence was altered to Leu–Gln and Val–Asn by introducing *Pst*I and *Hpa*I restriction sites into the plasmid (Figure 2).

Cloning of hybrid dimodular NRPS genes

Plasmid p$[A_{Ile}]_{(bacA1)}$–$[TCA_{Leu}TTe]_{(tycC6)}$ is a derivative of p$[A_{Ile}]_{(bacA1)}$–$[TE]_{(tycA)}$ based on pQE60 (Qiagen). The latter contains the A domain of *bac*A1 fused to the PCP and epimerization domains of *tyc*A by an artificial *Pst*I site (S.D. and M.A.M., unpublished observations). A 1605 bp PCR product containing DNA encoding for the *bac*A1 A domain was obtained using chromosomal DNA from *B. licheniformis* ATCC10716 and the following oligonucleotides: 5′-TTT***CCATGG***TTGCTAAA-CATTCATTAGA-3′ and 5′-TTC***CTGCAG***CGCCCCCGCCGTTCTG-3′ (italic, modified sequences; bold, restriction site). A PCR product containing DNA coding for the domain organization TCATTe from the tyrocidine synthetase gene *tyc*C6 was obtained using chromosomal DNA from *B. brevis* ATCC8185 and the following oligonucleotides: 5′-ATA***CTGCAG***GAGTATGTAGCGCCGC-3′ and 5′-ATA***GGATCC***TTT-CAGGATGAACAGTTCTTG-3′. After digestion of the product with *Pst*I

and *Bam*HI the 4125 bp DNA fragment was cloned into pA$_{(bacA1)}$–TycA digested with *Pst*I and *Bam*HI, to remove the 3′ terminal *tyc*A DNA, yielding p[A$_{Ile}$]$_{(bacA1)}$–[TCA$_{Leu}$TTe]$_{(tycC6)}$.

p[A$_{Phe}$]$_{(tycB2)}$–[TCA$_{Leu}$TTe]$_{(tycC6)}$ is a derivative of p[A$_{Ile}$]$_{(bacA1)}$–[TCA$_{Leu}$TTe]$_{(tycC6)}$. A PCR product containing DNA encoding the *tyc*B2 A domain was obtained using chromosomal DNA from *B. brevis* ATCC8185 and the following oligonucleotides: 5′-*AAT****CCATGG***TGACTGCGCATGAG-3′ and 5′-A*AT****CTGCAG***TGTTGCAGGCTTTCCTTCC-3′. After digestion with *Pst*I and *Nco*I the 1563 bp DNA fragment was cloned into p[A$_{Ile}$]$_{(bacA1)}$–[TCA$_{Leu}$TTe]$_{(tycC6)}$ digested with the same enzymes to remove the *bac*A DNA, yielding p[A$_{Phe}$]$_{(tycB2)}$–[TCA$_{Leu}$TTe]$_{(tycC6)}$.

p[A$_{Ile}$]$_{(bacA1)}$–[TCA$_{Phe}$]$_{(tycB2)}$–[TTe]$_{(tycC6)}$ is a derivative of p[A$_{Ile}$]$_{(bacA1)}$–[TCA$_{Leu}$TTe]$_{(tycC6)}$. The latter was reamplified using inverse PCR techniques that lead to the exclusion of parts of the second module ([TCA$_{Phe}$]$_{(tycB2)}$). The following primers were used: 5′-ACC***GTTAAC***GAATACGTGGCCCCGAG-3′ and 5′-*AAT****GTTAAC***CTC**CTGCAG**CGCCCC-3′. The DNA product was digested using the restriction enzyme *Hpa*I and religated yielding the 6018 bp plasmid p[A$_{Ile}$]$_{(bacA1)}$–[TTe]$_{(tycC6)}$. A PCR product containing DNA coding for parts of the *tyc*B2 module (domain organization TCA) was obtained using chromosomal DNA from *B. brevis* ATCC8185 and the following oligonucleotides: 5′-ACG***CTGCAG***GATTACGTCGCCCCGA-3′ and 5′-AGC***GTTAAC***TGTTGCAGGCTTTCCTTC-3′. The 3111 bp DNA fragment was digested with *Pst*I and *Hpa*I and cloned into *Pst*I- and *Hpa*I-digested p[A$_{Ile}$]$_{(bacA1)}$–[TTe]$_{(tycC6)}$ to yield p[A$_{Ile}$]$_{(bacA1)}$–[TCA$_{Phe}$]$_{(tycB2)}$–[TTe]$_{(tycC6)}$.

