Control of Directionality in Nonribosomal Peptide Synthesis: Role of the Condensation Domain in Preventing Misinitiation and Timing of Epimerization†

Uwe Linne and Mohamed A. Marahiel*

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

Received April 4, 2000; Revised Manuscript Received June 16, 2000

ABSTRACT: Product assembly by nonribosomal peptide synthetases (NRPS) is initiated by starter modules that comprise an adenylation (A) and a peptidyl carrier protein (PCP) domain. Elongation modules of NRPS have in addition a condensation (C) domain that is located upstream of the A domain. They cannot initiate peptide bond formation. To understand the role of domain arrangements and the influence of the domains present upstream of the A domains of the elongation modules of TycB on the initiation and epimerization activities, we constructed a set of proteins derived from the tyrocidine synthetases of *Bacillus brevis*, which represent several N-terminal truncations of TycB and the first module of TycC. The latter was fused with the thioesterase domain (Te) to give $TycC_1$-CAT-Te and to ensure product turnover. $TycB_{2-3}$-AT.CATE and $TycB_3$-ATE, lacking an N-terminal C domain, were capable of initiating peptide synthesis and epimerizing. In contrast, the corresponding constructs with a cognate N-terminal C domain, $TycB_{2-3}$-T.CATE and $TycB_3$-CATE, were strongly reduced in initiation and epimerization. Evidence is also provided that this reduction is due to substrate binding in an enantioselective binding pocket at the acceptor position of the C domains. By using $TycB_{2-3}$-AT.CATE and $TycB_3$-ATE, we were able to turn an elongation module into an initiation module, and to establish an in-trans system for the formation of new di- and tripeptides with recombinant NRPS modules. We also show that epimerization domains of elongation modules can in principle epimerize both aminoacyl-S-Ppant ($TycB_3$-ATE) and peptidyl-S-Ppant ($TycB_{2-3}$-AT.CATE) substrates, although the efficiency for epimerizing the noncognate aminoacyl-S-Ppant substrates appears to be lowered.

Many natural peptides of pharmacological interest are synthesized from bacteria and fungi by nonribosomal peptide synthetases (NRPS)[1] (*1, 2*). NRPS are huge multifunctional enzymes with a modular organization that represent an alternative itinerary to ribosomal peptide synthesis (*3, 4*). Size and sequence of the nonribosomally synthesized peptides are determined by the number and collinear arrangement of modules (*5*). Generally, for product assembly, three active sites (domains) are essential, resulting in a minimal elongation module: the A domain for ATP-dependent activation of the cognate amino acid as aminoacyl-*O*-AMP (*6*); the peptidyl-carrier protein [PCP; also described as the thiolation (T) domain] for covalent binding of the activated amino acids (*7, 8*); and the condensation (C) domain for peptide bond formation (*9*). Each PCP is posttranslationally modified at a strictly conserved serine with the cofactor 4′-phosphopantetheinyl. This modification is carried out by a special class of CoASH binding 4′-phosphopantetheinyl transferases (*10*−*12*). The activated amino acids are then tethered to this cofactor to yield the aminoacyl-S-Ppant PCP domain. The transfer of the growing aminoacyl (or peptidyl) chain from the upstream PCP domain to the attacking aminoacyl monomer in thioester linkage to the downstream PCP domain is catalyzed by the C domain.

The reaction sequence of nonribosomal peptide synthesis on a NRPS template is highly specific and leads to a single product. The first step is catalyzed by the so-called initiation module that triggers product formation. The following elongation modules, in contrast, are not capable of initiating the biosynthetic process. These modules remain in the aminoacyl-S-Ppant stage until translocation of the nascent peptidyl chain from the upstream PCP domain has occurred.

Optional modifying domains associated with epimerization (*13*), *N*-methylation (*14*), or heterocyclization reactions (*15*) significantly contribute to the large structural diversity of NRPS products (*16*). Release of the final product is catalyzed by a thioesterase-(Te)-like domain found at the C-terminal module of NRPS templates (*17, 18*). Often these Te domains catalyze an intramolecular cyclization of the peptides, resulting in cyclic products. For the rational design of new bioactive peptides synthesized by engineered NRPS, it would be very helpful to learn more about the molecular mecha-

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

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

[1] Abbreviations: A, adenylation domain; aminoacyl- or peptidyl-S-Ppant, aminoacylated thioester form of cofactor Ppant bound to the strictly conserved serine of T domains; ATP, adenosine 5′-triphosphate; C, condensation domain; CoASH, coenzyme A; DTE, 1,4-dithioerythritol; E, epimerization domain; EDTA, ethylendiaminetetraacetic acid; HPLC, high-performance liquid chromatography; LSC, liquid scintillation counting; MS, mass spectrometry; NRPS, nonribosomal peptide synthetase; PCR, polymerase chain reaction; Ppant, 4′-phosphopantetheine; PP_i, inorganic pyrophosphate; PCP (T domain), thiolation domain; TCA, trichloroacetic acid; Te, thioesterase domain; TLC, thin-layer chromatography.

10.1021/bi000768w CCC: $19.00
Published on Web 07/29/2000

nisms, the interdomain communications during product assembly, and the consequences of the domain rearrangement on their function (*19*). Nothing is known about the features responsible for classifying a module as an initiation or elongation module. Furthermore, in the case of epimerization domains (E domains), an ATE initiation module would be able to epimerize the aminoacyl-S-Ppant, while E domains in CATE elongation modules are thought to epimerize the corresponding peptidyl-S-Ppant acyl enzyme intermediates (*20*, *21*). It is still unclear if there are two types of epimerization domains (aminoacyl and peptidyl epimerases) or if the differences are induced by the domain organization. To address these questions, we designed a set of proteins derived from the tyrocidine synthetases (*22*) TycB and TycC with different deletions. Product formation was monitored in vitro, indicating that the presence of an N-terminal C domain is the determinant that prevents an elongation module from initiating peptide synthesis. Moreover, it was shown that the timing of the epimerization reaction on the peptidyl stage is also coupled to the presence of an N-terminal C domain.

EXPERIMENTAL PROCEDURES

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

All constructs were obtained by PCR amplification with chromosomal DNA of *Bacillus brevis* ATCC8185 as template with the following oligonucleotides (italic, modified sequences; boldface, restriction site): $tycB_{2-3}$-AT.CATE, 5′-*AAT**GCATGC***TGACTGCGCATGAG-3′ and 5′-*ATA**GGATCC***AATCCATTCCAGGATGTTTTCC-3′; $tycB_{2-3}$-T.CATE, 5′-*TTT**GCATGC***CTGCAACAGGAGC-3′ and 5′-*ATA**GGATCC***AATCCATTCCAGGATGTTTTCC-3′; $tycB_3$-CATE, 5′-*AAT**GCATGC***ATAGCGAACAGGCAG-3′ and 5′-*ATA**GGATCC***AATCCATTCCAGGATGTTTTCC-3′; $tycB_3$-ATE, 5′-*AAA**GC**ATGC***TGACAGCAGCAG-3′ and 5′-*ATA**GGATCC***AATCCATTCCAGGATGTTTTCC-3′; $tycC_1$-CAT, 5′-*ATA**CCATGG***AAAAGCAGGAAAACATCGC-3′ and 5′-*AAA**GGATCC***CGAAAGGAAGCGGGCC-3′.

The plasmid carrying the gene for the hybrid $TycC_1$-CAT-$TycC_6$-Te [$TycC_1$-CAT-Te (VI)] was obtained by digestion of pQE60-$tycC_6$-Te (*23*) (containing a *Bgl*II site at the 5′-end) with the restriction enzymes *Bgl*II/*Nde*I and ligation into pQE60-$tycC_1$-CAT previously cut with these enzymes.

