Protein templates for the biosynthesis of peptide antibiotics

Mohamed A Marahiel

Peptide synthetases of microbial origin can act as protein templates for the biosynthesis of unusual, often pharmacologically active, peptides of diverse structure and biological activity. Specific repeated modules in the synthetases each contain at least two distinct domains, required for substrate adenylation and thiolation, that define the sequence and length of the peptide product. The first crystal structure of an adenylation domain has provided insights into the mechanism of substrate recognition and activation.

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

E-mail: marahiel@ps1515.chemie.uni-marburg.de

Chemistry & Biology August 1997, **4**:561–567
http://biomednet.com/elecref/1074552100400561

Introduction

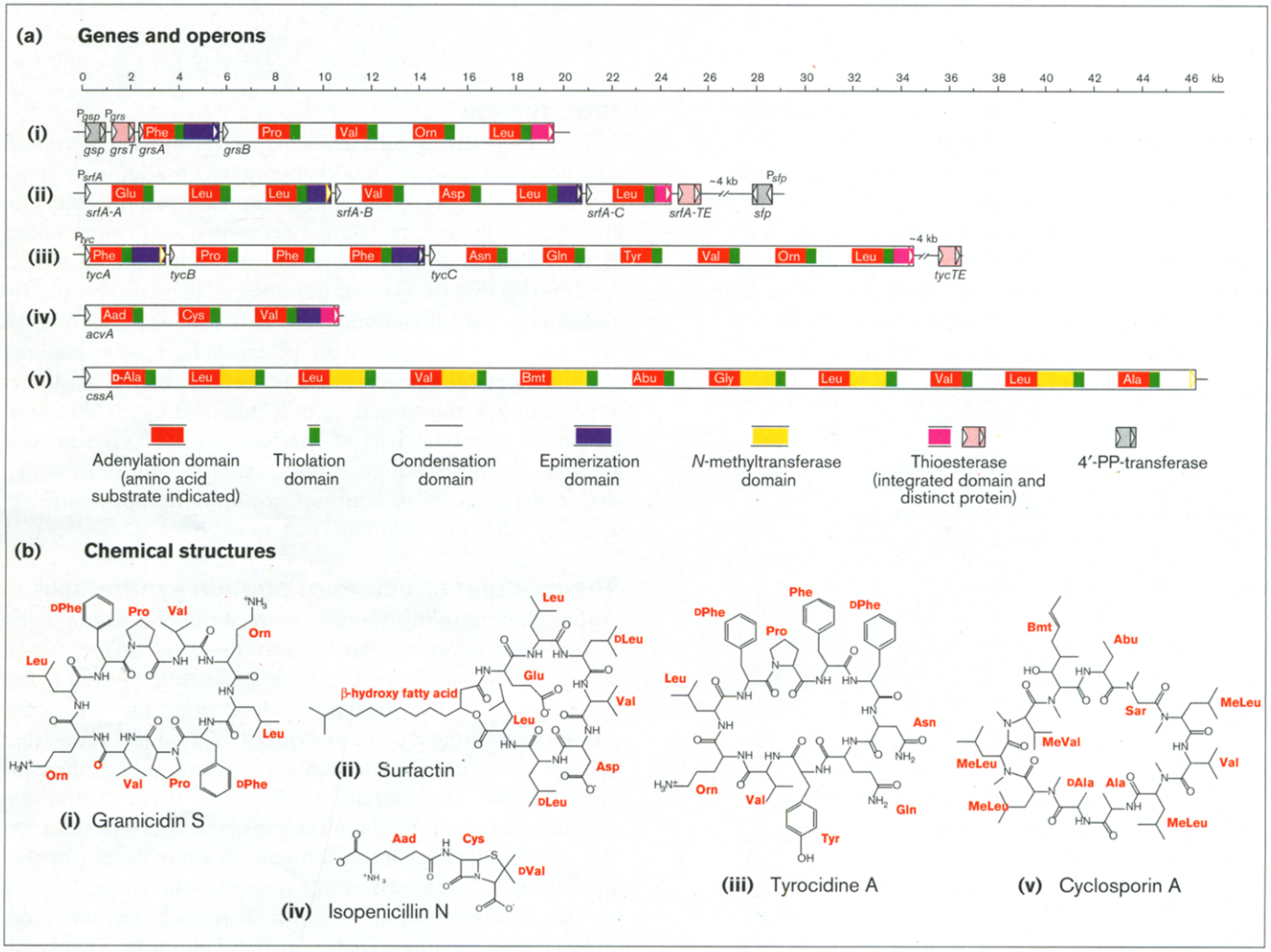
Non-ribosomally synthesized peptides form a class of microbial natural products that contain, in addition to the known proteinogenic amino acids, unusual residues that are not present in proteins [1–5]. These include D-amino acids, β-amino acids, and hydroxy- and *N*-methylated residues. Within this diverse class of peptides, depsipeptides, peptidolactones and lipopeptides, which can also be further modified by glycosylation or heterocyclic ring formation, the links between residues can be peptide bonds or can be made through the formation of lactones or esters. Because of the structural diversity of peptides in this class, many are pharmacologically active agents, ranging from antibiotic, cytostatic, antiviral or antitumor to immunosuppressive or enzyme inhibitory compounds.

