Structural and Functional Insights into a Peptide Bond-Forming Bidomain from a Nonribosomal Peptide Synthetase

ARTICLE *in* STRUCTURE · AUGUST 2007

Impact Factor: 5.62 · DOI: 10.1016/j.str.2007.05.008 · Source: PubMed

5 AUTHORS, INCLUDING:

Thomas A Knappe
Philipps University of Marburg
15 PUBLICATIONS **463** CITATIONS

SEE PROFILE

Lars-Oliver Essen
Philipps University of Marburg
118 PUBLICATIONS **4,790** CITATIONS

SEE PROFILE

Please cite this article in press as: Samel et al., Structural and Functional Insights into a Peptide Bond-Forming Bidomain from a Nonribosomal Peptide Synthetase, Structure (2007), doi:10.1016/j.str.2007.05.008

Structure
Article

Structural and Functional Insights into a Peptide Bond-Forming Bidomain from a Nonribosomal Peptide Synthetase

Stefan A. Samel,[1,2] Georg Schoenafinger,[1,2] Thomas A. Knappe,[1] Mohamed A. Marahiel,[1,*] and Lars-Oliver Essen[1,*]
[1] Department of Chemistry/Biochemistry, Philipps-Universität, Hans-Meerwein-Strasse, D-35032 Marburg, Germany
[2] These authors contributed equally to this work.
*Correspondence: marahiel@chemie.uni-marburg.de (M.A.M.), essen@chemie.uni-marburg.de (L.-O.E.)
DOI 10.1016/j.str.2007.05.008

SUMMARY

The crystal structure of the bidomain, peptidyl carrier protein (PCP)-C, from modules 5 and 6 of the nonribosomal tyrocidine synthetase, TycC, was determined at 1.8 Å resolution. The bidomain structure reveals a V-shaped condensation domain, the canyon-like active site groove of which is associated with the preceding PCP domain at its donor side. The relative arrangement of the PCP and the peptide bond-forming condensation domain places the active sites ~50 Å apart. Accordingly, this PCP-C structure represents a conformational state prior to peptide transfer from the donor-PCP to the acceptor-PCP domain, implying the existence of additional states of PCP-C domain interaction during catalysis. Additionally, PCP-C exerts a mode of cyclization activity that mimics peptide bond formation catalyzed by C domains. Based on mutational data and pK value analysis of active site residues, it is suggested that nonribosomal peptide bond formation depends on electrostatic interactions rather than on general acid/base catalysis.

INTRODUCTION

Nonribosomal peptide synthetases (NRPSs) are very large multidomain enzymes (0.15–1.5 mDa) that catalyze simple chemical reactions in a repetitive, assembly line-like manner to produce a broad variety of secondary peptidic metabolites, including many antimicrobial agents and siderophores (Walsh, 2004; Finking and Marahiel, 2004). These synthetases can be described as a linear arrangement of modules, which are responsible for the specific incorporation and often modification of one building block into the peptidic product (Sieber and Marahiel, 2005). As the order of NRPS modules mostly reflects the sequence of amino acids in the peptidic product (Figure 1), the reprogramming of NRPS by reshuffling their constituent modules and manipulation of their specificities offers a promising route for the creation of novel bioactive compounds (Nguyen et al., 2006; Stachelhaus et al., 1995).

At least three different domains are necessary to build up an NRPS module that elongates peptide intermediates: the adenylation (A) domain, the peptidyl carrier protein (PCP) domain, and the condensation (C) domain. The A domain (~550 aa) is responsible for the recognition and activation of the building block by adenylation, which in turn is then covalently attached as a thioester to the prosthetic 4′-phospho-pantetheine (Ppan) group of a subsequent PCP domain (~80 aa). After loading of the PCP domains of two neighboring modules with their cognate substrates, the C domain (~450 aa), located in between, catalyzes the condensation (i.e., peptide bond formation) between the two substrates. This condensation reaction is strictly unidirectional, leading to a downstream-directed synthesis of the NRPS product (Figure 1). As the NRPS activity proceeds, the growing peptide chain is continuously translocated toward the most C-terminal module in the synthetase, where it is eventually cleaved off by hydrolysis, intramolecular cyclization, or reduction (Walsh, 2004).

Even though X-ray and NMR structures of isolated A (Conti et al., 1997; May et al., 2002), PCP (Koglin et al., 2006; Weber et al., 2000), and C (Keating et al., 2002) domains have been solved, little is known about their interactions. With the exception of the initiation module, each PCP domain interacts within an NRPS not only with the adenylation domain, but also with the preceding and following C domains, where it either accepts or donates the growing peptide intermediate (Lai et al., 2006b). NMR titration experiments revealed profound structural plasticity of the PCP domain that changes its tertiary structure dependent on the domain type with which it needs to interact during different steps of NRPS action (Koglin et al., 2006). Furthermore, the prosthetic Ppan group merely spans ~20 Å, by far too little to simultaneously reach the upstream and downstream C domains. Accordingly, not only the domains have to be located so that their distance relative to each other is minimized, as exemplified by the fatty acid synthases (Maier et al., 2006; Jenni et al., 2006), but the PCP domain has to adapt its structure such that the prosthetic Ppan group reaches all the neighboring domains. As a consequence, any

STFODE 1352

Please cite this article in press as: Samel et al., Structural and Functional Insights into a Peptide Bond-Forming Bidomain from a Nonribosomal Peptide Synthetase, Structure (2007), doi:10.1016/j.str.2007.05.008

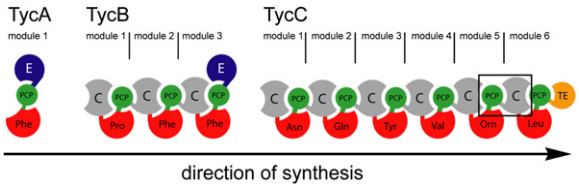

Figure 1. Schematic Depiction of the Tyrocidine Biosynthesis Cluster

Based on their function, the three distinct synthetases TycA-C (order of catalysis: A-B-C) can be divided into modules and domains. Each of the 10 modules specifically activates and incorporates 1 amino acid. Notably, PCP domains interact with up to four different domains, as exemplified by the TycB3 module, the PCP domain of which serves as the A domain, the upstream and downstream C domain, and an epimerization domain.

experimentally determined oligodomain structure might represent a distinct state of NRPS interdomain communication during nonribosomal peptide synthesis.

The C domains share a highly conserved histidine motif (HHxxxDG, core 3) (Marahiel et al., 1997), which is essential for condensation activity (Stachelhaus et al., 1998; Bergendahl et al., 2002). One model suggests that the second histidine of this motif acts as a general base catalyst, restoring the nucleophilicity of the acceptor substrate by deprotonation of its α-ammonium group prior to the initial attack of the amino group onto the carboxyl group of the donor substrate (Bergendahl et al., 2002).

In this work, the first X-ray structure of a PCP-C bidomain is presented, which was excised from modules 5 and 6 of the tyrocidine synthetase, TycC. Within the tyrocidine biosynthesis cluster and its three constituent NRPSs (TycA, TycB, and TycC), this C domain catalyzes the elongation of the nonapeptide DPhe-Pro-Phe-DPhe-Asn-Gln-Tyr-Val-Orn by L-Leu to the linear precursor of tyrocidine A (Figure 1). Although this reaction is not catalyzed in vitro by the isolated PCP-C bidomain, it was found to mediate peptide bond formation during the cyclization of synthetic peptides.

RESULTS AND DISCUSSION

Peptide Cyclization Catalyzed by the TycC5-6 PCP-C Bidomain

To investigate the substrate acceptance of internal C domains, the recombinant PCP-C bidomain from modules 5 and 6 of the tyrocidine NRPS, TycC (*Bacillus brevis,* ATCC 8185), was produced in its *apo*-form (TycC5-6) (i.e., without the Ppan modification at S43). The upstream PCP domain was chemoenzymatically primed with various acceptor substrates, as shown for **P1**-coenzyme A (CoA) in Figure 2. For this chemoenzymatic reaction, the permissive phospho-pantetheine transferase, Sfp, was used to catalyze the transfer of synthetic peptidyl-Ppan arms onto the PCP domain with the corresponding CoA substrates (Belshaw et al., 1999).

Even though no elongated peptide product could be detected after incubation with potential acceptor substrate mimics, HPLC analysis of the priming reaction surprisingly revealed the formation of the cyclic hexapeptide, **cP1** (Figure 2), in 20-fold excess over the background reaction. The head-to-tail connectivity of this product was delineated by MS-MS spectrometry (Figure 3). The use of an N-terminally acetylated variant of this hexapeptidyl-CoA substrate caused no product release—neither of the cyclic peptide, nor of the linear peptide released by hydrolysis of the PCP-bound thioester (data not shown). Furthermore, the peptidyl-Ppan modification of the isolated recombinant PCP domain of TycC5 alone led to no increased cyclization ratio compared to the background reaction.