The described plasmids were transformed in *E. coli* M15/pREP4. Expression and purification of the His_6-tagged apoproteins were performed as previously described (*9*). Overproduction and purification after single-step Ni^{2+}-affinity chromatography were confirmed by SDS−PAGE (*24*). The protein concentrations were assigned using the method of Bradford (*25*). After dialysis against assay buffer [50 mM HEPES (pH 8.0), 200 mM NaCl, 1 mM EDTA, 2 mM DTE, 10 mM $MgCl_2$], the proteins were shock-frozen in liquid nitrogen and could be stored at −80 °C over several months without significant loss of activity.

Posttranslational Modification of the Enzymes by Sfp and CoASH. Priming of heterologously expressed proteins was achieved by incubation with 200 μM CoASH and 25 nM recombinant *Bacillus subtilis* Ppant transferase Sfp (*12*) to the assays before addition of any amino acid. The reaction mixtures were allowed to incubate for 15 min at 37 °C.

ATP−PP_i Exchange Reaction. A domain activity, specificity, and correct folding were examined by applying the ATP−PP_i exchange assay. We tested all substrate amino acids (L-Trp, L-Phe, D-Phe, L-Asn) relevant for the subsequent product formation assays.

Reaction mixtures in assay buffer (final volume 100 μL) contained 300 nM enzyme and 1 mM amino acid. The reaction was initiated by addition of 2 mM ATP, 0.2 mM tetrasodium pyrophosphate, and 0.15 μCi of tetrasodium [^{32}P]-pyrophosphate. After 10 min incubation at 37 °C, the reaction was quenched by the addition of a stop-mix, containing 1.2% (w/v) activated charcoal, 0.1 M tetrasodium pyrophosphate, and 0.35 M perchloric acid. The charcoal was pelleted by centrifugation, washed once with 1 mL of water, and resuspended in 0.5 mL of water. The charcoal-bound radioactivity was quantified by liquid scintillation counting (LSC).

Radioassay for the Detection of Elongation and Product Release. Aliquots (500 nM) of enzyme were preincubated separately with their substrate amino acids (2 μM [^{14}C]-L-Phe (450 mCi/mmol), 100 μM L-Asn). After 10 min of incubation, product formation was initiated by mixing equal volumes of reaction mixtures counting TycB and TycC derivatives, respectively. At various time points, 200 μL aliquots were taken and immediately quenched by addition of 1 mL of ice-cold TCA (10%). After 15 min of incubation on ice, samples were centrifuged (4 °C, 13 000 rpm) for 20 min, washed 2 times with 1 mL of ice-cold TCA, redissolved in 150 μL of formic acid, and quantified by LSC.

Product Formation Assay. For the detection of peptide products, the TycB derivatives (I, II, III, and IV) and $TycC_1$-CAT-Te (VI) (500 nM each) were assayed at 37 °C with 100 μM L-Phe, D-Phe, or L-Trp and 100 μM L-Asn for up to 2 h in a final volume of 100 μL. The reaction was stopped by the addition of a 10-fold excess of methanol. After centrifugation (13 000 rpm, 15 min) and transfer of the supernatant to fresh tubes, the solvent was removed under vacuum and the residue dissolved in 200 μL of 10% methanol (v/v) and applied (20 μL) to HPLC-MS. Separation of the reaction products was achieved on a 250/3-Nucleosil-C18 reversed phase column by applying the following gradient at a flow rate of 0.3 mL min^{-1} [buffer A: 0.05% formic acid/water (v/v); buffer B: 0.045% formic acid/methanol (v/v)]: loading 10% buffer B, linear gradient up to 60% buffer B in 25 min, followed by a linear gradient up to 100% buffer B in 5 min, and then holding 100% buffer B for 5 min. The stereochemistry of the peptide products was confirmed by comparison with standards. The turnover of the enzymes was defined by co-injection with different amounts of standard.

Radio-TLC Assay for Detection of Phe-S-Ppant and Phe-Phe-S-Ppant Epimerization. To monitor the epimerization activity of the TycB derivatives (I, II, III, and IV), the enzymes were allowed to activate and covalently load [^{14}C]-L- or [^{14}C]-D-Phe; 500 nM holo-enzyme in assay buffer was incubated with 2 μM labeled amino acid. Samples were taken at defined time points and precipitated by the addition of ice-cold 10% TCA. After 15 min incubation on ice and centrifugation (4 °C, 13 000 rpm, 20 min), the pellet was

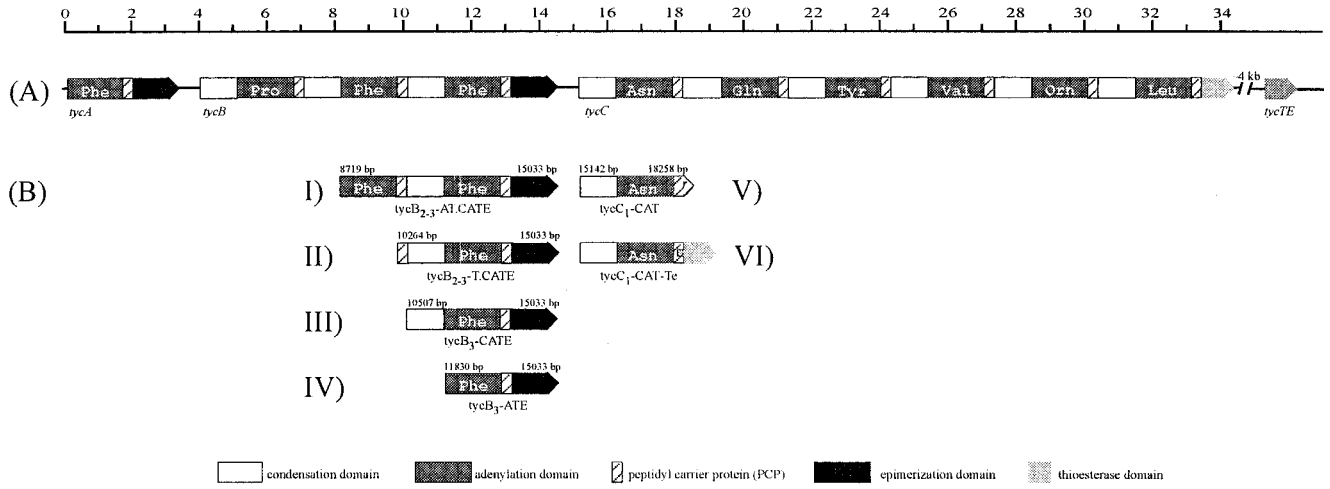

FIGURE 1: (A) Schematic representation of the tyrocidine operon of *Bacillus brevis* ATCC8185. Tyrocidine is assembled on three peptide synthetases: TycA (122 kDa), TycB (405 kDa), and TycC (724 kDa), encoded by the corresponding genes *tycA*, *tycB*, and *tycC*. (B) The gene fragments cloned are presented considering their relative location in the *tyc* operon.

washed 2 times with 1 mL of ice-cold 10% TCA, with 1 mL of ether/ethanol [3:1 (v/v)], and with 1 mL of ether. The pellet was dried at 37 °C. The ^{14}C-labeled amino acid bound to the precipitated enzyme as an acid-stable thioester was hydrolyzed by the addition of 100 μL of 100 mM potassium hydroxide and 10 min incubation at 75 °C. The extraction of the cleaved amino acid was carried out by the addition of 1 mL of methanol and centrifugation for 30 min at 4 °C/13 000 rpm. After transfer of the supernatant to fresh tubes, the solvent was removed under vacuum, and the pellet was dissolved in 20 μL of 50% ethanol (v/v) and applied to chiral TLC plates in the case of Phe-S-Ppant epimerization and to silica gel 60 TLC plates in the case of Phe-Phe-S-Ppant epimerization. The chiral TLC's were developed in acetonitrile/water/methanol [4:1:1 (v/v)] and the silica gel 60 TLC's in butanol/water/acetic acid/ethyl acetate [1:1:1:1 (v/v)] as solvent. The radioactivity was counted with a two-dimensional radio scanner (Raytest, Germany) and quantified by using the supplied RITA software.