The modular structure of peptide synthetases

Although diverse in structure, most of the non-ribosomally synthesized peptides share a common mode of synthesis by enzymes that employ a multiple-carrier thiotemplate mechanism. These enzymes form large multi-subunit complexes, termed peptide synthetases, which act as the protein templates for a remarkable system of peptide synthesis [5,6]. The peptide synthetases have a unique modular structure in which each module is responsible for the recognition and activation and, in some cases, for the modification (epimerization, *N*-methylation, etc.) of a single substrate residue of the final peptide product (see below). The modules, each of which is built up of several domains involved in substrate adenylation and acylation, are aligned in a sequence that is colinear with the sequence of the peptide product. With this template arrangement in peptide synthetases, the modules seem to operate independently of one another, but they act in concert to catalyze the formation of successive peptide bonds [7,8]. The reaction sequence carried out on each module starts with an ATP-dependent amino- or carboxy-acid activation and the release of pyrophosphate, followed by transfer of the acyl adenylate (acyl adenylate-AMP) to a specific thiol group and formation of a carboxy-thioester-bound intermediate. The thiol group belongs to the cofactor 4′ -phosphopantetheine (4′ -PP), which is covalently bound to each thiolation domain. The next step is the transpeptidation of the thioesterified substrates attached to the individual modules, thus affecting the amino- to carboxy-terminal stepwise assembly of the peptide chain.

Recent progress in the cloning, sequencing and biochemical characterization of an increasing number of bacterial operons and fungal gene clusters encoding modular peptide synthetases has provided valuable insights into

Figure 1



Examples of the genes and operons that encode peptide synthetases and the chemical structures of their peptide antibiotics products. **(a)** Schematic representation of some bacterial operons (*grsTAB, srfA-ABCTE, tycABCTE*) and fungal genes (*acvA; cssA*, formerly *simA*) showing the modular organization of encoded peptide synthetases. Red regions (about 550 amino acids) indicate the location of the adenylation domain within each module, whereas the green stripes (about 100 amino acids) show the thiolation (PCP) domains that have cofactor-binding sites. Purple regions define the locations of epimerization domains (about 430 amino acids) that are associated with modules incorporating D-amino acids. The D-Ala incorporating module of the *cssA* gene has no epimerization domain because an external racemase is involved in this conversion. The *cssA* gene encodes a synthetase with seven putative *N*-methyltransferase domains (about 430 amino acids, yellow regions) that coincides with the incorporation of *N*-methylated residues in the cyclic peptide and has a co-substrate *S*-adenosinemethionine (SAM)-binding motif. The white inter-module regions represent the putative condensation domains. Carboxy-terminal, module-associated thioesterase-like domains (dark pink) and those encoded by distinct genes (*grsT, srfA-TE, tycTE*; light pink) are indicated. The 4′ -phosphopantetheinyl transferase genes (*gsp, sfp*) associated with the *grs* and *srfA* operons are shown in grey. kb, kilobase; P, promoter. **(b)** The chemical structures of the peptide antibiotics **(i–v)** assembled on the corresponding peptide synthetases encoded by the genes **(i–v)** shown in (a). Orn, ornithine; Bmt, (4R)-4-[(E)-2-butenyl]-4-methyl-L-threonine; Abu, L-α-amino butyric acid; Sar, sarcosine; Aad, δ-L-α-amino adipic acid.

their molecular architecture [9–19]. Some of the first genes encoding peptide synthetases to be cloned, sequenced and analyzed are shown in Figure 1a, and the chemical structures of the peptide antibiotics produced by these peptide synthetases are shown in Figure 1b. For example, the peptide synthetases GrsA (126 kDa) and GrsB (510 kDa) have been shown to be required for the assembly of the decapeptide antibiotic gramicidin S, whereas three such multi-modular enzymes were found to be involved in the synthesis of tyrocidine and surfactin. In contrast to the operon organization of the bacterial genes (Figure 1; *grs, srfA, tyc*) encoding peptide synthetases, the fungal enzymes are encoded by large single genes, as shown for the 420 kDa δ-(L-α aminoadipyl)-L-cysteinyl-D-valine (ACV) synthetase of *Penicillium chrysogenum* and for the very large (1600 kDa) cyclosporin A