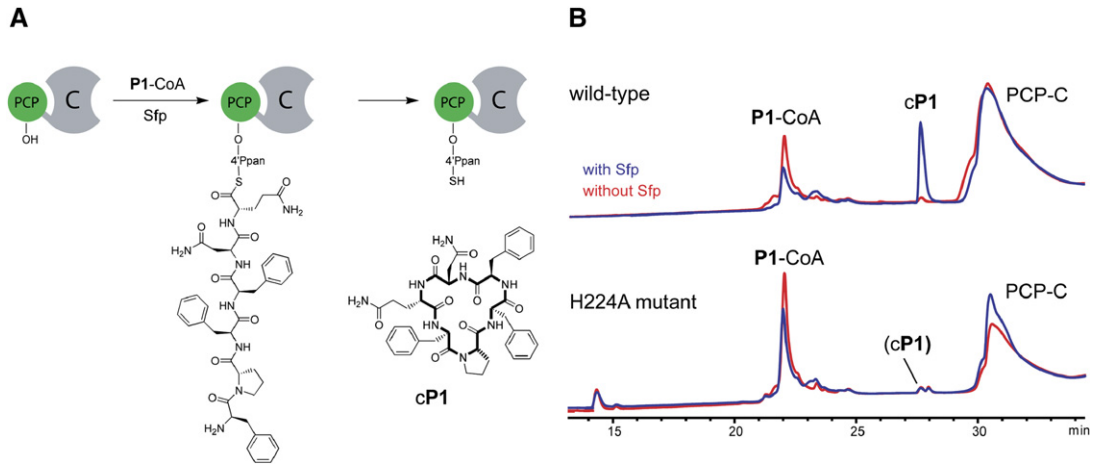

Figure 2. Biochemical Analysis of the TycC5-6 PCP-C Bidomain

(A) Schematic presentation of the cyclization assay performed. Apo-PCP-C is enzymatically primed with P1-CoA leading to the corresponding peptidyl-*holo* form. The C domain then catalyzes the release of the head-to-tail cyclized variant of the hexapeptide (cP1).

(B) HPLC traces of cyclization assays with the wild-type enzyme (top) leads to the formation of cP1, whereas, in the H224A mutant, cyclization activity appears to be abolished (bottom).

Please cite this article in press as: Samel et al., Structural and Functional Insights into a Peptide Bond-Forming Bidomain from a Nonribosomal Peptide Synthetase, Structure (2007), doi:10.1016/j.str.2007.05.008

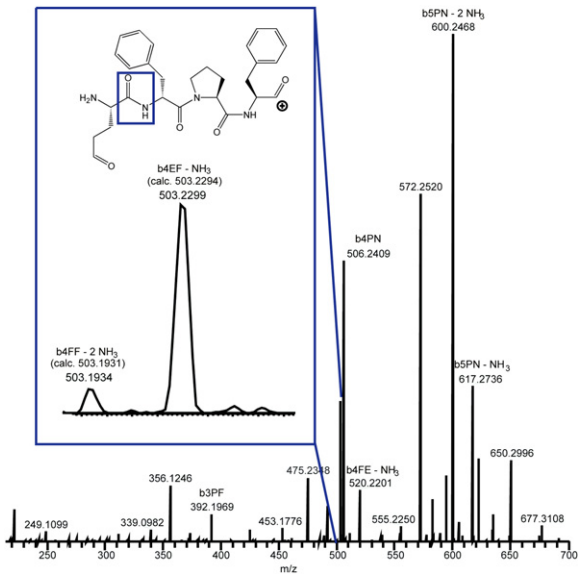

Figure 3. MS-MS Analysis of the Reaction Product

In the b-series, the deaminated fragment, b4EF-NH_3 can be found with a mass of 503.2299 Da. The presence of this species proves the cyclic connectivity of the product, since there is no amide bond between Gln and D-Phe in the absence of head-to-tail cyclization of the linear peptide substrate.

To demonstrate a direct role of the C domain in this cyclization reaction, mutants of the HHxxxDG motif (Marahiel et al., 1997) in PCP-C were prepared by site-directed mutagenesis. Since the second histidine of this motif had been suggested to be essential for catalysis (Bergendahl et al., 2002), it was mutated to HAxxxDG and HVxxxDG, respectively. Under the same assay conditions, PCP-C mutants in position 224 showed neither peptide cyclization nor thioester hydrolysis, even though an Sfp-dependent consumption of substrate due to the priming of the PCP domain could be monitored by LC-MS (Figure 2). In contrast, the corresponding AHxxxDG mutant in which the first histidine of the HHxxxDG motif (H223) was replaced still displayed ~28% of the wild-type activity.

The use of six other peptidyl-substrates (**P2–P7**; Table 1), the cyclization activity of PCP-C was further explored. With lengths between 5 and 8 amino acids, the type and the absolute configuration of the terminal amino acids was altered in comparison to **P1**. The PCP-C bidomain only cyclized substrates that contained D-configured N termini (**P1**, **P3**, and **P7**). No hydrolysis product was cleaved off the enzyme, except in the case of the octapeptidyl substrate, **P6**. The C-terminal L-Gln cannot be exchanged to L-Ala in these reactions. Whenever a C domain-dependent cyclization was observed, autocatalytic formation of the cyclic product occurred with relative amounts of 5%–17%, as determined by peak integration of UV/VIS-recorded chromatograms.

Overall, it is shown here that the TycC6 condensation domain exhibits a novel in vitro cyclization activity on peptides presented at its donor site. The primary sequence of the substrate, **P1**, was derived from the first six amino acids of the decapeptide, tyrocidine. The condensation domain tested, however, was from the 10th module (TycC6) of the tyrocidine NRPS cluster, where it elongates the nonapeptidyl substrate, DPhe-Pro-Phe-DPhe-Asn-Gln-Tyr-Val-Orn, by addition of L-Leu to the scaffold's C terminus. Synthetic CoA substrates with a C-terminal ornithine, however, lacked sufficient stability for in vitro assays, presumably caused by formation of a six-membered lactam due to an attack of the ornithine δ-amino group onto the CoA thioester.

For intramolecular head-to-tail cyclization of **P1**, a proximal arrangement of both termini must be sterically adopted within the active site of the C domain so that intramolecular peptide bond formation can proceed. Considering the linear tyrocidine precursor, it is assumed that it tends to prefold by intramolecular hydrogen bonds due to its delicate primary sequence in the N-terminal part (Kuo and Gibbons, 1980). The alteration of D- and L-configured amino acids, as well as their hydrophobicity and steric strain, are believed to be important for such a prearrangement, which might also partially occur in the truncated peptidyl substrates assayed here. Taking our current understanding of the selectivity of C domains (Bergendahl et al., 2002; Belshaw et al., 1999; Clugston et al., 2003) into account, one might hence suspect that the intramolecular cyclization mimics the mode of intermolecular peptide bond formation catalyzed by condensation domains with their cognate substrates. One indication for this is given by the observation that the mutants of the second histidine in the HHxxxDG core motif (H224) showed no cyclization activity at all.

Overall Structure of the TycC5-6 PCP-C Bidomain

The structure of the excised PCP-C bidomain was solved at 1.8 Å resolution by multiple anomalous diffraction (MAD) experiments from crystals of SeMet-labeled protein. Mass-spectrometric analysis of the SeMet-labeled TycC5-6 PCP-C bidomain showed that, as in the previous case of EntB (Drake et al., 2006), the absence of iron in the culture medium induced transfer of a Ppan group onto S43 of the PCP domain (data not shown). This post-translational modification initially prohibited growth of SeMet-labeled crystals and is apparently caused by the action of the endogenous phospho-pantetheinyl transferase, EntD, which is part of the *Escherichia coli* enterobactin biosynthesis cluster. Addition of 15 μM Fe(II) to the SeMet-labeling medium abolished this modification by suppressing the enterobactin biosynthesis cluster.

After refinement, each asymmetric unit of the PCP-C bidomain crystals contains one polypeptide chain defined from R3 to M455 and from Q461 to L522, 294 water molecules, and two sulfate ions and dioxane molecules, which originate from the crystallization buffer. The PCP domain of module 5 of TycC (residues M1–T82) measures approximately 20 Å × 24 Å × 30 Å, the C domain of module 6 (residues V101-L522) ca. 65 Å × 50 Å × 40 Å. Both domains are connected via an 18-residue linker (A83-P100) running along the PCP-C domain interface (Figure 4A).