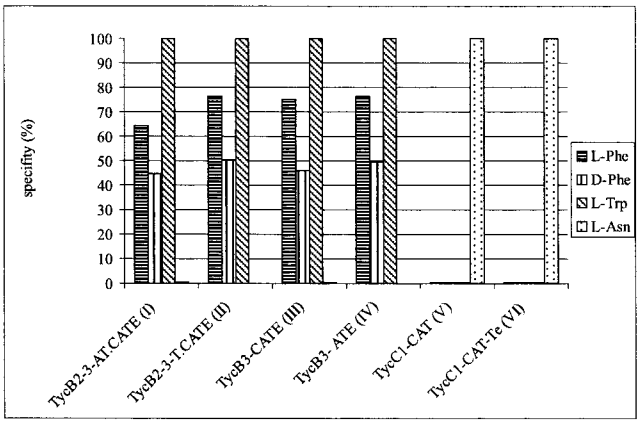

FIGURE 2: Relative ATP−PP$_i$ exchange activity of the recombinant NRPS modules $TycB_{2-3}$-AT.CATE (I), $TycB_{2-3}$-T.CATE (II), $TycB_3$-CATE (III), $TycB_3$-ATE (IV), $TycC_1$-CAT (V), and $TycC_1$-CAT-Te (VI). The highest exchange rate for each module was defined as 100%. The background measured by LSC was below 1%.

RESULTS

Generation and Purification of the Recombinant Enzymes. In this study, we designed a set of six recombinant proteins derived from the tyrocidine syntethases of *Bacillus brevis* (*22*). Four proteins refer to the C-terminus of TycB with N-terminal deletions of different length: $TycB_{2-3}$-AT.CATE (I), $TycB_{2-3}$-T.CATE (II), $TycB_3$-CATE (III), and $TycB_3$-ATE (IV) (see Figure 1).

The remaining two proteins refer to the N-terminus of TycC (see Figure 1). $TycC_1$-CAT (V) was constructed by deletion of the C-terminal modules of TycC, and $TycC_1$-CAT-Te (VI) was generated by ligating the thioesterase domain-encoding region of $TycC_6$ directly to the 3′-end of the $TycC_1$-CAT DNA encoding fragment. All proteins were individually expressed as C-terminal His_6-tag fusions in the heterologous host *E. coli* and purified by Ni^{2+}-affinity chromatography. All six proteins could be purified to homogeneity (confirmed by SDS−PAGE, data not shown) in comparable amounts.

Amino Acid Activation. Generally the substrate specificity of A domains is thought to be independent of the domain environment, and specificity is determined by the structure of the binding pocket (*26*). The A domains of $TycB_2$ and $TycB_3$ are both described as Phe- and Trp-activating domains while $TycC_1$ is an Asn-activating module (*22*). Because some hybrid modules do show marginal differences in their activation pattern compared to the single A domains (*27*), we verified the amino acid selectivity of the recombinant proteins by the PP$_i$ exchange reaction.

As shown in Figure 2 $TycB_{2-3}$-T.CATE (II), $TycB_3$-CATE (III), and $TycB_3$-ATE (IV) were found to activate L-Phe, D-Phe, and L-Trp. The activation patterns of $TycB_2$ and $TycB_3$ are similar, making it impossible to differentiate between the two A domains in the dimodule $TycB_{2-3}$-AT.CATE (I). $TycC_1$-CAT (V) and $TycC_1$-CAT-Te (VI) were verified as Asn-activating modules.

Assays for Investigating Substrate Thiolation and Phe Transfer from TycB Derivatives to $TycC_1$. To investigate the potential of the TycB derivatives (I, II, III, and IV) constructed to transfer to $TycC_1$-CAT (V) or $TycC_1$-CAT-Te (VI), the TycB derivatives (I, II, III, and IV) were preincubated with [^{14}C]-L-Phe and mixed with $TycC_1$-CAT

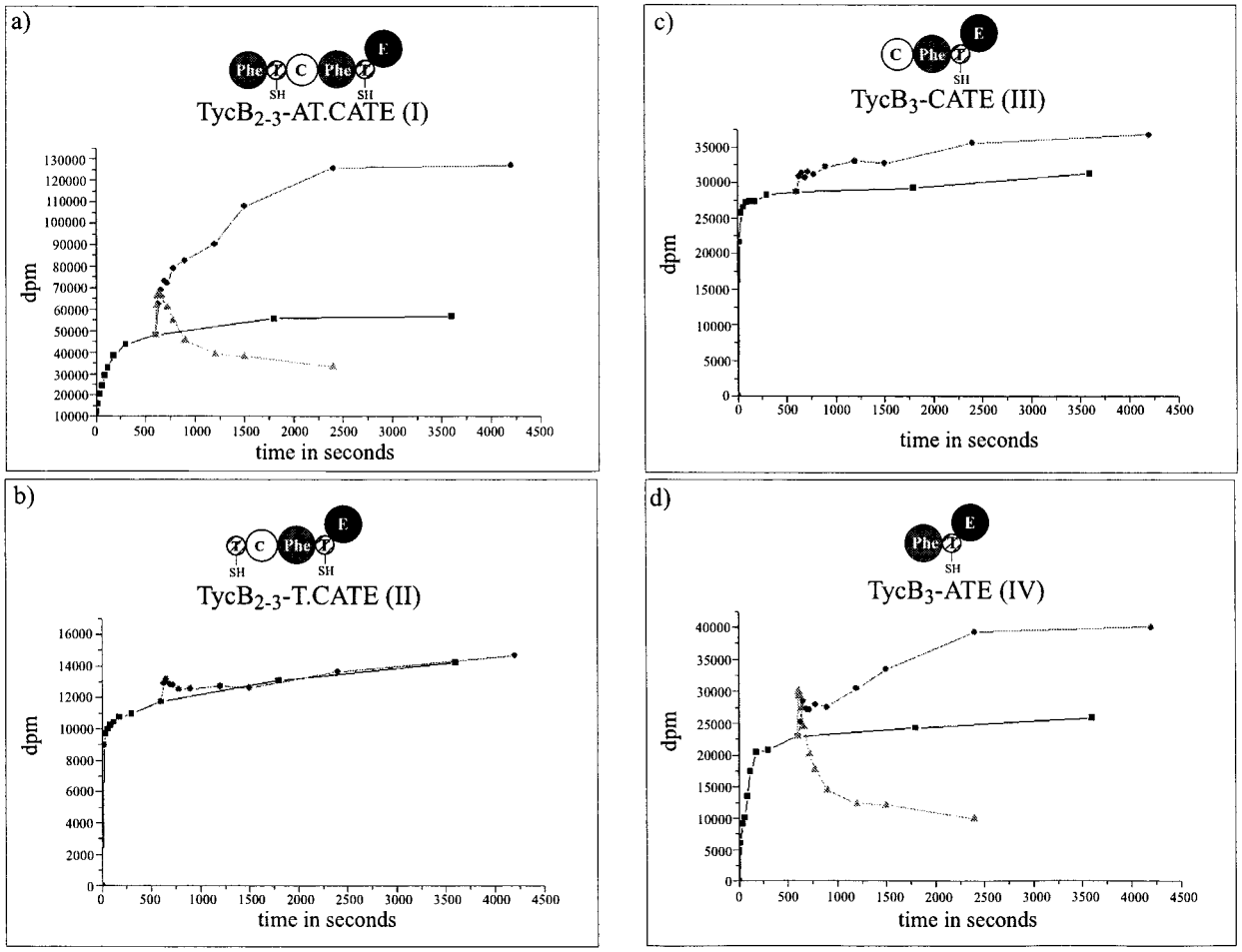

FIGURE 3: Thiolation and transfer kinetics of $TycB_3$ derivatives (a) $TycB_{2-3}$-AT.CATE (I), (b) $TycB_{2-3}$-T.CATE (II), (c) $TycB_3$-CATE (III), and (d) $TycB_3$-ATE (IV). The acceptor modules $TycC_1$-CAT (V) or $TycC_1$-CAT-Te (VI) were preloaded with L-Asn and added to each reaction mixture after 10 min of incubation. An excess of acceptor substrate L-Asn was used. The amount of acid-stable label is shown: squares, without addition of acceptor module; circles, after addition of $TycC_1$-CAT; triangles, after addition of $TycC_1$-CAT-Te. A schematic presentation of the donor enzyme is shown on the top of the corresponding diagram.