synthetase of *Tolypocladium niveum* (Figure 1, *acvA* and *cssA*, respectively) [9,14]. Alignment of the deduced amino acid sequences of peptide synthetases activating more than one amino acid residue revealed the presence of highly homologous repeating units. For example, the *tycC* gene encoding the 700 kDa tyrocidine synthetase 3, which activates and incorporates six amino acid residues (Asn–Gln–Tyr–Val–Orn–Leu) into the peptide antibiotic tyrocidine, comprises six units [18]. Similarly, the 1600 kDa cyclosporin A synthetase contains 11 distinct units on a single polypeptide chain [14]. This repetitive arrangement of functional units is the fingerprint of all peptide synthetases sequenced so far. Each of these units encompasses about 1000–1400 amino acids (120–160 kDa) and contains, irrespective of enzyme origin, several conserved signature sequences [10]. These functional units were later designated as modules after it had been shown through heterologous expression and biochemical analysis that they were able to recognize and activate individual amino acids [1,6,18]. In total agreement with this, the peptide synthetases that activate single amino acids, such as GrsA, TycA and SrfA-C (Figure 1), each contain single modules. On the basis of these studies, modules can be defined as semi-autonomous regions within peptide synthetases that carry all the information needed for recognition, activation and, in some cases, modification of a single substrate. Although the modules can act independently, they must act together, however, in a template-based mode of action to achieve a specific peptide-elongation reaction.

Recent biochemical studies and further sequence analysis of genes encoding peptide synthetases of known functions have shown that the modules can be subdivided into different functional domains [1,8,18]. For example, a minimal module of a peptide synthetase must contain at least two domains: an adenylation domain and a thiolation domain, which together form a region of about 650 amino acids. The adenylation domain (Figure 1; red regions, about 550 amino acids each) is the essential region of each module. It has been shown to bear the substrate-recognition and ATP-binding sites and is therefore solely responsible for activation of the cognate amino acid as its acyl adenylate through ATP hydrolysis. The adenylation domain, which represents the central part of each peptide synthetase module, is a member of a large superfamily of adenylate-forming enzymes that includes, in addition to peptide synthetases, the luciferases and acyl-CoA synthetases. All members of this superfamily activate their substrate carboxyl or amino acids as the corresponding acyl adenylates and have highly conserved signature sequences that have been shown to be important for ligand binding (see below).

The second functional domain needed to build up a minimal peptide synthetase module is the thiolation domain (Figure 1; green regions, about 100 amino acids each). It is also designated as the peptidyl carrier protein (PCP), by analogy to the acyl carrier protein (ACP) of fatty acid and polyketide synthases. As is the case for ACPs, this PCP domain is the site for the binding of the cofactor 4′-PP. Within the domain, 4′-PP is covalently bound to the sidechain of an invariant serine residue that is located within a highly conserved signature sequence (Figure 2). The substrate–AMP intermediates associated with the adenylation domains (Figure 2; red regions) are acylated during peptide synthesis to the cysteamine group of the 4′-PP cofactor as an activated thioester. In fact, for protein-template-directed peptide synthesis to occur, only such minimal modules comprised of an adenylation and a PCP (thiolation) domain are needed. The specific interaction of these substrate-charged minimal modules with each other promotes the amino- to carboxy-terminal peptide elongation reaction and forms the basis for the proposed multiple-carrier thiotemplate mechanism. Recently, the PCP domain of tyrocidine synthetase 1 (TycA) was expressed as a protein fragment of about 100 amino acids and shown to be catalytically active [8,20]. Its post-translational modification from inactive apo-PCP to the active holoform by the cofactor 4′-PP (see below) and its acylation by phenylalanine were shown. These studies on functionally active, but dissected, adenylation and PCP domains provide strong support for the modular structure of peptide synthetases and for the multiple-carrier thiotemplate mechanism [21].

The size of a minimal module is substantially increased when additional tailoring domains (Figure 1a; epimerization domains, purple; *N*-methylation domains, yellow) are required within a peptide synthetase module. As shown for the modules that incorporate D-amino acids (Figure 1, genes encoding GrsA, TycA, TycB, srfA-A, SrfA-B and ACV-Val) an additional domain of about 400 amino acids (Figure 1, purple regions) is attached to the carboxy-terminal end of the corresponding minimal module. Also, for modules that incorporate *N*-methylated amino acids, a domain of about 420 amino acids (Figure 1, yellow regions in the *cssA* gene, which incorporate seven *N*-methylated residues into cyclosporin A) was found to be integrated between the adenylation and thiolation domains. This tailoring domain shows a significant degree of homology to other *N*-methyltransferases and carries a signature sequence that is homologous to the *S*-adenosylmethionine (SAM)-binding site [13]. The number of epimerization or *N*-methylation domains in the synthetase coincides with the number of D-configurated or *N*-methylated residues in the product, respectively, and these domains seem to act at the stage of thioester-activated substrates.