STFODE 1352

Please cite this article in press as: Samel et al., Structural and Functional Insights into a Peptide Bond-Forming Bidomain from a Nonribosomal Peptide Synthetase, Structure (2007), doi:10.1016/j.str.2007.05.008

Table 1. Substrates Used in Cyclization Assays

CoA Substrate		Cyclization (Ratio)[a]	Hydrolysis
P1	DPhe-Pro-Phe-DPhe-Asn-Gln	++ (20:1)	–
P2	LPhe-Pro-Phe-DPhe-Asn-Gln	–	–
P3	DAla-Pro-Phe-DPhe-Asn-Gln	++ (7:1)	–
P4	LAla-Pro-Phe-DPhe-Asn-Gln	–	–
P5	DPhe-Pro-Phe-DPhe-Asn-Ala	–	–
P6	DPhe-Pro-Phe-DPhe-Ala-Ala-Asn-Gln	–	++
P7	DPhe-Pro-Phe-Asn-Gln	+ (n.d.)[b]	–

Hexapeptides with D-configured N termini and C-terminal glutamine are cyclized by PCP-C with ratios ranging from 7- to 20-fold in comparison with background reactions. Hydrolysis was only observed when the octapeptidic substrate P6 was used.
[a] Minus signs indicate that no cyclization product was observed.
[b] n.d. indicates values not determined and refers to the ratio of cyclization to hydrolysis.

The structure of the PCP domain closely corresponds to the A/H state (Koglin et al., 2006). This state was found in the PCP domain of the third module of TycC as an intermediary conformational state that is present in both the Ppan-modified (holo [H]) and unmodified (apo [A]) PCP domain (Koglin et al., 2006). A superposition with the A/H-like state of the PCP domain of module 3 of TycC (Weber et al., 2000) gives an rmsd of 1.67 Å for 56 C_α positions. In the holo form of PCP domains, the A/H state is in conformational equilibrium with the H state, which becomes increasingly stabilized by interactions with other holo-PCP-recognizing domains, like the editing thioesterase II (Koglin et al., 2006). This H state diverges significantly from the conformation of the TycC5 PCP domain, as 78 C_α positions superimpose with an rmsd of only 5.5 Å, mostly along helix αI and αIV with a displacement of 2.3 Å for 18 C_α positions. Accordingly, the PCP-C domain interface observed in the crystal structure has to represent a state, where adoption of the A/H state for the PCP domain is compatible for loading with an aminoacyl residue by the preceding A domain.

The only as-yet-known structure of a C domain is VibH from the vibriobactin synthesis cluster (Keating et al., 2002), which recognizes in *trans* both the donor substrate, PCP-bound 2,4-dihydroxy-benzoic acid, and the acceptor substrate, norspermidine. Like the isolated VibH domain, the TycC6 condensation domain consists of two mainly separated and structurally similar subdomains, an N-terminal (V101–S268) and a C-terminal subdomain (A269–L522). Both subdomains are arranged in a V-shaped fashion and belong to the chloramphenicol-acetyltransferase (CAT) fold. There are only two major contact sites between these two CAT-like subdomains, thus giving rise to a large, canyon-like active site groove: one is mainly made at the floor of the active site canyon and comprises the loop, β8-β9 (T359–V374), with a short, internal 3_{10} element and the conserved arginine R361 that stabilizes the loop conformation by H-bond interactions between its side chain and the peptide carbonyls of L366, E367, and I369. Besides a salt bridge between the side chains of D365 and R252, most H-bond interactions between the floor loop and the N-terminal subdomain are derived from the main chain, as this loop is made up of mostly hydrophobic residues. The second region is strand β11 that is donated from the C-terminal to the N-terminal subdomain, and thereby extends the four-stranded β sheet (β1-β6-β5-β4) of the latter. This extension (N438–F465) spans like a bridge over the active site canyon (Figures 4A and 7C), and appears to be rather flexible due to high thermal B-factors as compared with the remaining condensation domain.

A superposition of the TycC6 C domain (residues 101–522) with VibH (PDB code: 1L5A; pairwise sequence identity, 19%) shows structural similarity with an overall rmsd of 1.58 Å for 197 C_α-carbons (Figure 4B). Due to the fact that the arrangement of the two CAT-like subdomains is off-rotated by 12°, in the superposition only the C-terminal CAT-like subdomains are fitted properly, whereas the orientation of the N-terminal CAT-like subdomain of the TycC6 C domain as a whole—compared with the corresponding VibH N-terminal subdomain—is slightly distorted. The hinge-like region, around which this swiveling motion is centered, corresponds to S268 (VibH: S174) in the short connection between helices α5 and α6. Accordingly, structural comparisons of the corresponding subdomains were performed (Figure 4C), yielding a lower rmsd value of 1.35 Å for 149 C_α-carbons of the C-terminal CAT-like subdomains and an increased value of 1.70 Å for 116 C_α-carbons of the N-terminal CAT-like subdomains.

Two sulfate ions are located in the vicinity of the two domains' active site residues, S43 and H224, respectively (Figure 4A). The sulfate ion next to the PCP domain forms salt bridges to the conserved residues, H42 (Nδ, 3.11 Å) and R45 (Nε, 2.92 Å; Nη, 2.89 Å), and is additionally stabilized by the dipole moment of helix αII, which points with its N terminus onto this sulfate. Obviously, this sulfate anion occupies the supposed position of the phosphate group of the Ppan arm, if the latter is esterified to S43 of the PCP domain. The sulfate ion in the C domain's active

Please cite this article in press as: Samel et al., Structural and Functional Insights into a Peptide Bond-Forming Bidomain from a Nonribosomal Peptide Synthetase, Structure (2007), doi:10.1016/j.str.2007.05.008

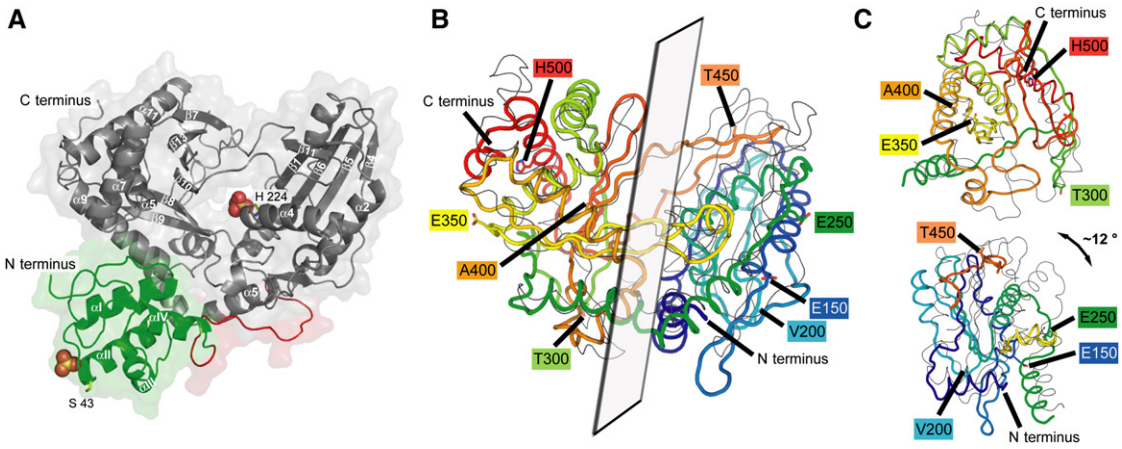

Figure 4. Overall Structure of PCP-C and Structural Comparison with VibH

(A) Overall structure of the TycC5-6 PCP-C bidomain. The PCP domain is colored green, the linker is red, and the C domain is gray. The catalytically active residues, S43 and H224, and two sulfate ions bound to the active sites are shown.

(B) Comparison of the TycC6 condensation domain with VibH. The TycC6 C domain is rendered in rainbow ranging from blue (N terminus) to red (C terminus). The structure of VibH is represented by a black ribbon. For clarity, every 50th residue has been labeled in Figures 3B and 3C.

(C) Illustration of the superposition of the N-terminal and C-terminal subdomains, respectively, viewed from the "inside" after intersection along the gray plane shown in Figure 3B. Figures 4–7 were generated with PyMol 0.98 (DeLano Scientific) (DeLano, 2002).

site forms a salt bridge to the catalytically active residue, H224 (Nε2, 2.87 Å), and an H bond to the amide nitrogen of G229 (2.68 Å).