Overproduction and purification of dimodular hybrid NRPSs

E. coli BL21(λDE3)/p*gsp* was transformed with the plasmids described above, and the constructs expressed as described previously [6]. p*gsp* is a derivative of pREP4 carrying the 4′phosphopantetheinyl transferase gene (*gsp*) under the control of a T7 promotor. Coexpression with *gsp* allows NRPS holo-enzymes to be produced *in vivo* by modification with CoA. Cells were induced with 0.2 mM IPTG at A_{600} 0.7 and allowed to grow for an additional 1.5 h before being harvested. Purification of the (*His*$_6$)-tagged proteins was basically carried out as described previously [6]. Overproduction and purification was analyzed with Coomassie Brilliant Blue stained SDS polyacrylamide gels [42]. Protein concentration was measured using the Bradford method [43]. Proteins could be stored on ice for at least a week with no observable loss of activity.

ATP–PP$_i$ exchange reaction

The ATP–PP$_i$ exchange reaction that monitors formation of adenylates was carried out as described previously [6]. The enzyme concentration was 20 pmol in a total volume of 100 µl. Substrate amino acids were used at a final concentration of 1 mM. To measure kinetic constants, time courses of the initial velocity of ATP–PP$_i$ exchange were performed at varying amino acid concentrations. The ATP concentration was 5 mM. Kinetic constants were determined using Lineweaver–Burk graphs.

Trichloroacetic acid precipitation assay

The TCA precipitation assay used to monitor substrate thiolation was carried out as described previously [6]. The enzyme concentration was 50 pmol in a total volume of 100 µl. Substrate amino acids were generally used in a threefold excess to enzyme concentration. For kinetics according to Figure 9, 50 pmol enzyme **III** was incubated at 37°C for 30 min with 150 pmol [^{14}C]-labeled isoleucine (substrate of the first module) in a total volume of 100 µl. A solution of 1 mM nonlabeled substrate amino acid of the second module in a total volume of 50 µl was added. Incorporation level of labeled isoleucine was followed by taking samples at distinct times according to the TCA precipitation assay.

During elongation in **III**, isoleucine is loaded onto the PCP domain of module 2. Reloading of the module 1 PCP domain therefore results in a 100% increase in isoleucine incorporation. If elongation and termination of processed dipeptides occur simultaneously, an immediate decline in isoleucine incorporation is observed.

Analogous kinetics were also performed on **II**, underlining that the protein elongates and terminates both dipeptides with tryptophan and phenylalanine in the donor position and leucine and isoleucine in the acceptor position (data not shown). For protein **I**, the assay was hindered by the fact that substrate amino acids are thioesterified on both modules so that dipeptides are formed immediately (data not shown).

Dipeptide formation

Assays for dipeptide formation were carried out in a total volume of 100 µl containing 50 pmol enzyme, 1 mM amino acids, 2 mM ATP in a solution of 50 mM HEPES and 20 mM $MgCl_2$. Negative controls were assayed with no ATP or with only one of the two amino acids. Assays were incubated for 3 h at 37°C before the reaction was quenched by adding 100 µl n-butanol followed by rigorous shaking. The addition of butanol led to precipitation of proteins that were then removed with a pipette tip. The whole mixture was evaporated in a speed-vac and resuspended in 100 µl 10% methanol before being used for HPLC analysis.

To determine product turnover of the enzymes, samples were taken at distinct times and prepared in the same way as described above. The amount of product formed was compared to a quantified standard solution of the purchased dipeptide on HPLC and HPLC MS carried out on a Hewlett Packard Series 1100 MSD. We used a Nucleosil 120-3 C18 reverse phase column.

Radioisotopes and chemicals

Tetrasodium [^{32}P]-pyrophosphate was purchased from NEN Life Science Products. [U-^{14}C]-isoleucine (260 Ci/mol, 100 µCi/ml), [U-^{14}C]-leucine (292 Ci/mol, 100 µCi/ml), [U-^{14}C]-valine (200 Ci/mol, 100 µCi/ml), [U-^{14}C]-phenylalanine (450 Ci/mol, 100 µCi/ml), [U-^{14}C] tyrosine (497 Ci/mol, 100 µCi/ml) and [5-^{3}H]-tryptophan (22.1 Ci/mmol, 1 nCi/ml) were purchased from Hartmann Analytik. Authentic dipeptides used as standards were purchased from Bachem or Sigma, respectively.