(V) or $TycC_1$-CAT-Te (VI) preloaded with L-Asn. Samples were taken at defined time points and immediately quenched by the addition of 10% TCA. Subsequently the amount of acid-stable label was determined by LSC. These values were compared with those obtained for TycB derivatives (I, II, III, and IV) without addition of a $TycC_1$ derivative (V+VI) (see Figure 3). $TycB_{2-3}$-T.CATE (II) showed a reduced activity for thioesterification when compared with the other $TycB_3$ derivatives (III+IV) (see Figure 3b−d).

In the case that labeled Phe [a Phe-Phe dipeptide using $TycB_{2-3}$-AT-CATE (I) as donor] is transferred to the $TycC_1$ derivative (V+VI) and the TycB derivative (I, II, III, and IV) is reloaded with the radiolabeled substrate amino acid phenylalanine, one would expect an increase of enzyme-bound radioactivity. In the case of $TycB_{2-3}$-AT.CATE (I) and $TycB_3$-ATE (IV), we found that the enzyme-bound radioactivity increases strongly after addition of $TycC_1$-CAT (V) (see Figure 3a,d), indicating that these TycB derivatives (I+IV), which lack an N-terminal C domain, can initiate an elongation reaction. Approximately 25 min after addition of the acceptor module $TycC_1$-CAT (V), a stable maximum was reached, revealing no or only negligible product release. Using $TycB_{2-3}$-T.CATE (II) or $TycB_3$-CATE (III) instead, no significant increase in enzyme-bound radioactivity was detectable (see Figure 3b,c). Obviously these TycB derivatives (II+III) carrying a cognate N-terminal C domain were not capable of triggering peptide bond formation by translocation of Phe to $TycC_1$.

$TycB_{2-3}$-AT.CATE (I) and $TycB_3$-ATE (IV) were also assayed for the initiation of tri- and dipeptide formation with the hybrid enzyme $TycC_1$-CAT-Te (VI). The Te domain was fused to the latter construct to catalyze the release of products from $TycC_1$, thereby accelerating recycling of the enzyme for a new elongation round. We found that upon addition of $TycC_1$-CAT-Te (VI) to $TycB_{2-3}$-AT.CATE (I) or $TycB_3$-ATE (IV), the amount of acid-stable label marginally increases before it strongly decreases (see Figure 3a,d). Obviously, the peptide product is rapidly released by the Te domain.

Product Identification and Quantification by HPLC-MS. For identification of the products formed and quantitative analysis of the initiation activity of the TycB derivatives (I, II, III, and IV), an HPLC-MS assay was utilized. This assay was carried out with L-Phe, D-Phe, and L-Trp, all of which are alternative substrates for the A domains of $TycB_2$ and $TycB_3$ (see Figure 2).

The predicted products formed with Phe and Asn were the D-Phe-L-Asn dipeptide (mass 279 Da) using the enzymes $TycB_{2-3}$-T.CATE (II), $TycB_3$-CATE (III), or $TycB_3$-ATE (IV), whereas the tripeptide L-Phe-D-Phe-L-Asn (mass 426