A structure-based sequence alignment between the VibH and TycC6 C domains and other NRPS C and epimerase domains (Figure 5A) indicates that, in the E domain, there are several sites distributed all over the protein where minor loop insertions and deletions occur. According to this alignment, there is no hint of an insertion unique to epimerase domains that may block the acceptor side of the substrate channel, as suggested previously (Keating et al., 2002).

Interactions between the C Domain and its PCP Donor Domain

There are several indications that the observed PCP-C domain arrangement is not the peptide bond-forming conformational state, where the Ppan-bound nonapeptide is passed from the donor side of the C domain's active site. First, the relative orientation of the PCP and C domains (Figure 4A) places the catalytic residues, S43 and H224, at a distance of 47 Å (Figure 5B), which is more than the Ppan arm can span (~20 Å). Second, Lai et al. (2006a) identified several residues of the EntB PCP domain from the enterobactin biosynthesis cluster that are crucial for productive interaction with the downstream C domain. Those residues of EntB, M249, F264, and A268, correspond to M47, L63, and F67 of the TycC5 PCP domain. As these residues are proximal to the Ppan attachment site, S43, they are far off-positioned from the surface of the TycC6 condensation domain (Figure 6A).

Nevertheless, there are several well-defined interactions found at the domains' interface. Besides the connection via the linker region PCP-C domain, interactions occur in two separated clusters, thereby occluding 1089 $Å^2$ from bulk solvent access. In the first cluster (Figure 6A), H bonds are formed between the side chains of Y6 and K397 (2.67 Å) and between water molecule HOH88 and the side chains of Q29 (3.19 Å), Q394 (2.84 Å), K397 (2.70 Å), and the peptide group preceding G328 (3.08 Å), respectively. Interactions in the second cluster are mainly made by H bonds. Besides bridging water molecules, there are also direct interactions of residues of the PCP domain with those of the C domain. For instance, R16 forms H bonds to both G405 (2.60 Å) and D408 (2.83 Å), and E56 forms salt bridges to H406 (3.00 Å) and K273 (4.31 Å). Interactions of the linker comprise hydrophobic interactions formed by F88 with the residues W261 and F265 of the C domain. Furthermore, several residues of the linker are involved in an intricate H-bond network with both, the PCP and C domain (Figure 6B). Not only does N86 interact with T82 (3.15 Å) and the carboxylic group of I79 (3.18 Å), but its Oδ atom also forms an H bond with the amide nitrogen of F88 (2.91 Å). Likewise, the side chain of D257 interacts with the amide nitrogen of V93 (2.91 Å), whereas the indole amine function of W261 is H bonded to the amide oxygen of I90.

Overall, it is possible that the domain arrangement observed here represents the domains' relative positions in which the PCP domain interacts with either the preceding A or C domain. For example, loading of the Ppan arm of the PCP domain with a cognate amino acid by action of the A domain may trigger an A/H→H transition of the PCP conformation (Koglin et al., 2006; Lai et al., 2006b). Such a transition would set the PCP domain free from the observed PCP-C interface (Figure 6C) so that it can interact with either the upstream C domain along its acceptor side or the downstream C domain at its donor side. The 18-residue-long PCP-C linker is only partly involved in PCP-C interactions and could obviously support such domain rearrangements, as its final seven residues

STFODE 1352

Please cite this article in press as: Samel et al., Structural and Functional Insights into a Peptide Bond-Forming Bidomain from a Nonribosomal Peptide Synthetase, Structure (2007), doi:10.1016/j.str.2007.05.008

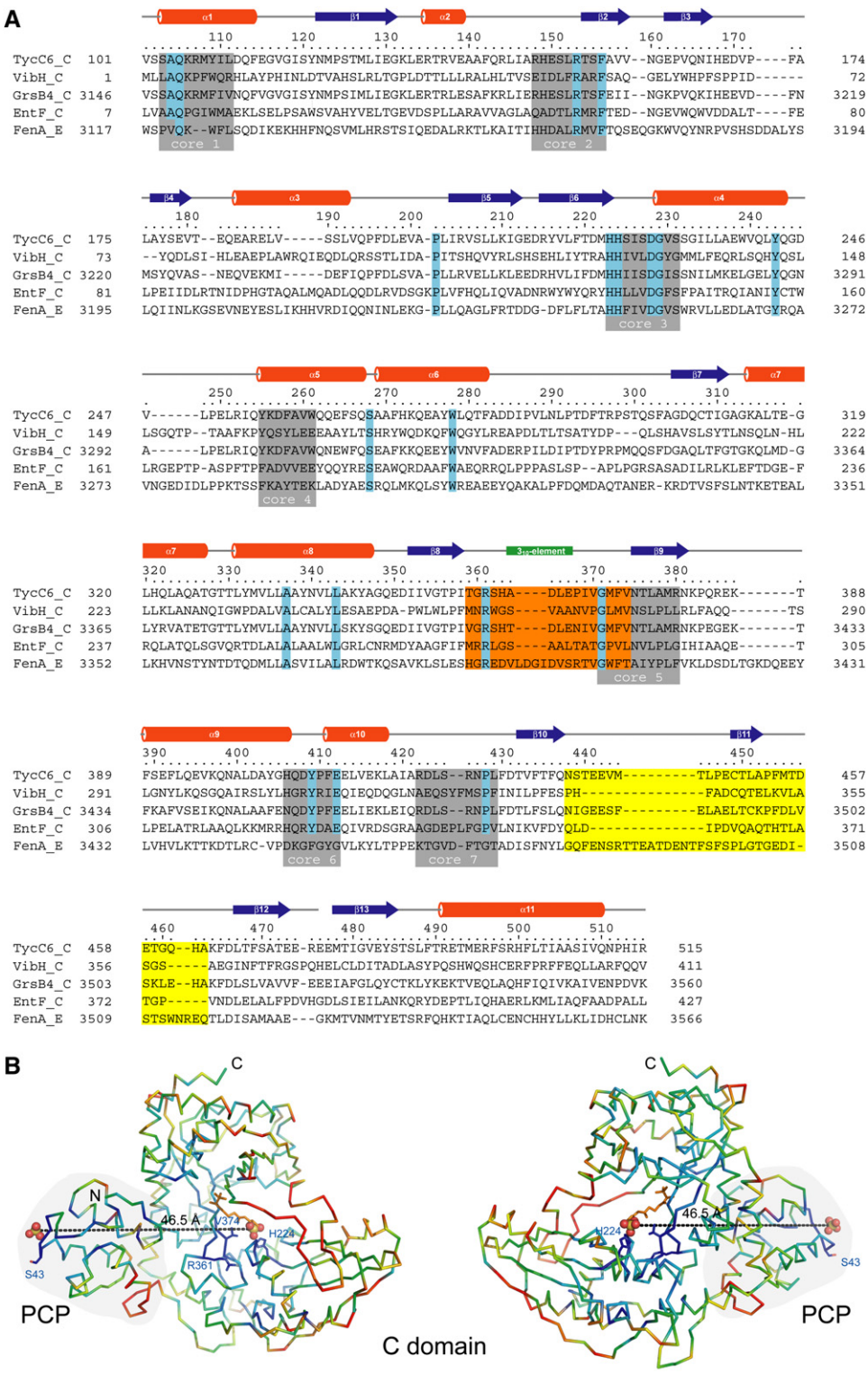

Figure 5. Sequence Alignment and Residue Conservation in the PCP-C Bidomain

(A) Sequence alignment of TycC6 condensation domain with condensation domains from other NRPS-modules. The excerpt of the master sequence alignment shows four condensation domains and one epimerase domain. The TycC6 sequence is compared to the condensation domains VibH from the vibriobactin synthetase cluster of *Vibrio cholerae* (19% sequence identity; GenBank entry: AAD48879), the fourth C domain from gramicidin S synthetase B from *Brevibacillus brevis* (64%; Genbank entry: CAA 43838), C domain from the enterobactin synthetase EntF from *E. coli*. (18%;

Please cite this article in press as: Samel et al., Structural and Functional Insights into a Peptide Bond-Forming Bidomain from a Nonribosomal Peptide Synthetase, Structure (2007), doi:10.1016/j.str.2007.05.008