Supplementary material

Supplementary material including SDS polyacrylamide gels of overproduced dimodular hybrid NRPSs is available at http://current-biology.com/supmat/supmatin.htm.

Acknowledgements

We are indebted to Veit Bergendahl for help during the HPLC-MS analysis and thank Torsten Stachelhaus and Henning D Mootz for comments on the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft and Fonds der Chemischen Industrie.

References

1. Marahiel, M.A., Stachelhaus, T. & Mootz, H.D. (1997). Modular peptide synthetases involved in non-ribosomal peptide synthesis. *Chem. Rev.* **97**, 2651-2673.
2. Cane, D.E. & Walsh, C.T. (1999). The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. *Chem. Biol.* **6**, R319-R325.
3. Stein, T., *et al.*, & Morris, H.R. (1996). The multiple carrier model of nonribosomal peptide biosynthesis at modular multienzymatic templates. *J. Biol. Chem.* **271**, 15428-15435.
4. 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.
5. Stachelhaus, T., Hüser, A. & Marahiel, M.A. (1996). Biochemical characterization of peptidyl carrier protein (PCP), the thiolation domain of multifunctional peptide synthetases. *Chem. Biol.* **3**, 913-921.
6. 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.
7. Stindl, A. & Keller, U. (1994). Epimerization of the D-valine portion in the biosynthesis of actinomycin D. *Biochemistry* **33**, 9358-9364.