For the interaction of the various domains and for the ordered channelling of the peptide product on a peptide-synthetase template, one would expect there to be specific intramolecular and/or intermolecular communication

between the different modules. For such productive interactions to take place, conserved regions that are located between the modules (see Figure 1; white regions, about 400 amino acids) are believed to be important [17,22]. Although no direct biochemical data concerning their function exist, these putative regions were designated as the condensation domains; they can be identified upstream of all internal adenylation domains that are involved in peptide-elongation reactions (Figure 1) and are absent from all other modules involved in peptide initiation (Figure 1, GrsA, TycA and AcvA-Aad) which strongly supports their suggested role [9,10,18]. No information is available, however, on either the mechanism of peptide elongation or the interaction of modules and how this interaction may affect the direction of polymerization. Nevertheless, during non-ribosomal peptide synthesis, all peptide intermediates remain covalently attached to the protein template throughout the elongation reaction. Also, little is known about the termination of non-ribosomal peptide synthesis, but cyclization, thioesterase-driven hydrolysis and the specific transfer of the mature peptide chain to other functional groups are all believed to be possible termination reactions.

An integrated region (about 250 amino acids) with homology to thioesterases was found to be located only at the carboxy-terminal end of several modules that are responsible for the incorporation of the last amino acid in the non-ribosomally synthesized peptide (Figure 1; dark pink regions at the carboxy-terminal modules of GrsB, SrfA-C, TycC and AcvA) . Because this domain is located at the carboxy-terminal end, one may speculate that it has a role in the hydrolytic cleavage of the mature peptide. Not all genes encoding peptide synthetases have such a domain at their carboxy-terminal ends, however (e.g. Figure 1, *cssA* which encodes cyclosporin A synthetase). In addition, one should keep in mind that such a thioesterase-like domain is present in systems producing linear peptides (tripeptide ACV), peptides branched via an ester bond (surfactin) or cyclic peptides (gramicidin S, tyrocidine), so it could serve another function; it could be an acyltransferase, as these enzymes and the thioesterases have a similar catalytic center, with the signature sequence Gly–X–Ser–X–Gly (GXSXG) [6,23].

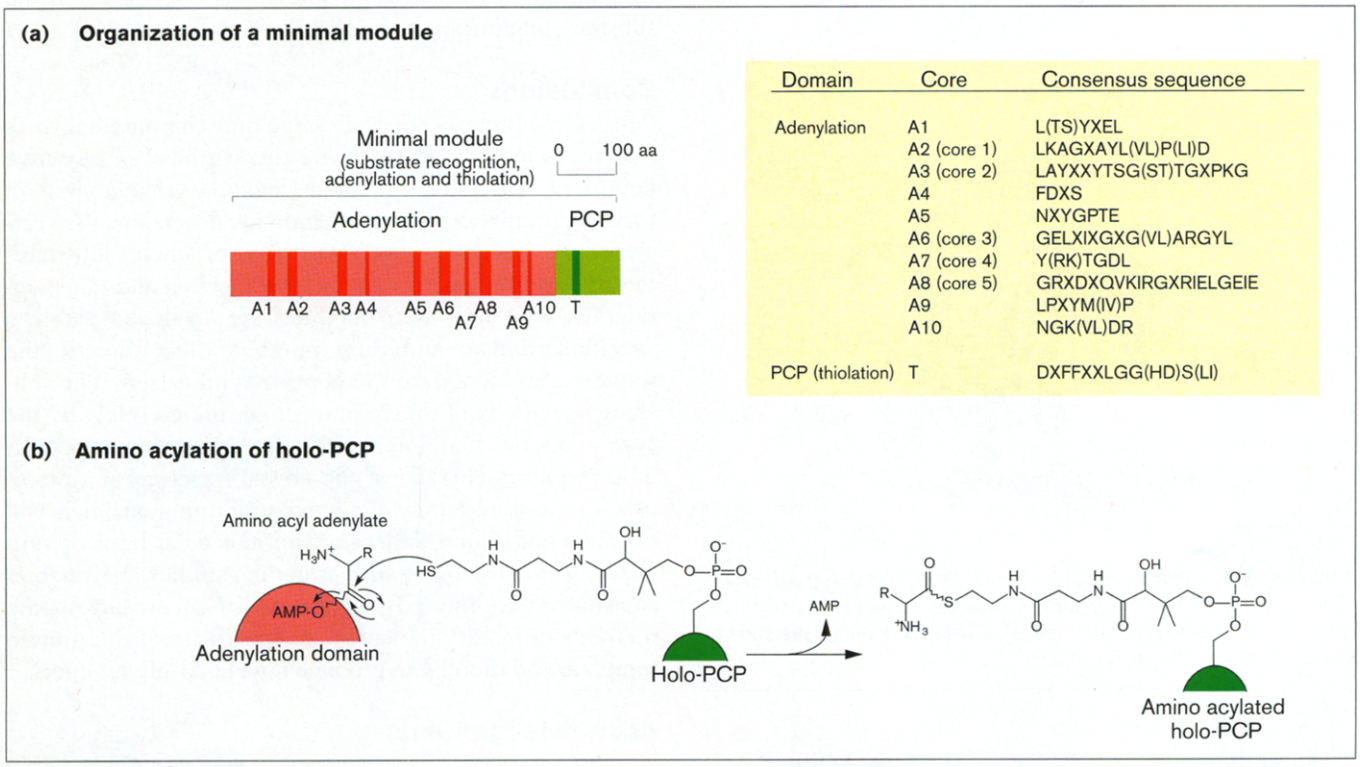
An investigation of the bacterial operons (*grs*, *srfA* and *tyc*) outside the regions encoding modular peptide synthetases revealed the presence of distinct genes (Figure 1; *grsT*, *srfA-TE* and *tycTE*, light pink), which encode a new class of thioesterases that are homologous to the integrated domains at the carboxy-terminal modules of peptide synthetases. These genes, which encode small proteins (220–340 amino acids) that are homologous to fatty acid type II thioesterases, may be located either at the 5′ end (*grsT* in the *grs* operon) or at the 3′ end (*srfA-TE* in the *srfA* operon, and *tycTE* in the *tyc* operon) of the corresponding bacterial operon [12,23]. The function of these operon-associated thioesterases in non-ribosomal peptide synthesis is still unknown, as is the case for the integrated thioesterase domains; further studies are needed to unravel their mysterious role.