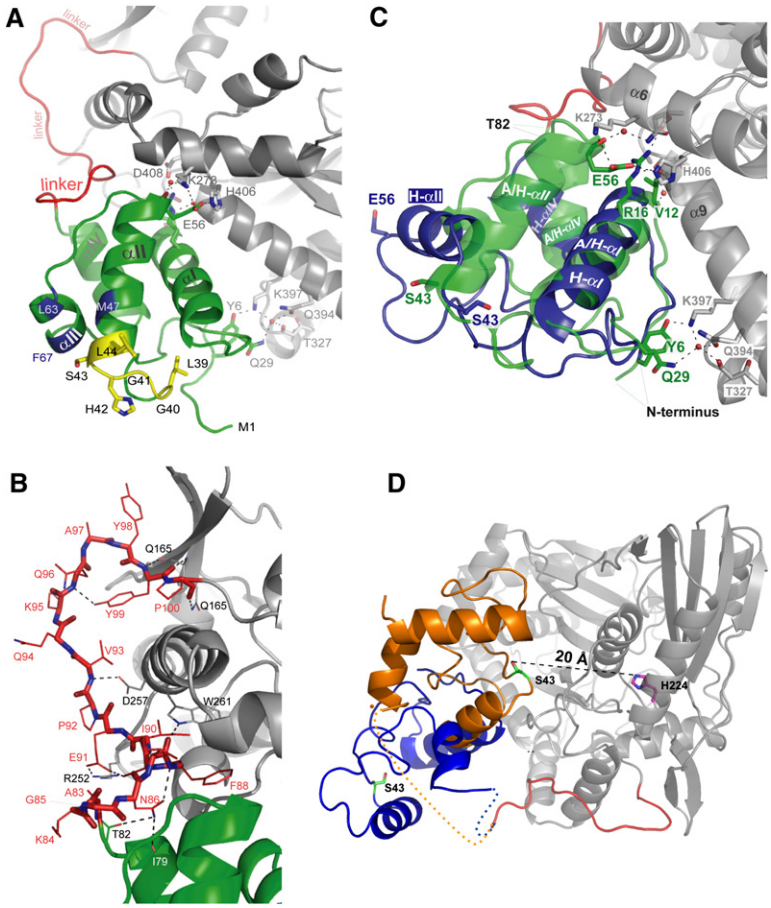

Figure 6. Interdomain Interactions in the PCP-C Bidomain

(A) Overview of the PCP domain and its interactions with the C domain. The core motif commonly found in PCP domains is highlighted in yellow, and the linker is displayed in red. Residues corresponding to those identified to be crucial for interaction with the downstream C domain by Lai et al., 2006a are colored blue.

(B) Detailed view of the 18 residue linker connecting the PCP and C domains and its interactions with the PCP domain and the C domain.

(C) Superposition of the observed A/H state and a modeled H state of the TycC5 PCP domain. As E56 is translocated by 28 Å, at least this residue would be no longer involved in interactions with the downstream C domain.

(D) Tentative model of a catalytically competent PCP-C domain interaction. The TycC5 PCP domain is shown in the modeled H state (blue, orientation as in Figure 6C; orange, orientation compatible with a positioning of the Ppan arm in the active site of the C domain).

(Q94–P100) already lack defined interactions with the C or PCP domain, and were found to be rather mobile in the PCP-C structure, as judged by high B factors.

Given the high promiscuity of PCP domains for interacting with other NRPS domains, such as the A domain, upstream and downstream C domains (Figure 1), and, optionally, cyclization or epimerization domains, it may also be possible that, instead, the observed PCP-C domain interaction is induced by crystal packing. Due to a lack of conservation for residues on the surface of the C domain, which could interact with the conserved surface patch around S43 of the PCP domain, a model for a catalytically competent PCP-C domain interaction would still be speculative. In any case, the PCP domain in either the A/H or H state would have to reverse its orientation to locate the Ppan arm, carrying residue S43 within a distance of 20 Å relative to the catalytic residue, H224 (Figure 6D). This also implies that the linker region has to exert considerable conformational flexibility to accommodate different PCP-C domain arrangements.

The Active Site Canyon of the TycC6 Condensation Domain

The substrate channel is accommodated in a canyon formed between the CAT-like subdomains. It is covered by the bridging region, including β strand 11, which is highly sequence divergent (Figure 5B), and is reminiscent of the flexible lid regions lining the active site canyons of the peptide-cyclizing thioesterase domains from NRPS (Samel et al., 2006; Bruner et al., 2002). Interestingly, a structural superposition of the carnitine acetyltransferase/CoA complex (PDB code: 1T7Q) (Hsiao et al., 2004) onto the N-terminal CAT-like subdomain of the TycC6 C domain places the pantetheine moiety in such a way that the thiol group almost coincides with the location of the sulfate ion within the active site (Figure 7A). This implies that the thioester of a peptide from the donor PCP domain would be closely positioned to the catalytically active residue, H224, which points toward the active site canyon. This residue is hydrogen bonded to the buffer-derived sulfate ion (2.87 Å), which could be interpreted

Genbank entry: AAA92015), and the C-terminal epimerase domain from the fengycin synthetase A from *Bacillus subtilis* (25%; Genbank entry: AAB80955). The numbering of the TycC6 C domain and its secondary structure elements are shown atop of each block. The core consensus sequences of the C domain are indicated by gray boxes. The residues of the "floor loop" and the "bridging region" are highlighted by orange and yellow boxes, respectively, whereas residues, which are conserved in at least the condensation domains, are highlighted blue.

(B) Mapping of sequence conservation (blue, conserved; red, variable) onto the structure of the PCP-C bidomain as viewed from the donor (left) and acceptor side (right) of the C domain. The position of the crystallographically defined sulfate ion is shown for orientation.

STFODE 1352

Please cite this article in press as: Samel et al., Structural and Functional Insights into a Peptide Bond-Forming Bidomain from a Nonribosomal Peptide Synthetase, Structure (2007), doi:10.1016/j.str.2007.05.008