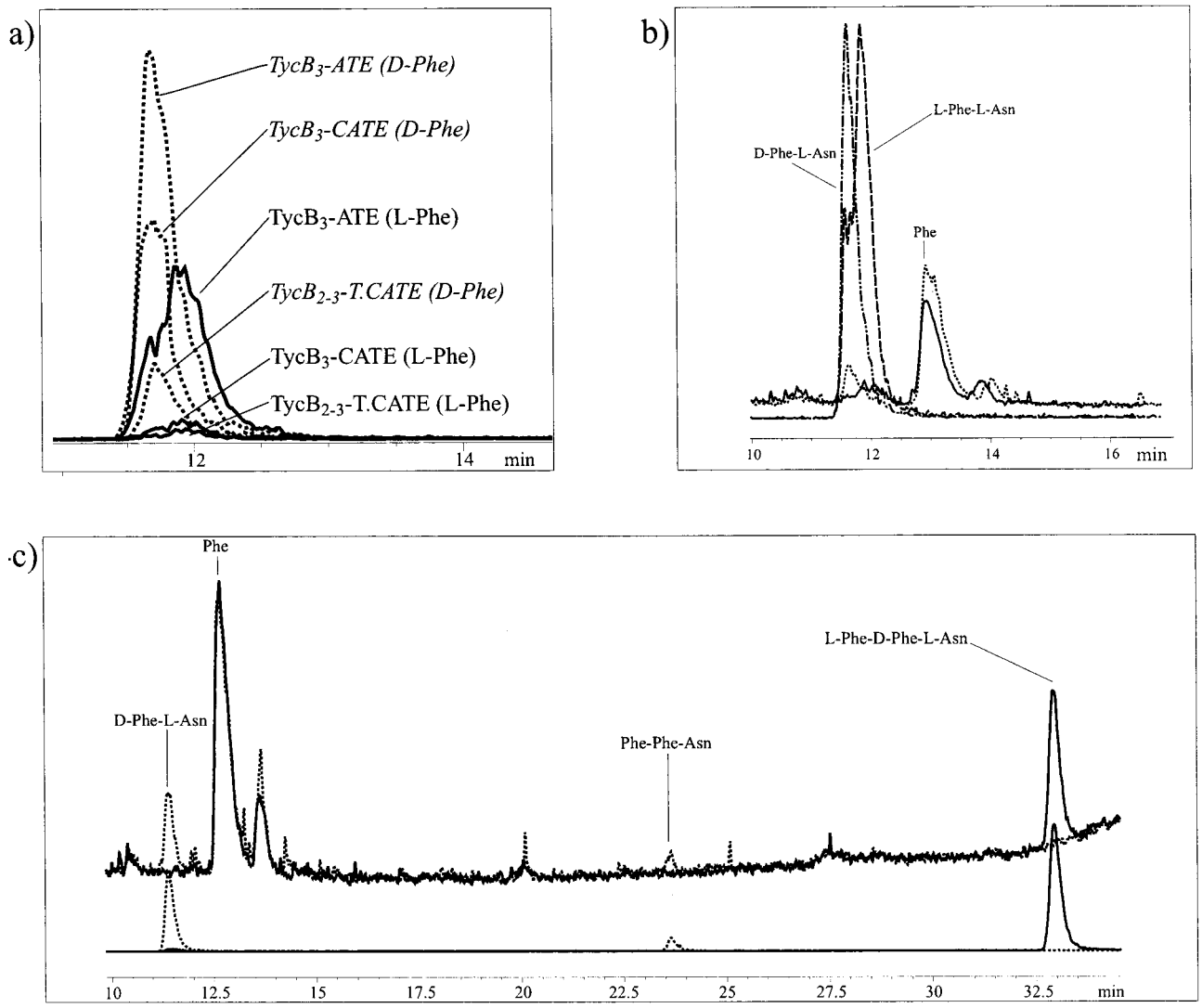

FIGURE 4: Products formed by the TycB derivatives (I, II, III, and IV) as initiating enzymes analyzed by HPLC-MS (mass range detected: 100−600 Da). In all cases, $TycC_1$-CAT-Te (VI) (substrate: L-Asn) was used as acceptor enzyme. The fused Te domain recycled the acceptor enzyme by cleaving off the formed di- and tripeptides of the holo-PCP domain. (a) Relative amount of Phe-Asn formed by $TycB_{2-3}$-T.CATE (II), $TycB_3$-CATE (III), and $TycB_3$-ATE (IV) with L-Phe (solid line) or D-Phe (dotted line, italic letters) as substrate. Only the extracted ion $[M+H^+_{(Phe-Asn)} = 280]$ is shown. (b) Dipeptides formed by $TycB_3$-ATE (IV) with L-Phe (solid line) or D-Phe (dotted line) as substrate. The product peaks (Phe-Asn, printed in full scale mode) were extracted out of the spectra by using the single ion mode $[M+H^+_{(Phe-Asn)} = 280]$. Using L-Phe as substrate, a mixture of the diastereomers L-Phe-L-Asn and D-Phe-L-Asn in a ratio of approximately 2:1 could be detected. When D-Phe was utilized as substrate, only D-Phe-L-Asn was produced. (c) Products formed by $TycB_{2-3}$-AT.CATE (I) with L-Phe (solid line) or D-Phe (dotted line) as substrates. Below the product, peaks extracted out of the original spectra by using the single ion mode are shown $[M+H^+_{(Phe-Asn)} = 280; M+H^+_{(Phe-Phe-Asn)} = 427]$. With L-Phe as substrate L-Phe-D-Phe-L-Asn was formed as the only product. Utilizing D-Phe, no tripeptide L-Phe-D-Phe-L-Asn was produced, but the dipeptide D-Phe-L-Asn, resulting in misinitiation, was synthesized.

Da) should be produced when the dimodular protein $TycB_{2-3}$-AT.CATE (I) is used. Loading with Trp and Asn should lead to the corresponding Trp-Asn di- (mass 318 Da) and tripeptides (mass 504 Da). The identification and stereochemistry of the dipeptide L/D-Phe-L-Asn and the tripeptide L-Phe-L/D-Phe-L-Asn were confirmed by mass analysis and peptide standards. However, the corresponding Trp products were only detected by mass analysis, and no additional information about the stereochemistry is available yet. The results of the HPLC-MS analysis are shown in Figure 4a−c.

In all assays, omitting the acceptor protein TycC1-CAT-Te (VI) from the reaction mixture abolished product formation. In Figure 4a, the amounts of dipeptide Phe-Asn formed by the TycB derivatives (II, III, and IV) incubated with TycC1-CAT-Te (VI) with L-Phe and D-Phe as substrates are shown: $TycB_{2-3}$-T.CATE (II) (L-Phe: $k_{obs} = 0.04\ min^{-1}$; D-Phe: $k_{obs} = 0.21\ min^{-1}$), $TycB_3$-CATE (III) (L-Phe: $k_{obs} = 0.07\ min^{-1}$; D-Phe: $k_{obs} = 0.72\ min^{-1}$), and $TycB_3$-ATE (IV) (L-Phe: $k_{obs} = 0.91\ min^{-1}$; D-Phe: $k_{obs} = 1.16\ min^{-1}$). The data obtained for L-Trp are similar to L-Phe (data not shown). It can be concluded that utilizing L-Phe or L-Trp as substrates, the two enzymes carrying a cognate N-terminal C domain, $TycB_{2-3}$-T.CATE (II) and $TycB_3$-CATE (III), formed only small amounts of the dipeptides Phe-Asn and Trp-Asn. In contrast, using $TycB_3$-ATE (IV), which lacks this N-terminal C-domain, results in the formation of about 15-fold higher levels of dipeptide products. Surprisingly, these latter dipeptides were found to be a mixture of the diastereomers L-Phe-L-Asn and D-Phe-L-Asn in a ratio of approximatley 2:1 (see Figure 4b). When D-Phe was used

as substrate, only the formation of D-Phe-L-Asn could be detected. Additionally, using D-Phe instead of L-Phe, the dipeptide formation rate in the assays with the two enzymes carrying an N-terminal C domain (II+III) increased by a factor of 10. In the case of $TycB_3$-ATE (IV), this increase was only marginal.

$TycB_{2-3}$-AT.CATE (I) produced the tripeptide L-Phe-D-Phe-L-Asn utilizing L-Phe (see Figure 4c) and Trp-Trp-Asn with L-Trp as substrate (data not shown) as expected. Moreover, as shown in Figure 4d, utilizing D-Phe as substrate, no tripeptide L-Phe-D-Phe-L-Asn, but only the dipeptide D-Phe-L-Asn (k_{obs}=1.1 min^{-1}), was produced with nearly the same velocity as in the assays with $TycB_3$-ATE (IV) (k_{obs} = 1.16 min^{-1}). Additionally, we found very small amounts of a product with a mass of 426 Da, which could be D-Phe-D-Phe-L-Asn (see Figure 4c).

Generally, the level of dipeptide (Phe-Asn) formation was in the same range for all TycB derivatives (I, II, III, and IV) with D-Phe as substrate. Utilizing the cognate substrates L-Phe or L-Trp, the two enzymes with an N-terminal C domain (II+III) revealed a 10-fold loss in initiation activity. $TycB_{2-3}$-AT.CATE (I) produced only the expected tripeptides L-Phe-D-Phe-L-Asn and Trp-Trp-Asn.

Epimerization Activity of the TycB Derivatives. Epimerization domains catalyze the L-to-D and D-to-L conversion of aminoacyl-S-Ppant or peptidyl-S-Ppant substrates. In the case of GrsA, it was shown that the equilibrium position is at a L-Phe:D-Phe ratio of 1:1.9 which is reached after a few seconds. It does not matter if L-Phe or D-Phe is used as substrate (*28*).

The extent of Phe-S-Ppant epimerization was investigated by the dependence of time with [^{14}C]-L-Phe and [^{14}C]-D-Phe as substrates for the enzymes $TycB_{2-3}$-T.CATE (II), $TycB_3$-CATE (III), and $TycB_3$-ATE (IV). The results are shown in Figure 5a.

The results revealed that $TycB_3$-ATE (IV) reached the equilibrium position (~40% D-Phe) after approximately 10 min, while $TycB_3$-T.CATE (II) and $TycB_3$-CATE (III) did not reach this value even after 2 h, when L-Phe was utilized. When loaded with D-Phe, $TycB_{2-3}$-T.CATE (II) and $TycB_3$-CATE (III) reached the equilibrium position after approximately 1 h, the same value obtained for $TycB_3$-ATE (IV) with D-Phe as substrate.

Also, the extent of Phe-Phe-S-Ppant epimerization was investigated with [^{14}C]-L-Phe as substrate for the dimodule $TycB_{2-3}$-AT.CATE (I). As shown in Figure 5b, the equilibrium (~62% L-Phe-D-Phe) was reached very fast, within 30 s, the first time point measured.