The post-translational modification of peptide synthetases

Peptide synthetases, as well as other multi-enzyme complexes that function in acyl-group activation, such as fatty acid and polyketide synthases, require post-translational modification of their ACP(s) or their PCP domains for them to become catalytically active [20]. In peptide synthetases, the PCP domain is converted from the inactive apo-PCP to the active holoform by transfer of the 4′-PP moiety from CoA to the sidechain of a conserved serine residue within the PCP-pantothienylation signature sequence Leu–Gly–Gly–His/Asp–Ser–Leu (LGGH/DSL) (Figure 2). The essential role of this conserved serine residue in cofactor-binding and linking of the activated amino acid substrate as a carboxyl thioester to the 4′-PP cofactor was validated by several investigations, including site-directed mutagensis and affinity labelling studies [8,24,25]. The specific modification with 4′-PP of every PCP domain within each peptide synthetase module is catalyzed by a newly discovered enzyme superfamily, the 4′-PP transferases. They catalyze the nucleophilic attack of the conserved seryl hydroxyl on the pyrophosphate linkage of CoA, resulting in the transfer of the 4′-PP moiety onto the attacking serine and the release of 3′,5′-ADP.

Further studies on the post-translational modification of dissected PCP domains and ACP proteins *in vitro* using radiolabelled CoA led to the discovery of other proteins that catalyze the conversion of apo-ACP/PCP to their corresponding holoforms [26–28]. Among this group of 4′-PP transferases are the proteins encoded by *sfp* and *gsp*, genes that were found to be located in proximity to bacterial operons encoding peptide synthetases (Figure 1; grey). These enzymes seem to utilize CoA as a common substrate but gain specificity through protein–protein interactions. For example, the apo-ACP protein of *Escherichia coli* is post-translationally modified to the holoform only by the ACP-synthase (ACPS) and not by EntD, a second 4′-PP transferase in *E. coli* that has been shown to be specific for EntF (peptide synthetase) and EntB (isochorsmate lyase), which are needed during the synthesis of the siderophore enterobactin [20,28]. Also, the apo-PCP domain of TycA was found to be a poor substrate for ACPS and an excellent substrate for Sfp and Gsp. In agreement with these *in vitro* results is the finding that disruption of the *sfp* gene in *Bacillus subtilis*, which results in a surfactin-negative phenotype, can be rescued by supplementing the *gsp* gene *in trans* [29]. This implies that there is a specific protein partnership for the conversion of an apoprotein to its holoform and indicates that there are probably additional, as yet

Figure 2



The minimal module of a peptide synthetase and its role in the amino acylation of holo-PCP. **(a)** The adenylation (red) and PCP (thiolation; green) domains that comprise the minimal module of a peptide synthetase. The relative location and sequence of the highly conserved signature sequences found in all adenylation domains (A1–A10) of peptide synthetases are indicated; see box. The location of the signature sequence for the cofactor binding site (T) within the PCP (thiolation) domain is also shown. The signature amino acid sequences are represented by the single-letter amino acid codes. aa, amino acids. **(b)** The involvement of the adenylation and thiolation domains in the amino acylation of holo-PCP.