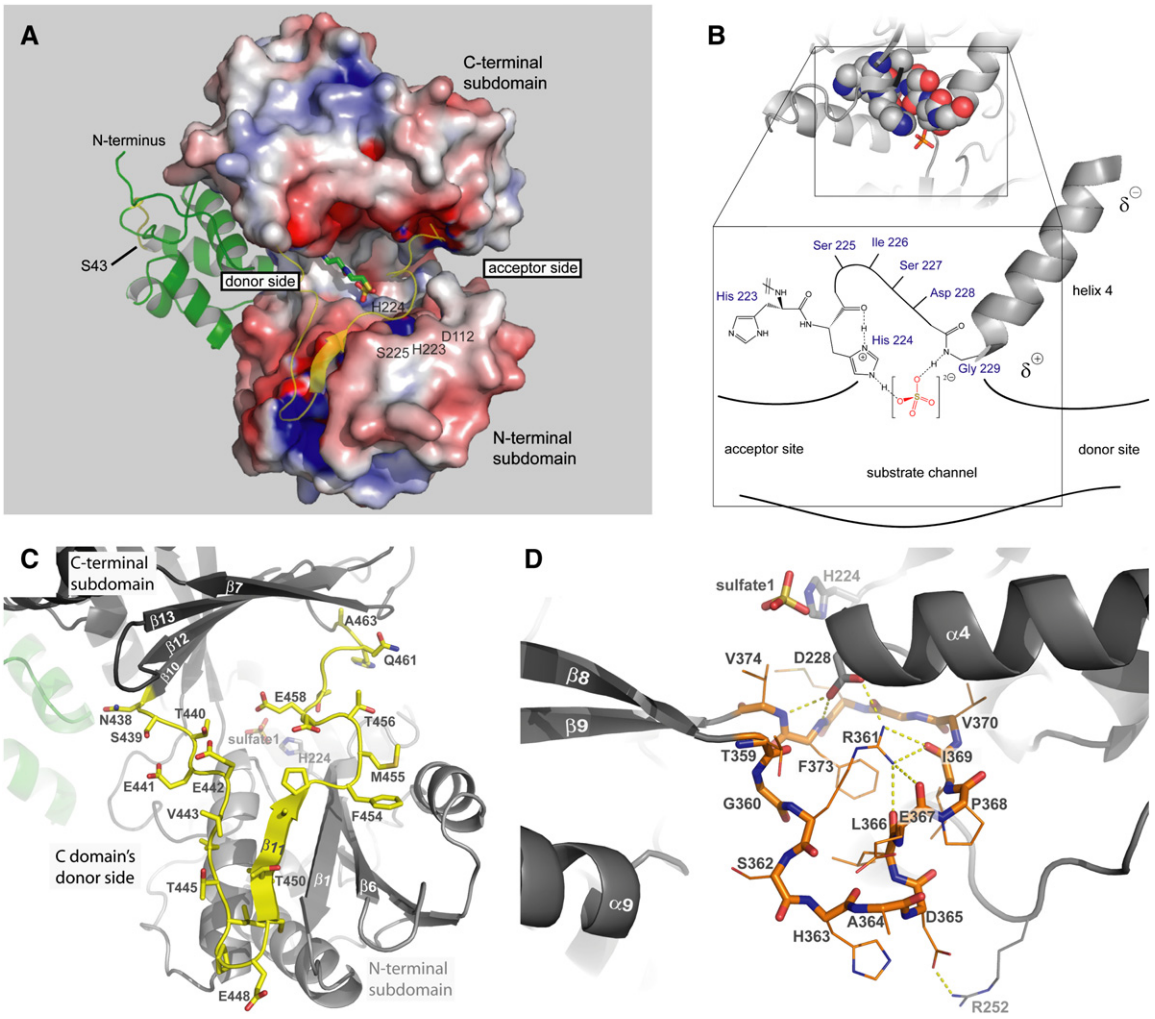

Figure 7. The Active Site of the TycC6 Condensation Domain

(A) Top view of the C domain's substrate channel. A pantetheine cofactor from the structure of carnitine acetyltransferase (Hsiao et al., 2004) (PDB code: 1T7Q, green) has been fitted into this electrostatic surface plot. The electrostatic surface representation of the TycC6 C domain was calculated by APBS (Baker et al., 2001) (blue, +5 kT/e; red, −5kT/e; ion concentration, 0.1 M). Interestingly, the terminal thiol group points toward the supposedly catalytically active residue H224 from a distance of 4 Å. Residue H223 forms hydrogen bonds with the side chains of S225 and D122. The bridging residues (Q438–A463) are omitted from the surface representation and are shown as a yellow cartoon.

(B) Schematic view of the proposed active site as seen in the crystal structure. A sulfate ion is bound to H224 and G229 via hydrogen bonds. H223 points away from the substrate channel, surrounded by helix α1 and the loop between β6 and α4. Given that H224 is the catalytically active residue, the amide NH of G229 and the dipole momentum of helix 4 might stabilize the tetrahedral anionic intermediate of the condensation reaction.

(C) The "bridging region" extends the N-terminal subdomain's β sheet and thereby covers the C domain's active site. This part of the protein exhibits a high degree of sequence variability. The stretch T456–G460 is highly flexible so that it was only modeled on stereochemical restraints with zero occupancy. In the distant center there is the sulfate ion next to the active site residue, H224.

(D) The active site's bottom is constituted by the "floor loop" residues. Selected H bond interactions comprising the conserved residues D228 and R361 are shown.

as a structural surrogate of the sp^3 intermediate of the peptide bond-forming reaction (Figure 7B). In contrast, the histidine residue of core motif 3 not essential for catalysis, H223, is turned away from the active site canyon and forms H bonds to D112 (2.95 Å) and S225 (2.71 Å).

Previous models of the reaction mechanism of NRPS condensation domains have suggested that peptide bond formation involves general acid/base catalysis with the second histidine of core motif 3, corresponding to H224, that acts as the most likely catalytic base (Bergendahl et al., 2002). Given the structure of the TycC6 C domain, pK values of residues were calculated by using the H++ server (Bashford and Karplus, 1990; Gordon et al., 2005). Interestingly, the pK of H224 was calculated to be 11.8, indicating that this residue is protonated in the active site under physiological conditions. This contradicts a suggested role as a catalytic base, unless structural changes upon substrate binding are considered

Please cite this article in press as: Samel et al., Structural and Functional Insights into a Peptide Bond-Forming Bidomain from a Nonribosomal Peptide Synthetase, Structure (2007), doi:10.1016/j.str.2007.05.008

(e.g., a relative reorientation of the CAT-like subdomains along a hinge at S268 or a closure of the active site by a collapse of the bridge-like region; Figure 7C). Kinetic studies on the C domain of the first module of TycB revealed that its condensation activity was similarly abolished in H224V mutants (Stachelhaus et al., 1998; Bergendahl et al., 2002), as described above for the unnatural cyclization reaction catalyzed by the TycC6 C domain. Keating et al. (2002), however, found that the histidine corresponding to H224 in the TycC6 C domain is not essential for the condensation reaction catalyzed by the free-standing VibH C domain, which accepts a free amine, norspermidine, instead of the α-amino thioester peptides used by the internal C domains of NRPS. Taking into account that a catalyzed deprotonation of the α-amino group in thioester peptides is not necessary, per se, due to their already low pK (~7) (Anderson and Packer, 1974), H224 might instead serve for the electrostatic stabilization of the putative, zwitterionic sp^3 reaction intermediate. Such an electrostatic stabilization of the sp^3 intermediate, which would comprise a protonated, secondary amine and an oxyanion at the carbonyl group (Yang and Drueckhammer, 2000), might be augmented by the dipole moment of helix α4. In the crystal structure (Figure 7B), this helix points toward the sulfate ion of the active site and forms a hydrogen bond (2.68 Å) with the peptide group of its N-terminal residue, G229.

The high degree of structural and sequence conservation of the floor loop connecting the two CAT-like subdomains might imply its direct involvement in peptide bond formation by NRPS C domains (Figures 7D and 5B). Noteworthy interactions between this region and the similarly conserved HHxxxDG motif are H bonds from the conserved aspartate residue, D228, to the main chain amides of F373 (2.95 Å) and V374 (3.05 Å) and a salt bridge to the conserved arginine, R361 (3.07 Å). As D228 is not suitably positioned to interact with the sulfate ion, these interactions formed by D228 are apparently crucial for the structural integrity of the active site (Keating et al., 2002).

Interestingly, such a suggested lack of importance of general acid/base catalysis during peptide bond formation was recently demonstrated for the ribosome (Bieling et al., 2006). Here, a comparison of the pH activity profiles for the peptidyltransferase reaction with aminoacyl- and hydroxyacyl-tRNA analogs revealed a broad pH independence, suggesting that proper positioning of substrates and suitable electrostatic interactions in the peptidyl transferase site of ribosomes might be sufficient for catalysis. If this holds true for ribosomes, one might expect a similar situation for NRPS C domains: The formation of a tetrahedral reaction intermediate is predicted to be more favorable for the aminolysis of thioesters (NRPS: peptidyl-4′-Ppan-PCP) than for oxoesters (ribosomes: peptidyl-tRNA), because the calculated activation energy differs by about 4.5 kcal/mol (Yang and Drueckhammer, 2000).

Conclusions

Obviously, NRPSs exert a high degree of conformational freedom to support the assembly-line synthesis of nonribosomal peptides. Particularly, the PCP domain acts as a mobile carrier of substrate intermediates between the adenylation, upstream and downstream condensation domains as well as other catalytic entities, such as the epimerization domains. Accordingly, the shown PCP-C bidomain structure will only correspond to one of the several distinct PCP-C domain arrangements; as such, more structural work on PCP-NRPS domain interactions is needed.

EXPERIMENTAL PROCEDURES

Reagents and Buffers

All solvents were purchased from Fluka and were of the highest commercially available purity. Fmoc-protected amino acids were purchased from Novabiochem. All other chemicals were purchased from Sigma-Aldrich.

Cloning of the TycC5-6 PCP-C Bidomain

The gene encoding the PCP-C bidomain from modules 5 and 6 of TycC was amplified by PCR from *B. brevis* (ATTC 8185) genomic DNA (GenBank entry: O30409) with the primers 5′-TAACCATGGTTA GATCTGAGTATGTAGCGCCGCGC-3′ and 5′-AAAGGATCCAAGCAT GTCGATCTCGCCC-3′ (restriction sites underlined). The PCR product was cloned via BamHI and NcoI into a pQE-61 vector—a derivative of the pQE-60 vector (QIAgen)—resulting in the plasmid pQE61-PCP-C, which encodes a C-terminally His_6-tagged recombinant protein of 540 amino acids. The DNA sequence is identical to the TycC gene from base G 15,346 to T 16,899. On the protein level, TycC5-6 PCP-C (here numbered from E5 to L522) corresponds to the original sequence from residue E 5116 to L 5633.

Quik-change site-directed mutagenesis (Stratagene) of the PCP-C gene utilized the primers 5′-CTCTTTACCGACATGCATGCCAGCATTT CCGATGGCG-3′ and 5′-CGCCATCGGAAATGCTGGCATGCATGTCG GTAAAGAG (H224A), 5′-CTCTTTACCGACATGCATGTCAGCATTTTC CGATGGCG-3′ and 5′-CGCCATCGGAAATGCTGACATGCATGTCGG TAAAGAG-3′ (H224V), and 5′-CTCTTTACCGACATGGCTCACAGCA TTTTCCGATGGCG-3′ and 5′-CGCCATCGGAAATGCTGTGAGCCAT GTCGGTAAAGAG-3′ (H223A, changed codons shown in lower case letters). The obtained constructs were confirmed by dideoxy sequencing.