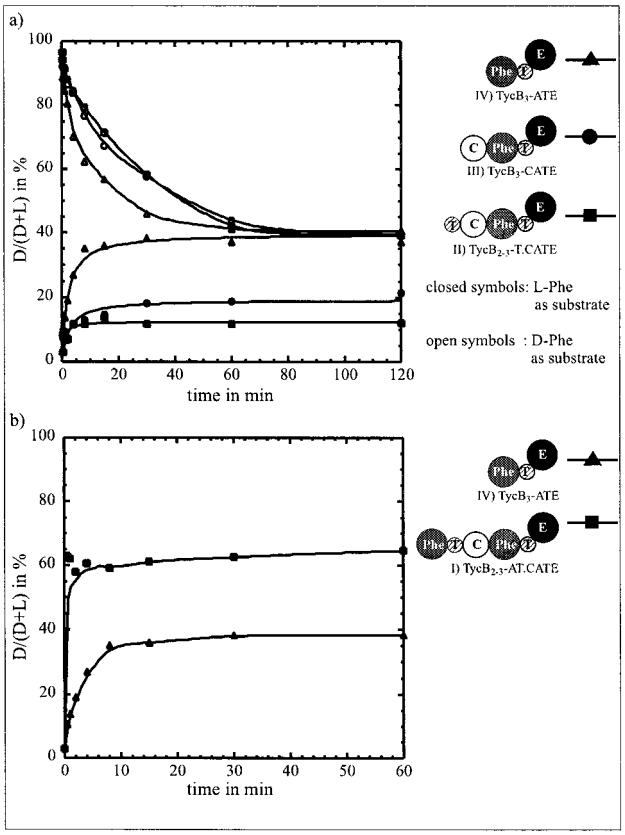

FIGURE 5: (a) L-to-D and D-to-L conversion in the Phe-S-Ppant thioester complex of $TycB_{2-3}$-T.CATE (II) (squares), $TycB_3$-CATE (III) (circles), and $TycB_3$-ATE (IV) (triangles). Open symbols represent [^{14}C]-D-Phe as substrate, closed symbols [^{14}C]-L-Phe. Holo-enzymes were allowed to load with L- or D-[^{14}C]-Phe. The reaction mixtures were incubated at 37 °C, and samples were taken at defined time points. After being quenched with 10% TCA, thioester complexes were hydrolyzed and products applied to chiral TLC plates. After development of the TLC plates in the solvent system acetonitrile/water/methanol [4:1:1 (v/v)], radioactivity was detected and quantified by using two-dimensional radio scanning. (b) L-to-D conversion in the Phe-Phe-S-Ppant thioester complex of $TycB_{2-3}$-AT.CATE (I) (squares) compared to the L-to-D conversion in the Phe-S-Ppant thioester complex of $TycB_3$-ATE (IV) (triangles, see panel a). Holo-enzymes $TycB_{2-3}$-AT.CATE (I) were allowed to load with L-[^{14}C]-Phe. The reaction mixtures were incubated at 37 °C, and samples were taken at defined time points. After being quenched with 10% TCA, thioester complexes were hydrolyzed and products applied to silica gel 60 TLC plates. After development of the TLC plates in the solvent system butanol/water/acetic acid/ethyl acetate [1:1:1:1 (v/v)], radioactivity was detected and quantified by using two-dimensional radio scanning. For comparing the velocities of aminoacyl versus peptidyl-S-Ppant epimerization, additionally the L-to-D conversion in the Phe-S-Ppant thioester complex of $TycB_3$-ATE (squares) is shown.

DISCUSSION

The reaction sequence catalyzed by peptide synthetases is highly template dependent, yielding a product of a distinct length and order. Omitting one amino acid substrate results in a total breakdown of product formation, indicating that initiation of peptide synthesis cannot occur at internal modules of the magasynthetase complexes. In this study, we have defined the role of C domains in initiation and elongation of peptide synthesis by a NRPS as well as timing of epimerization. We have shown that it is possible to convert an elongation module (CATE) into an initiation module (ATE) by deleting the N-terminal C domain.

Within the *tyc* operon the genes *tycA*, *tycB*, and *tycC* (see Figure 1) encode the tyrocidine synthetases TycA, TycB, and TycC that comprise one, three, and six modules. It is obvious that TycB has three elongation modules, whereas TycC has six, all containing N-terminal C domains. In contrast, the initiation synthetase TycA does not have such a C domain. Recently an editing function of C domains was postulated (*29*). It was presumed that C domains harbor a kind of binding pocket with low substrate selectivity at the donor site (the "upstream" site) and a higher selectivity at the acceptor site (the "downstream" site). As a consequence, it was postulated that a C domain binds the aminoacyl-S-

Ppant of its module at the acceptor site until a condensation reaction with the upstream-bound amino acid (or peptide) at the donor site is accomplished. The formed peptidyl-S-Ppant would then no longer be a substrate of the acceptor site of this C domain, so it is released and can react with the next downstream aminoacyl-S-Ppant at the following C domain. Such a scenario would imply the blocking of an uncontrolled initiation reaction of an elongation module by an upstream C domain and explain the observed directionality of peptide synthesis. In this work, we constructed truncated recombinant NRPS to examine this model experimentally. The system we have developed also adresses another issue of de novo dipeptide formation in-cis (Phe-Phe) and di- and tripeptide formation in-trans (Phe-Asn, Phe-Phe-Asn). We present compiling evidence that an elongation module can be switched to an initiation module simply by deleting the cognate N-terminal C domain.

The initiation activities of $TycB_{2-3}$-T.CATE (II) and $TycB_3$-CATE (III) are very low compared to $TycB_{2-3}$-AT.CATE (I) and $TycB_3$-ATE (IV) with L-Phe as substrate. $TycB_{2-3}$-T.CATE (II) and $TycB_3$-CATE (III) carry N-terminal C domains, which are missing in the initiating recombinant enzymes $TycB_{2-3}$-AT.CATE (I) and $TycB_3$-ATE (IV), indicating that the N-terminal C domain is the reason for the enzymes not to initiate. However, using D-Phe instead of L-Phe as substrate, the two enzymes with the N-terminal C domain can initiate dipeptide formation as well. Interestingly, $TycB_{2-3}$-AT.CATE (I) produces the dipeptide D-Phe-L-Asn when loaded with D-Phe with nearly the same rate as $TycB_3$-ATE (IV). These findings support the idea, that the upstream C domain obviously binds the aminoacyl-S-Ppant substrate in an enantioselective pocket until the condensation reaction occurred (see Figure 6A). In the cell-free biosynthesis of surfactin, it was previously shown that although D-Leu was found to be activated, it inhibited the entire biosynthetic process (*30*). This observation can now be explained by the inability effect of C domain specificity at its acceptor site for L-Leu to promote peptide bond formation with a D-Leu-S-Ppant substrate.

At the donor site, one could imagine that the C domains are selective for the length and the composition of the incoming substrates. If C domains were able to discriminate between an aminoacyl-S-Ppant and a peptidyl-S-Ppant, an initiation reaction of an elongation module would also be expected to be blocked because the next downstream C domain would await a peptidyl-S-Ppant instead of an aminoacyl-S-Ppant. Thus, in this case the blocking of misinitiation would be achieved by selectivity at the donor site of the C domain. The C domains of $TycB_3$ and $TycC_1$ normally process a tripeptidyl-S-Ppant and tetrapeptidyl-S-Ppant (see Figure 1), respectively. In this work, however, we were able to show that the C domains of $TycC_1$ and $TycB_3$ do accept both aminoacyl- as well as peptidyl-S-Ppant substrates. Therefore, we conclude that the C domains are not specific for the length and composition of the incoming substrate.