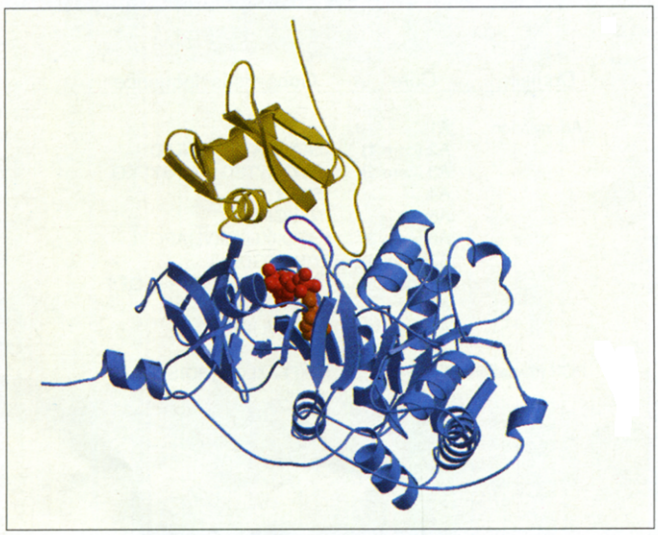
undiscovered, 4′-PP transferases associated with other multiple-4′-PP-requiring pathways.

The crystal structure of the adenylation domain PheA: a prototype for the superfamily of adenylate-forming enzymes

A large group of enzymes of diverse origin, including all adenylation domains of modular peptide synthetases and distinct proteins such as luciferases, acetyl-CoA synthases and 4-coumaryl CoA ligases, have the ability to activate their carboxylic acid or amino acid substrates to their acyl adenylates through ATP hydrolysis. They have in common a homologous domain of about 550 amino acids that has a set of highly conserved signature sequences (Figure 2a, signature sequences A1–A10 in the red region) that were shown to be of functional significance. Recently, the crystal structures have been solved for two members of this family, the firefly luciferase and the phenylalanine-adenylation domain of GrsA (PheA; Figure 1, *grs* operon) [30,31]. Although the two proteins share only 16% sequence identity and they adenylate two structurally different substrates, the overall topology of their crystal structures is very similar. One would therefore expect a very similar three-dimensional structure for all adenylation domains of peptide synthetase origin that show 30–60% sequence identity to PheA.

The PheA crystal structure, in contrast to that of firefly luciferase, was determined in the presence of the ligands AMP and phenylalanine. As shown in Figure 3, the adenylation domain of GrsA folds into two subdomains, a large amino-terminal and a small compact carboxy-terminal region with only a few interactions between them. A sandwiched layer of ordered water molecules mediates a network of indirect interactions between the two subdomains. In the PheA structure, the small carboxy-terminal domain is rotated relative to the amino-terminal domain by about 90°, seen with respect to the ligand-free structure of firefly luciferase, in which this rotation was not observed. The rotation in the PheA structure may represent a different stage in the catalytic mechanism of the adenylation reaction, because it also moves the carboxy-terminal subdomain closer (by about 5 Å) to the amino-terminal region. Although most of the residues involved in

Figure 3



Ribbon diagram of the adenylation domain PheA, showing the large amino-terminal domain (blue) and the small carboxy-terminal domain (yellow). The ligands AMP (red) and phenylalanine (orange) are shown in a space-filling representation.

substrate binding are provided by the large amino-terminal domain, an invariant lysine residue (Figure 2; core A10, Lys517) provides a key interaction in the co-ordination of the α-carboxylate group of the substrate amino acid (Figure 3). This residue is essential in adenylate-forming enzymes, as shown by site-directed mutagensis, and is located within a loop in the carboxy-terminal subdomain. The structure also reveals the role of several other invariant or highly conserved residues located within the signature sequences (Figure 2; A1–A10) of adenylation domains. Most of them were found to form interactions with AMP and with the α-amino and α-carboxyl groups of the phenylalanine substrate. The residues involved in building the substrate-binding pocket, as expected, were not found to be highly conserved. In the crystal structure of PheA, eight residues (Ala301, Ala322, Ile299, Thr278, Trp239, Ala236, Ile330 and Cys331) line the hydrophobic pocket, which is closed at the bottom by the indole ring of Trp239. The pocket is also connected with the solvent at one side by a water channel.