8. Zocher, R., *et al*., Keller, U. (1986). Biosynthesis of cyclosporin A: partial purification and properties of a multifunctional enzyme from *Tolypocladium inflatum*. *Biochemistry* **25**, 550-553.
9. Konz, D., Klens, A., Schörgendorfer, K. & Marahiel, M.A. (1997). The bacitracin biosynthesis operon of *Bacillus licheniformis* ATCC 10716: molecular characterization of three multi-modular peptide synthetases. *Chem. Biol.* **4**, 927-937.
10. Schneider, A. & Marahiel, M.A. (1998). Genetic evidence for a role of thioesterase domains, integrated in or associated with peptide synthetases, in non-ribosomal peptide biosynthesis in *Bacillus subtilis*. *Arch. Microbiol.* **169**, 404-410.
11. Konz, D. & Marahiel, M.A. (1999). How do peptide synthetases generate structural diversity? *Chem. Biol.* **6**, R39-R48.
12. Lambalot, R.H., *et al.,* Walsh, C.T. (1996). A new enzyme superfamily – the phosphopantetheinyl transferases. *Chem. Biol.* **3**, 923-936.
13. Reuter, K., Mofid, M.R., Marahiel, M.A. & Ficner, R. (1999). Crystal structure of the surfactin synthetase-activating enzyme Sfp: a prototype of the 4′-phosphopantetheinyl transferase superfamily. *EMBO J.* **18**, 6823-6831.
14. Weber, G., Schörgendorfer, K., Schneider-Scherzer, E. & Leitner, E. (1994). The peptide synthetase catalyzing cyclosporine production in *Tolypocladium niveum* is encoded by a giant 45.8-kilobase open reading frame. *Curr. Genet.* **26**, 120-125.
15. Aharonowitz, Y., *et al.,* Zhang, J. (1993). Delta-(L-alpha-aminoadipyl)-L-cysteinyl-D-valine synthetase, the multienzyme integrating the four primary reactions in beta-lactam biosynthesis, as a model peptide synthetase. *Biotechnology (NY)* **11**, 807-810.
16. van Wageningen, A.M., *et al*., Solenberg, P.J. (1998). Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. *Chem. Biol.* **5**, 155-162.
17. Stachelhaus, T., Schneider, A. & Marahiel, M.A. (1995). Rational design of peptide antibiotics by targeted replacement of bacterial and fungal domains. *Science* **269**, 69-72.
18. Guenzi, E., Galli, G., Grgurina, I., Pace, E., Ferranti, P. & Grandi, G. (1998). Coordinate transcription and physical linkage of domains in surfactin synthetase are not essential for proper assembly and activity of the multienzyme complex. *J. Biol. Chem.* **273**, 14403-14410.
19. 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.
20. 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.
21. Cane, D.E., Walsh, C.T. & Khosla, C. (1998). Harnessing the biosynthetic code: combinations, permutations and mutations. *Science* **282**, 63-68.
22. Mootz, H.D. & Marahiel, M.A. (1999). Design and application of multimodular peptide synthetases. *Curr. Opin. Biotechnol.* **10**, 341-348.
23. 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.
24. Dieckmann, R., Pavela-Vrancic, M., von Döhren, 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.
25. 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* **8**, 407-418.
26. Gokhale, R.S. & Khosla, C. (2000). Role of linkers in communication between protein modules. *Curr. Opin. Chem. Biol.* **4**, 22-27.
27. Riggs, S.D. & Smithgall, T.E. (1999). SH2-kinase linker mutations release Hck tyrosine kinase and transforming activities in rat-2 fibroblasts. *J. Biol. Chem.* **274**, 26579-26583.
28. Kealey, J.T., Liu, L., Santi, D.V., Betlach, M.C. & Barr, P.J. (1998). Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. *Proc. Natl Acad. Sci. USA* **95**, 505-509.
29. Pavela-Vrancic, M., Dieckmann, R., Döhren, H.V. & Kleinkauf, H. (1999). Editing of non-cognate aminoacyl adenylates by peptide synthetases. *Biochem. J.* **342**, 715-719.
30. Weinreb, P.H., Quadri, L.E., Walsh, C.T. & Zuber, P. (1998). Stoichiometry and specificity of *in vitro* phosphopantetheinylation and aminoacylation of the valine-activating module of surfactin synthetase. *Biochemistry* **37**, 1575-1584.
31. Gokhale, R.S., Tsuji, S.Y., Cane, D.E. & Khosla, C. (1999). Dissecting and exploiting intermodular communication in polyketide synthases. *Science* **284**, 482-485.
32. Shaw-Reid, C.A., Kelleher, N.L., Losey, H.C., Gehring, A.M., Berg, C. & Walsh, C.T. (1999). Assembly line enzymology by multimodular nonribosomal peptide synthetases: the thioesterase domain of *E. coli* EntF catalyzes both elongation and cyclolactonization. *Chem. Biol.* **6**, 385-400.
33. Gokhale, R.S., Hunziker, D., Cane, D.E. & Khosla, C. (1999). Mechanism and specificity of the terminal thioesterase domain from the erythromycin polyketide synthase. *Chem. Biol.* **6**, 117-125.
34. Mootz, H.D., Schwarzer, D. & Marahiel, M.A. (2000). Construction of hybrid peptide synthetases by module and domain fusions. *Proc. Natl Acad. Sci. USA*, in press.
35. 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.
36. Suo, Z., Walsh, C.T. & Miller, D.A. (1999). Tandem heterocyclization activity of the multidomain 230 kDa HMWP2 subunit of *Yersinia pestis* yersiniabactin synthetase: interaction of the 1-1382 and 1383-2035 fragments. *Biochemistry* **38**, 14023-14035.
37. Stachelhaus, T. & Marahiel, M.A. (1995). Modular structure of peptide synthetases revealed by dissection of the multifunctional enzyme GrsA. *J. Biol. Chem.* **270**, 6163-6169.
38. Konz, D., Doekel, S. & Marahiel, M.A. (1999). Molecular and biochemical characterization of the protein template controlling biosynthesis of the lipopeptide lichenysin. *J. Bacteriol.* **181**, 133-140.
39. Fujikawa, K., Sakamoto, Y., Suzuki, T. & Kurahashi, K. (1968). Biosynthesis of tyrocidine by a cell-free enzyme system of *Bacillus brevis* ATCC 8185. II. Amino acid substitution in tyrocidine. *Biochim. Biophys. Acta* **169**, 520-533.
40. Ruttenberg, M.A. & Mach, B. (1966). Studies on amino acid substitution in the biosynthesis of the antibiotic polypeptide tyrocidine. *Biochemistry* **5**, 2864-2869.
41. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. (2nd ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
42. Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.
43. Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254.