Some natural biosynthetic systems, however, seem to carry N-terminal C domains. These examples include, for instance, cyclosporin synthetase (CYSYN) (*31*, *32*) and actinomycin synthetase II (ACMS II) (*33*). Lysergyl peptide synthetase 1 (LPS 1) does not initiate (*34*), although it lacks an N-terminal C domain. The D-Ala residue of cyclosporin is not epimerized by CYSYN with an integrated epimerization domain, but introduced by an A domain specific for D-Ala. Perhaps D-Ala could not be bound by the enantioselective

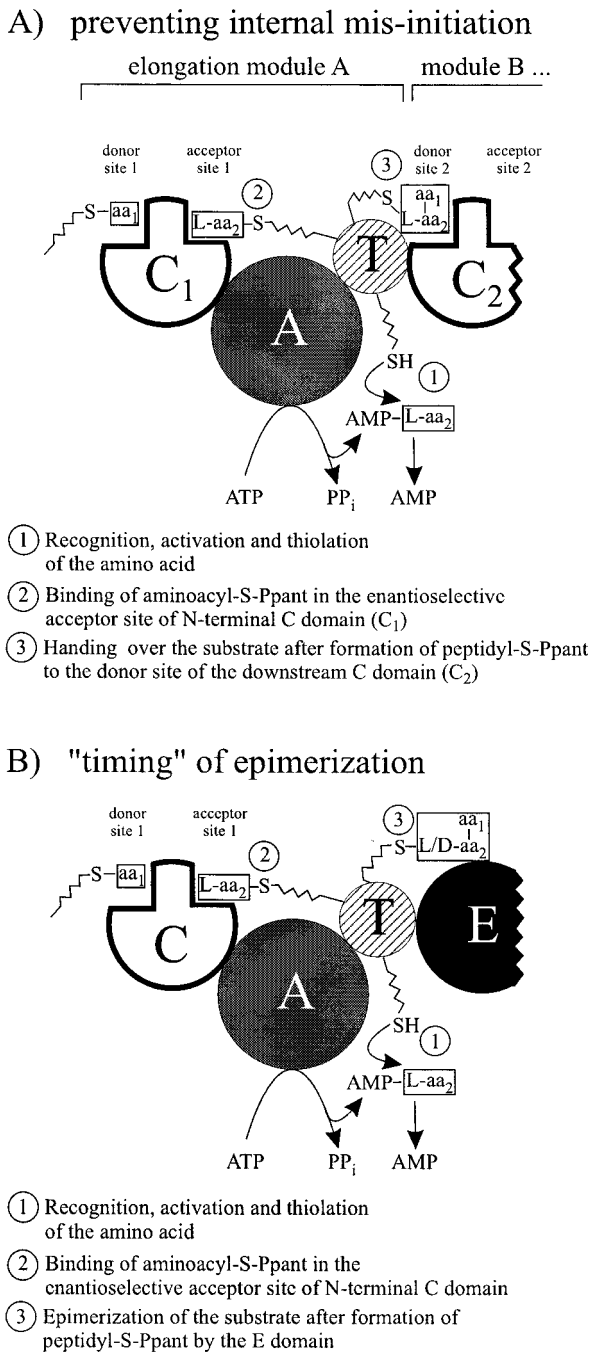

FIGURE 6: Model for the role of N-terminal C domains in preventing internal misinitiation and timing of epimerization. The formed L-aminoacyl-S-Ppant (after recognition and activation of the amino acid by the A domain) is a substrate of the cognate N-terminal C domain and directly after formation bound in an enantioselective acceptor site of this C domain. From the condensation reaction with the upstream aminoacyl- or peptidyl-S-Ppant, a peptidyl-S-Ppant is formed. Being no longer a substrate of the acceptor site, the peptidyl-S-Ppant is handed over to the donor site of the following C domain. The corresponding D-aminoacyl-S-Ppant is not a substrate of the C domain's enantioselective acceptor site. The consequence would be a misinitiation resulting in shortened products. Epimerization of an aminoacyl-S-Ppant by an internal E domain is blocked by the same mechanism involved in preventing misinitiation. After the formation of peptidyl-S-Ppant and release out of the acceptor site, it can be epimerized by the E domain.

acceptor site of the putative N-terminal condensation domain. This would be in agreement with our findings. Otherwise, the putative N-terminal C domain of CYSYN is not well conserved and may be needed either for protein−protein interactions with the D-Ala epimerase that provides the first amino acid or for cyclization reaction rather than for initiation. ACMS II, that was shown to initiate dipeptide formation in vitro (*33, 35*), carries a functional N-terminal C domain. However, for this reaction no rates were published, and the efficiency of product formation was reported to be reduced compared to the natural system. This also corresponds to our findings, as $TycB_{2-3}$-T.CATE (II) and $TycB_3$-CATE (III) showed a strongly decreased initiation activity. In the case of LPS 1 and LPS 2 (*34*), which were assayed in vitro, no product was detected in the absence of D-lysergic acid, the substrate of the initiation module LPS 2. Therefore, studies on how this system prevents misinitiation should awaite further information on the quaternary structure and protein−protein interactions in NRPS.

Another aspect of our described system is the investigation of the role of a N-terminal C domain for peptidyl-S-Ppant versus aminoacyl-S-Ppant epimerization. E domains are found in N-terminal chain initiating modules (e.g., TycA-ATE) as well as in internal elongation modules [e.g., $TycB_3$-CATE (III)] (*3, 36*) (see Figure 1). It was described that the epimerization in -CATE elongation modules occurs at the peptidyl stage. No racemized aminoacyl-S-Ppant was detected in these cases (*20, 37*). The question arises if there are two types of E domains or if the different catalytic properties are induced by the domain environment. Our findings indicate that generally the internal E domain of $TycB_3$ has the ability to epimerize an aminoacyl-S-Ppant (see Figure 5a). When loaded with L-Phe, the epimerization reactions of $TycB_{2-3}$-T.CATE (II) and $TycB_3$-CATE (III), however, are very slow compared to that of $TycB_3$-ATE (IV). Utilizing D-Phe as substrate, the epimerization activity is also lowered slightly in the constructs with the N-terminal C domain, but the differences between $TycB_{2-3}$-T.CATE (II)/$TycB_3$-CATE (III) and $TycB_3$-ATE (IV) are not as significant as with L-Phe as substrate. So it seems that the enantioselectivity of the substrate binding pocket of the C domain is not absolute.

We conclude that the binding of L-aminoacyl-S-Ppant at the acceptor site of the C domain prevents the E domain of an -CATE elongation module from epimerizing the substrate before peptidyl-S-Ppant has been formed (see Figure 6B). This model explains why all E domains in elongation modules seem to act on peptidyl-S-Ppant.

As shown in Figure 5b, the peptidyl-S-Ppant epimerization in the case of the dimodular synthetase $TycB_{2-3}$-AT.CATE (I) is very fast compared to the aminoacyl-S-Ppant epimerization in the case of $TycB_3$-ATE (IV). The equilibrium positions reached are also different (∼62% L-Phe-D-Phe for peptidyl-S-Ppant epimerization and ∼40% D-Phe for aminoacyl-S-Ppant epimerization). Interestingly, both the equilibrium position and the time necessary for reaching it are for peptidyl-S-Ppant-epimerization with $TycB_{2-3}$-AT.CATE (I) in the same range as in the case of the natural aminoacyl-epimerase GrsA-ATE (equilibrium position of ∼66% D-Phe reached after a few seconds) (*28*).

Thus, epimerization domains of elongation modules can epimerize both aminoacyl- and peptidyl-S-Ppant substrates,

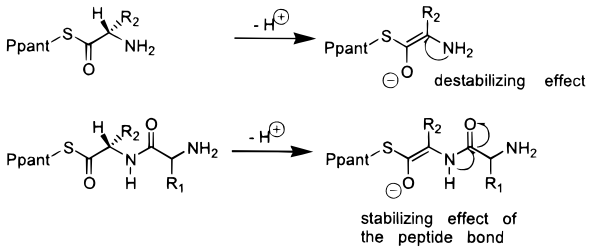

FIGURE 7: Aminoacyl-S-Ppant versus peptidyl-S-Ppant epimerization. A model showing the destabilizing effect of the free electron pair on the amino group of aminoacyl-S-Ppant. In contrast, when epimerizing peptidyl-S-Ppant, the peptide bond stabilizes the free electron pair of the amino group by mesomeric effects and lowers the destabilization of the carbanion intermediate compared to aminoacyl-S-Ppant.

but the efficiency of epimerizing the unnatural aminoacyl-S-Ppant substrates is lowered. This theory is supported by another result: The majority of dipeptide formed with $TycB_3$-ATE (IV) using L-Phe as substrate is L-Phe-L-Asn and not the expected D-Phe-L-Asn. With D-Phe as substrate, only D-Phe-L-Asn was detected. However, in the tripeptide product (peptidyl-S-Ppant epimerization), the second Phe is only found in the D configuration when L-Phe is utilized as substrate.