Taking into account the outstandingly high degree of structural similarity between the crystal structures of the PheA domain and firefly luciferase [30,31], one would expect that other homologous adenylation domains of peptide synthetases use residues located at the same positions to create their specific substrate pockets. Analysis of these corresponding residues through multiple sequence alignments of several adenylation domains [31] revealed, at least in some cases, a correlation between the polarity of the substrate sidechain and the sidechains of some residues forming the hypothetical substrate pocket. These interesting structural findings could allow specific alterations of residues in the substrate binding site to modify the enzyme activities.

Conclusions

Peptide synthetases are very large and complex enzymes that act as protein templates for the synthesis of bioactive peptides. They show a unique modular arrangement of functional units composed of individual domains that catalyze the successive condensation of their substrates through adenylation, thiolation, modification and transpeptidation. The way these modules are organized within a specific template and their number affect directly the sequence and length of the peptide synthesized. The substrate specificity of the system is determined solely by the type of adenylation domain incorporated within a module. The recent elucidation of the crystal structure of a prototype of the superfamily of adenylate-forming enzymes will enhance our understanding of the molecular basis of substrate specificity. From such structural studies, information is expected to emerge that also will help in developing methods to modify the substrate specificity of the protein template and thereby to produce novel peptide products.

Acknowledgements

I am very grateful to Peter Brick for providing Figure 3 and to Torsten Stachelhaus, Dirk Konz and Craig Townsend for critical reading of the manuscript and helpful discussions. Work carried out in my laboratory was supported by the Deutsche Forschungsgemeinschaft, EC-project (Contract-Nr. BI04-CT95-01769), Biochemie Gesellschaft m.b.H. and the Fonds der Chemischen Industrie.

References

1. Stachelhaus, T. & Marahiel, M.A. (1995). Modular structure of peptide synthetases revealed by dissection of the multifunctional enzyme GrsA. *J. Biol. Chem.* **270**, 6163-6169.
2. Stachelhaus, T., Schneider, A. & Marahiel, M.A. (1996). Engineered biosynthesis of peptide antibiotics. *Biochem. Pharmacol.* **52**, 177-186.
3. Kleinkauf, H. & von Döhren, H. (1996). A non-ribosomal system of peptide biosynthesis. *Eur. J. Biochem.* **236**, 335-351.
4. Zocher, R. & Keller, U. (1997). Thiol template peptide synthesis systems in bacteria and fungi. *Adv. Microbiol. Physiol.* **38**, 85-131.
5. Zuber, P. & A., Marahiel, M. A.(1997). Structure, function, and regulation of genes encoding multidomain peptide synthetases. *Biotechnology of Antibiotics,* (W. Strohl, ed.) pp. 187-216 Marcel Dekker, Inc., N.Y.
6. Marahiel, M.A., Stachelhaus, T. & Mootz, H.D. (1997). Modular peptide synthetases involved in non-ribosomal peptide synthesis. *Chem. Rev.*, in press.
7. Stachelhaus, T., Schneider, A. & Marahiel, M.A. (1995). Rational design of peptide antibiotics by targeted replacement of bacterial and fungal domains. *Science* **269**, 69-72.
8. Stachelhaus, T., Hüser, A. & Marahiel, M.A. (1996). Biochemical characterization of peptidyl carrier protein (PCP), the thiolation domain of multifunctional peptide synthetases. *Chem. Biol.* **3**, 913-921.
9. Smith, D.J., Alison, J.E. & Turner, G. (1990). The multifunctional peptide synthetase performing the first step of penicillin biosynthesis in *Penicillium chrysogenum* is a 421073 dalton protein similar to *Bacillus brevis* peptide antibiotic synthetases. *EMBO J.* **9**, 2743-2750.
10. Turgay, K., Krause, M. & Marahiel, M.A. (1992). Four homologous domains in the primary structure of GrsB are related to domains in a superfamily of adenylate-forming enzymes. *Mol. Microbiol.* **6**, 529-546.
11. Cosmina, P., Rodriguez, F., de Ferra, F., Perego, M., Venema, G. & van Sinderen, D. (1993). Sequence and analysis of the genetic locus responsible for surfactin synthesis in *Bacillus subtilis*. *Mol. Microbiol.* **8**, 821-831.