Overproduction and Purification of the TycC5-6 PCP-C Bidomain

E. coli M15[pREP4] cells were transformed with pQE61-PCP-C and grown in LB medium at 37°C and 180 rpm to an OD_{595} of 0.6. The cultures were cooled to 28°C and induced with 0.5 mM IPTG. After 4 hr, cells were harvested and resuspended in buffer A (50 mM HEPES; 25 mM NaCl; pH 7) and then frozen in liquid nitrogen.

Frozen cells were thawed on ice and lysed with a fluidizer (Avestin) and centrifuged (JA-20; 15,000 rpm; 2 × 15 min; 4°C). The supernatant was sterile filtered before it was subjected to Ni^{2+}-NTA column purification (QIAgen), as described previously (Samel et al., 2006). Size exclusion chromatography was carried out on a Superdex 200 column (GE Healthcare). The protein eluted at an apparent mass of 67 kDa, indicating its monomeric state. Fractions were pooled and concentrated to 12 mg/ml and frozen in liquid nitrogen.

Expression of the SeMet-Labeled TycC5-6 PCP-C Bidomain

Cells were incubated in M9 minimal medium, which was supplemented with 15 μM ferrous sulfate ($FeSO_4$), at 30°C and 180 rpm. At an OD_{595} of 0.9, amino acids were added; 15 min later, L-selenomethionine was added; another 15 min later, expression was induced by IPTG at a final concentration of 0.5 mM. The expression was stopped after 5 h by harvesting, resuspending, and freezing the cells in liquid nitrogen.

STFODE 1352

Please cite this article in press as: Samel et al., Structural and Functional Insights into a Peptide Bond-Forming Bidomain from a Nonribosomal Peptide Synthetase, Structure (2007), doi:10.1016/j.str.2007.05.008

Table 2. Data Collection and Refinement Statistics

Data Collection Statistics

Dataset	Cell	Resolution (Å)	Measured, Unique Reflections	R_{merge}[a]	I/σ (I)	Completeness
Native (λ = 0.808 Å)	a = b = 84.97 Å, c = 164.97 Å	20–1.85	261739, 51834	0.043 (0.582)	20.9 (2.2)	0.989 (1.000)
SeMet (λ_1 = 0.9777 Å)	a = b = 85.09 Å, c = 164.76 Å	25–2.50	90992, 21485	0.066 (0.234)	18.1 (5.1)	0.979 (0.990)
SeMet (λ_2 = 0.9780 Å)	a = b = 85.05 Å, c = 164.69 Å	25–2.50	90919, 21484	0.050 (0.247)	18.0 (4.9)	0.980 (0.992)
SeMet (λ_3 = 0.9774 Å)	a = b = 85.17 Å, c = 164.91 Å	25–2.50	90144, 21555	0.060 (0.250)	17.9 (4.8)	0.979 (0.990)

Refinement Statistics

Dataset	Resolution	Reflections ($F > 0$)	R Value, R_{free}(%)[a]	No. of Protein Atoms, Waters, Ions and Heterogens	Rmsd Bonds (Å)	Rmsd Bond Angles, Dihedral Angles (°)
Native	19.8–1.85	49809	20.8 (26.4), 23.8 (29.7)	4160, 294, 1, 22	0.006	1.094

[a] Values in parentheses correspond to the highest resolution shell.

Synthesis of Peptidyl-Thioesters

CoA peptides were synthesized by automated solid-phase peptide synthesis in 0.1 mmol scale with an Advanced ChemTech APEX 396 synthesizer; 2-chlorotrityl resin (IRIS biotech) and a standard Fmoc coupling strategy were used as described previously (Belshaw et al., 1999). The crude product was preciptitated from 30 ml cold diethylether and subjected to preparative reverse-phase chromatography (Agilent-HPLC, C_{18}-column from Macherey-Nagel, linear gradient between 5% and 95% MeCN in water [all containing 0.1 % TFA] within 40 min). Automated fraction collection was triggered by UV absorption at 215 nm. The desired fractions were identified by MALDI-TOF, dried in vacuo, and dissolved in assay buffer (50 mM HEPES, 100 mM NaCl, pH 7.0) to a concentration of 5 mM. Purity was verified by LCMS.

TycC5-6 PCP-C Bidomain Peptide Cyclization Assays

All stock solutions used in the assays were prepared in assay buffer containing 50 mM HEPES and 100 mM NaCl at pH 7.0. *Apo*-TycC5-6 PCP-C (75 μM) was incubated with 2 μM Sfp in the presence of 10 mM $MgCl_2$ and 125 μM peptidyl-CoA in a total volume of 50 μl at 25°C for 2 hr. Reactions were stopped by adding 20 μl of 4% TFA, and subjected to LCMS analysis (C_{18} column (Macherey-Nagel), linear gradient from 10% to 60% MeCN (all containing 0.1 % TFA) within 40 min. Compounds were identified by their UV signals (210 nm) and the corresponding masses (ESI).

The connectivity of the cyclic hexapeptide **cP1** DPhe-Pro-Phe-DPhe-Asn-Gln (measured MW for **cP1-H**$^+$, 781.3664 Da; calculated MW, 781.3673 Da) was elucidated by MS-MS analysis with an API Qstar Pulsar I Q-q-TOF mass spectrometer (Applied Biosystems).

Crystallization of the TycC5-6 PCP-C Bidomain

Initial crystallization screens were performed at 19°C with a Cartesian Microsys 4004 crystallization robot. Crystals were identified from one condition (1.6 M $(NH_4)_2SO_4$; 0.1 M MES, pH 6.5; 10% dioxane). Crystals suitable for crystallographic analysis were obtained within 3 days from hanging drop setups (Hampton Research) with a protein concentration of 7.5 mg/ml and a reservoir solution containing 1.6 M ammonium sulfate, 0.04 M MES (pH 6.5), and 4% dioxane.

SeMet-labeled PCP-C was purified as described above, with the only difference being that the gel filtration was carried out in the presence of 1.5 mM DTT. The optimum crystallization condition for the SeMet derivative varied slightly from that of native PCP-C (1.6 M $(NH_4)_2SO_4$; 0.15 M MES, pH 6.2; 4% dioxane).

Data Collection and Structure Determination

The TycC5-6-PCP-C crystals were flash frozen in 1.6 M ammonium sulfate, 0.04 M MES (pH 6.5), 4% dioxane, 15% glycerol as cryoprotection buffer, and X-ray data were collected at 100K. Diffraction data of native and SeMet crystals were collected at beamline BW7A, EMBL, Hamburg, Germany. The diffraction data were indexed and integrated with DENZO (HKL Research) in space group $P4_32_12$ (a = b = 84.97 Å, c = 164.97 Å). The crystals have a calculated solvent content of 42.5% and comprise one monomer per asymmetric unit. The use of MAD data of the SeMet-labeled PCP-C crystals structure solution succeeded at 2.5 Å resolution with the Auto-RICKSHAW pipeline (Panjikar et al., 2005), identifying 13 of 15 selenium atom positions by SHELX and yielding interpretable electron density after solvent flattening by DM (CCP4, 1994) (Table 2). A first molecular model was constructed by ARP6.1 (Perrakis et al., 2001) and manually completed with COOT (Emsley and Cowtan, 2004). The resulting model was further refined by the program REFMAC5 (CCP4, 1994) until the refinement converged at an R factor/R_{free} of 20.8%/23.8% (Table 2).

Sequence Comparison and Theoretical Analysis of NRPS C Domains

A structure-based multiple sequence alignment was carried out in the absence of crystallographically determined ion positions with the program 3DCOFFEE (Notredame et al., 2000; O'Sullivan et al., 2004). At first, the TycC6 C domain was structurally aligned with the known structure of VibH by using the SAP routine (Taylor and Orengo, 1989) and subsequently compared to 103 other C domains, 13 bifunctional C/E domains, and 23 E domains.

The calculation of the pK values of (de)protonable residues in PCP-C and VibH used the quasimacroscopic model for electrostatics as implemented in MEAD (Bashford and Karplus, 1990) and interfaced via the H++ server (i.e., no explicit counter ions or water molecules were included). In short, a continuous bulk solvent region is assumed with a salinity of 0.15 M (Gordon et al., 2005). The external (internal) dielectricity constant is set to 80 (6). The electrostatic calculations

utilized the Poisson-Boltzmann model and AMBER parametrization for the molecular model. As the TycC5 PCP domain is ~23 Å away from H224 in the active center of the TycC6 C domain, pK_a values calculated for active site residues differed by less than 0.1 units in the presence or absence of the PCP domain.