Little is known about the exact mechanism of the epimerization reaction, but for epimerization of C_α, a proton should be removed, generating a carbanion (*38*). One possible reason for the decreased epimerization activity of $TycB_3$-ATE (IV) for the aminoacyl-S-Ppant substrate could be that the generated carbanion intermediate is less stabilized when formed on an aminoacyl stage compared to the peptidyl stage, which is the natural substrate of this E domain (see Figure 7). The high electron density of the free electron pair on the amino group of aminoacyl-S-Ppant destabilizes the carbanion intermediate. At the peptidyl stage, the free electron pair on the amino group is stabilized by the carbonyl group of the peptide bond, leading to a better stabilization of the carbanion intermediate.

In summary, we have shown that the N-terminal positioning of a cognate C domain has consequences for the initiation activity and the timing of the epimerization reaction of an -ATE- module. The C domain binds the substrate amino acid in an enantioselective binding pocket, blocking further reactions such as epimerization or initiation. Correspondingly, deleting the N-terminal C domain switches a natural elongation module to an initiation module and turns a peptidyl into an aminoacyl epimerase. This has great impact for all future engineering steps with the aim to alter the substrate specificity of NRPS templates. First, the high number of known elongation modules with different substrate specificities could be used as starter modules by deleting the N-terminal C-domains. In this way, for example, shortened peptides can be achieved. Second, for module exchange it is necessary to swap -CA- or -CAT- units and not distinct A domains preventing a misinitiation of the new hybrid NRPS modules. Third, E domains of elongation modules in principle can be utilized as aminoacyl epimerases by deleting the N-terminal C domain; however, a loss of catalytic efficency will be likely. Fourth, we could collect further evidence that the C domains as well as the Te domain of $TycC_6$ (see Figure 1) are obviously not selective for the length and the order of the incoming peptidyl chains.

ACKNOWLEDGMENT

We thank Henning Mootz and Torsten Stachelhaus for valuable suggestions and discussions as well as for critical reading of the manuscript. Henning Mootz and Dirk Schwarzer provided the plasmid carrying $TycC_6$-Te. Veit Bergendahl assisted in HPLC-MS analysis. Inge Schüler supported this work by excellent technical assistance.

REFERENCES

1. Aharonowitz, Y., Bergmeyer, J., Cantoral, J. M., Cohen, G., Demain, A. L., Fink, U., Kinghorn, J., Kleinkauf, H., MacCabe, A., Palissa, H., et al. (1993) *BioTechnology 11*, 807−810.
2. van Wageningen, A., Kirkpatrick, P., Williams, D., Harris, B., Kershaw, J., Lennard, N., Jones, M., Jones, S., and Solenberg, P. (1998) *Chem. Biol. 5*, 155−162.
3. Marahiel, M. A., Stachelhaus, T., and Mootz, H. D. (1997) *Chem. Rev. 97*, 2651−2673.
4. Cane, D. E., and Walsh, C. T. (1999) *Chem. Biol. 6*, R319−325.
5. Stein, T., Vater, J., Kruft, V., Otto, A., Wittmann-Liebold, B., Franke, P., Panico, M., McDowell, R., and Morris, H. R. (1996) *J. Biol. Chem. 271*, 15428−15435.
6. Conti, E., Stachelhaus, T., Marahiel, M. A., and Brick, P. (1997) *EMBO J. 16*, 4174−4183.
7. Stachelhaus, T., Hüser, A., and Marahiel, M. A. (1996) *Chem. Biol. 3*, 913−921.
8. Weber, T., Baumgartner, R., Renner, C., Marahiel, M. A., and Holak, T. A. (2000) *Struct. Fold. Des. 8*, 407−418.
9. Stachelhaus, T., Mootz, H. D., Bergendahl, V., and Marahiel, M. A. (1998) *J. Biol. Chem. 273*, 22773−22781.
10. Lambalot, R. H., Gehring, A. M., Flugel, R. S., Zuber, P., LaCelle, M., Marahiel, M. A., Reid, R., Khosla, C., and Walsh, C. T. (1996) *Chem. Biol. 3*, 923−936.
11. Walsh, C. T., Gehring, A. M., Weinreb, P. H., Luis, E. N., and Flugel, R. S. (1997) *Curr. Opin. Chem. Biol. 1*, 309−315.
12. Reuter, K., Mofid, M. R., Marahiel, M. A., and Ficner, R. (1999) *EMBO J. 18*, 6823−6831.
13. Stein, T., Kluge, B., Vater, J., Franke, P., Otto, A., and Wittmann-Liebold, B. (1995) *Biochemistry 34*, 4633−4642.
14. Zocher, R., Nihira, T., Paul, E., Madry, N., Peeters, H., Kleinkauf, H., and Keller, U. (1986) *Biochemistry 25*, 550−553.
15. Konz, D., Klens, A., Schorgendorfer, K., and Marahiel, M. A. (1997) *Chem. Biol. 4*, 927−937.
16. Konz, D., and Marahiel, M. A. (1999) *Chem. Biol. 6*, R39−48.
17. Schneider, A., and Marahiel, M. A. (1998) *Arch. Microbiol. 169*, 404−410.
18. Shaw-Reid, C. A., Kelleher, N. L., Losey, H. C., Gehring, A. M., Berg, C., and Walsh, C. T. (1999) *Chem. Biol. 6*, 385−400.
19. Cane, D. E., Walsh, C. T., and Khosla, C. (1998) *Science 282*, 63−68.
20. Stindl, A., and Keller, U. (1994) *Biochemistry 33*, 9358−9364.
21. Shiau, C. Y., Baldwin, J. E., Byford, M. F., and Schofield, C. J. (1995) *FEBS Lett. 373*, 303−306.
22. Mootz, H. D., and Marahiel, M. A. (1997) *J. Bacteriol. 179*, 6843−6850.
23. Mootz, H. D., Schwarzer, D., and Marahiel, M. A. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 5848−5853.
24. Laemmli, U. K. (1970) *Nature 227*, 491−493.
25. Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−254.
26. Stachelhaus, T., Mootz, H. D., and Marahiel, M. A. (1999) *Chem. Biol. 6*, 493−505.
27. Doekel, S., and Marahiel, M. A. (2000) *Chem. Biol. 7*, 373−384.
28. Stachelhaus, T., and Walsh, C. T. (2000) *Biochemistry 39*, 5775−5787.
29. Belshaw, P. J., Walsh, C. T., and Stachelhaus, T. (1999) *Science 284*, 486−489.
30. Ullrich, C., Kluge, B., Palacz, Z., and Vater, J. (1991) *Biochemistry 30*, 6503−6508.
31. Dittmann, J., Wenger, R. M., Kleinkauf, H., and Lawen, A. (1994) *J. Biol. Chem. 269*, 2841−2846.
32. Weber, G., Schörgendorfer, K., Schneider-Scherzer, E., and Leitner, E. (1994) *Curr. Genet. 26*, 120−125.
33. Schauwecker, F., Pfennig, F., Schroder, W., and Keller, U. (1998) *J. Bacteriol. 180*, 2468−2474.
34. Walzel, B., Riederer, B., and Keller, U. (1997) *Chem. Biol. 4*, 223−230.
35. Schauwecker, F., Pfennig, F., Grammel, N., and Keller, U. (2000) *Chem. Biol. 7*, 287−297.
36. von Döhren, H., Keller, U., Vater, J., and Zocher, R. (1997) *Chem. Rev. 97*, 2675−2705.
37. Shiau, C. Y., Baldwin, J. E., Byford, M. F., Sobey, W. J., and Schofield, C. J. (1995) *FEBS Lett. 358*, 97−100.
38. De Crecy-Lagard, V., Marliere, P., and Saurin, W. (1995) *C. R. Acad. Sci., Ser. III 318*, 927−936.

BI000768W