12. Fuma, S., *et al.*, & Yamane, K. (1993). Nucleotide sequence of 5′ portion of *srfA* that contains the region required for competence establishment in *Bacillus subtilis*. *Nucleic Acids Res.* **21**, 93-97.
13. Haese, A., Schubert, M., Herrmann, M. & Zocher, R. (1993). Molecular characterization of the enniatin synthetase gene encoding a multifunctional enzyme catalysing *N*-methyldepsipeptide formation in *Fusarium scirpi*. *Mol. Microbiol.* **7**, 905-914.
14. Weber, G., Schörgendorfer, K., Schneider-Scherzer, E. & Leitner, E. (1994). The peptide synthetase catalyzing cyclosporin production in *Tolypocladiun niveum* is encoded by a giant 45.8-kilobase open reading frame. *Curr. Genet.* **26**, 120-125.
15. Scott-Craig, J.S., Panaccione, D.G., Pocard, J.-A. & Walton, J.D. (1992). The cyclic peptide synthetase catalyzing HC-toxin production in the filamentous fungus *Cochliobolus carbonum* is encoded by a 15.7- kilobase open reading frame. *J. Biol. Chem.* **267**, 26044-26049.
16. D'Souza, C., Nakano, M.M., Corbell, N. & Zuber, P. (1993). Amino-acylation site mutations in amino acid-activating aomains of surfactin synthetase: effects on surfactin production and competence development in *Bacillus subtilis*. *J. Bacteriol.* **175**, 3502-3510.
17. de Crécy-Lagard, V., *et al.*, & Thibaut, D. (1997). Pristamycin I biosynthesis in *Streptomyces pristenaespiralis*: molecular characterization of the first two structural peptide synthetase genes. *J. Bacteriol.* **179**, 705-713.
18. Mootz, H.D. & Marahiel, M.A. (1997). The tyrocidine biosynthesis operon of *Bacillus brevis*: complete nucleotide sequence and biochemical characterization of functional internal adenylation domains. *J. Bacteriol.*, in press.
19. Merriman, T.R., Merriman, M.E. & Lamont, I.L. (1995). Nucleotide sequence of *pvdD*, a pyoverdine biosynthetic gene from *Pseudomonas aeruginosa*: PvdD has similarity to peptide synthetases. *J. Bacteriol.* **177**, 252-258.
20. Lambalot, R.H., *et al.*, & Walsh, C.T. (1996). A new enzyme superfamily - the phosphopantetheinyl transferases. *Chem. Biol.* **3**, 923-936.
21. Stein, T., *et al.*, & Morris, H.R. (1996). The multiple carrier model of non-ribosomal peptide synthesis at modular multienzymatic enzymes. *J. Biol. Chem.* **271**, 15428-15435.
22. de Crécy-Lagard, V., Marlière, P. & Saurin, W. (1995). Multi-enzymatic non-ribosomal peptide biosynthesis: identification of the functional domains catalyzing peptide elongation and epimerization. *Life Sci.* **318**, 927-936.
23. Krätzschmar, J., Krause, M. & Marahiel, M.A. (1989). Gramicidin S biosynthesis operon containing the structural genes *grsA* and *grsB* has an open reading frame encoding a protein homologous to fatty acid thioesterases. *J. Bacteriol.* **171**, 5422-5429.
24. Stachelhaus, T. & Marahiel, M.A. (1995). Modular structure of genes encoding multifunctional peptide synthetases required for non-ribosomal peptide synthesis. *FEMS Microbiol. Lett.* **125**, 3-14.
25. Gocht, M. & Marahiel, M.A. (1994). Analysis of core sequences in the D- Phe activating domain of the multifunctional peptide synthetase TycA by site-directed mutagenesis. *J. Bacteriol.* **176**, 2654-2662.
26. Lambalot, R.H. & Walsh, C.T. (1995). Cloning, overproduction and characterization of the *Escherichia coli* holo-acyl carrier protein synthase. *J. Biol. Chem.* **270**, 24658-24661.
27. Gehring, A.M., Lambalot, R., Vogel, K.W., Drueckhammer, D.G. & Walsh, C. (1997). Ability of *Streptomyces spp.* acyl carrier proteins and coenzyme A analogs to serve as substrates *in vitro* for *E. coli* holo-ACP synthase. *Chem. Biol.* **4**, 17-24.
28. Gehring, A.M., Bradley, K.A. & Walsh, T.W. (1997). Entrobactin biosynthesis in *E. coli*: EntB (isochorismate lyase) is a bifunctional enzyme that phosphopantetheinylated by EntD then acylated by EntE using ATP and 2,3-dihydroxybenzoate. *Biochemistry,* **36**, 8495-8503.
29. Borchert, S., Stachelhaus, T. & Marahiel, M.A. (1994). Induction of surfactin production in *Bacillus subtilis* by *gsp*, a gene located upstream of the gramicidin S operon in *Bacillus brevis*. *J. Bacteriol.* **176**, 2458-2462.
30. Conti, E., Franks, N.P. & Brick, P. (1996). Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes. *Structure* **4**, 287-298.
31. Conti, E., Stachelhaus, T., Marahiel, M.A. & Brick, P. (1997). Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO J.* **16**, 4174-4183.