ACKNOWLEDGMENTS

The authors thank Uwe Linne for the mass spectrometric analysis of PCP-C, Andrea Schmidt for support at synchrotron beamline BW7A at EMBL, DESY, Hamburg, Germany, and are very grateful to Martin Hahn for collecting C domain sequences and database work. This work was supported by grants from the Deutsche Forschungsgemeinschaft (ES152/2, MA811-14), the Graduiertenkolleg "Proteinfunktion auf atomarer Ebene," Marburg (S.A.S.), and the Studienstiftung des deutschen Volkes (G.S.).

Received: April 3, 2007
Revised: May 15, 2007
Accepted: May 16, 2007
Published: July 17, 2007

REFERENCES

Anderson, R.F., and Packer, J.E. (1974). The radiolysis of aqueous solutions of homocysteine thiolactone. Int. J. Radiat. Phys. Chem. *6*, 33–46.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. USA *98*, 10037–10041.

Bashford, D., and Karplus, M. (1990). pKa of ionizable groups in proteins: atomic detail from a continuum electrostatic model. Biochemistry *29*, 10219–10225.

Belshaw, P.J., Walsh, C.T., and Stachelhaus, T. (1999). Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis. Science *284*, 486–489.

Bergendahl, V., Linne, U., and Marahiel, M.A. (2002). Mutational analysis of the C-domain in nonribosomal peptide synthesis. Eur. J. Biochem. *269*, 620–629.

Bieling, P., Beringer, M., Adio, S., and Rodnina, M.V. (2006). Peptide bond formation does not involve acid-base catalysis by ribosomal residues. Nat. Struct. Mol. Biol. *13*, 423–428.

Bruner, S.D., Weber, T., Kohli, R.M., Schwarzer, D., Marahiel, M.A., Walsh, C.T., and Stubbs, M.T. (2002). Structural basis for the cyclization of the lipopeptide antibiotic surfactin by the thioesterase domain SrfTE. Structure *10*, 301–310.

CCP4 (Collaborative Computational Project, Number 4) (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

Clugston, S.L., Sieber, S.A., Marahiel, M.A., and Walsh, C.T. (2003). Chirality of peptide bond-forming condensation domains in nonribosomal peptide synthetases: the C5 domain of tyrocidine synthetase is a (D)C(L) catalyst. Biochemistry *42*, 12095–12104.

Conti, E., Stachelhaus, T., Marahiel, M.A., and Brick, P. (1997). Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. EMBO J. *16*, 4174–4183.

DeLano, W.L. (2002). The PyMOL molecular graphics system. (San Carlos, CA: DeLano Scientific LLC). http://www.pymol.org.

Drake, E.J., Nicolai, D.A., and Gulick, A.M. (2006). Structure of the EntB multidomain nonribosomal peptide synthetase and functional analysis of its interaction with the EntE adenylation domain. Chem. Biol. *13*, 409–419.

Emsley, P., and Cowtan, K. (2004). COOT: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Finking, R., and Marahiel, M.A. (2004). Biosynthesis of nonribosomal peptides. Annu. Rev. Microbiol. *58*, 453–488.

Gordon, J.C., Myers, J.B., Folta, T., Shoja, V., Heath, L.S., and Onufriev, A. (2005). H++: a server for estimating pKas and adding missing hydrogens to macromolecules. Nucleic Acids Res. *33*, W368–W371.

Hsiao, Y.S., Jogl, G., and Tong, L. (2004). Structural and biochemical studies of the substrate selectivity of carnitine acetyltransferase. J. Biol. Chem. *279*, 31584–31589.

Jenni, S., Leibundgut, M., Maier, T., and Ban, N. (2006). Architecture of a fungal fatty acid synthase at 5 Å resolution. Science *311*, 1263–1267.

Keating, T.A., Marshall, C.G., Walsh, C.T., and Keating, A.E. (2002). The structure of VibH represents nonribosomal peptide synthetase condensation, cyclization and epimerization domains. Nat. Struct. Biol. *9*, 522–526.

Koglin, A., Mofid, M.R., Lohr, F., Schafer, B., Rogov, V.V., Blum, M.M., Mittag, T., Marahiel, M.A., Bernhard, F., and Dotsch, V. (2006). Conformational switches modulate protein interactions in peptide antibiotic synthetases. Science *312*, 273–276.

Kuo, M.-C., and Gibbons, W.A. (1980). Nuclear Overhauser effect and cross-relaxation rate determinations of dihedral and transannular interproton distances in the decapeptide tyrocidine A. Biophys. J. *32*, 807–836.

Lai, J.R., Fischbach, M.A., Liu, D.R., and Walsh, C.T. (2006a). A protein interaction surface in nonribosomal peptide synthesis mapped by combinatorial mutagenesis and selection. Proc. Natl. Acad. Sci. USA *103*, 5314–5319.

Lai, J.R., Koglin, A., and Walsh, C.T. (2006b). Carrier protein structure and recognition in polyketide and nonribosomal peptide biosynthesis. Biochemistry *45*, 14869–14879.

Maier, T., Jenni, S., and Ban, N. (2006). Architecture of mammalian fatty acid synthase at 4.5 Å resolution. Science *311*, 1258–1262.

Marahiel, M.A., Stachelhaus, T., and Mootz, H.D. (1997). Modular peptide synthetases involved in non-ribosomal peptide synthesis. Chem. Rev. *97*, 2651–2673.

May, J.J., Kessler, N., Marahiel, M.A., and Stubbs, M.T. (2002). Crystal structure of DhbE, an archetype for aryl acid activating domains of modular nonribosomal peptide synthetases. Proc. Natl. Acad. Sci. USA *99*, 12120–12125.

Nguyen, K.T., Ritz, D., Gu, J.-Q., Alexander, D., Chu, M., Miao, V., and Brian, P. (2006). Combinatorial biosynthesis of novel antibiotics related to daptomycin. Proc. Natl. Acad. Sci. USA *103*, 17462–17467.

Notredame, C., Higgins, D.G., and Heringa, J. (2000). T-Coffee: a novel method for multiple sequence alignments. J. Mol. Biol. *302*, 205–217.

O'Sullivan, O., Suhre, K., Abergel, C., Higgins, D.G., and Notredame, C. (2004). 3DCoffee: combining protein sequences and structures within multiple sequence alignments. J. Mol. Biol. *340*, 385–395.

Panjikar, S., Parthasarathy, V., Lamzin, V.S., Weiss, M.S., and Tucker, P.A. (2005). Auto-Rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. Acta Crystallogr. D *61*, 449–457.

Perrakis, A., Harkiolaki, M., Wilson, K.S., and Lamzin, V.S. (2001). ARP/wARP and molecular replacement. Acta Crystallogr. *D57*, 1445–1450.

Samel, S.A., Wagner, B., Marahiel, M.A., and Essen, L.-O. (2006). The thioesterase domain of the fengycin biosynthesis cluster: a structural base for the macrocyclization of a non-ribosomal lipopeptide. J. Mol. Biol. *359*, 876–889.

Sieber, S.A., and Marahiel, M.A. (2005). Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics. Chem. Rev. *105*, 715–738.

Stachelhaus, T., Mootz, H.D., Bergendahl, V., and Marahiel, M.A. (1998). Peptide bond formation in nonribosomal peptide biosynthesis. J. Biol. Chem. *273*, 22773–22781.

Please cite this article in press as: Samel et al., Structural and Functional Insights into a Peptide Bond-Forming Bidomain from a Nonribosomal Peptide Synthetase, Structure (2007), doi:10.1016/j.str.2007.05.008

Stachelhaus, T., Schneider, A., and Marahiel, M.A. (1995). Rational design of peptide antibiotics by targeted replacement of bacterial and fungal domains. Science *269*, 69–72.

Taylor, W.R., and Orengo, C.A. (1989). Protein structure alignment. J. Mol. Biol. *208*, 1–22.

Walsh, C.T. (2004). Polyketide and nonribosomal peptide antibiotics: modularity and versatility. Science *303*, 1805–1810.

Weber, T., Baumgartner, R., Renner, C., Marahiel, M.A., and Holak, T.A. (2000). Solution structure of PCP, a prototype for the peptidyl carrier domains of modular peptide synthetases. Struct. Fold. Des. *8*, 407–418.

Yang, W., and Drueckhammer, D.G. (2000). Computational studies of the aminolysis of oxoesters and thioesters in aqueous solution. Org. Lett. *2*, 4133–4136.

Accession Numbers

Coordinates and structure factors of the TycC5-6 PCP-C bidomain are deposited in the RCSB protein data bank (accession number: 2JGP).

STFODE 1352