The nonribosomal code

Hans von Döhren[1], Ralf Dieckmann[2] and Maja Pavela-Vrancic[3]

How genes are expressed and translated into proteins (using mRNA, codons and tRNAs as adaptor molecules) forms the basis of the 'genetic code'. Many peptides are synthesized nonribosomally, however, by large protein complexes that also serve as templates. Recent advances have shed light on what the nonribosomal code is and how it can be read.

Addresses: [1]TU Biochemie OE2, Franklinstrasse 29, D-10587 Berlin, Germany. [2]BrainTec GmbH, Biotechnologiepark, D-14943 Luckenwalde, Germany. [3]Faculty of Natural Sciences, Mathematics and Education, University of Split, N Tesle 12, 21000 Split, Croatia.

Correspondence: Hans von Döhren
E-mail doehren@chem.tu-berlin.de

Many pharmacologically important peptide natural products are synthesized nonribosomally in bacteria and fungi, including antibiotics, cytostatic agents and antiviral compounds. The structural diversity of these peptides is remarkably large, which is due in part to incorporation of a large number of unusual, nonproteinogenic residues. More than 300 of these residues have been identified to date, and they include D-amino acids and β-amino acids, as well as a variety of hydroxy acids and *N*-methylated acids.

Although these peptides have diverse structures, they are almost all synthesized by large, multifunctional protein complexes termed nonribosomal peptide synthetases (NRPSs), which use a multiple-carrier thiotemplate mechanism. The sequence of each peptide is encoded within the modular structure of the respective NRPS, which is why the enzymes have also been called 'protein templates'. Structurally, these complexes are organized in a modular fashion, and each module is responsible for activating and incorporating a single substrate residue of the final peptide product. A basic or minimal module consists of an amino-acid activating (adenylation) domain, an acyl carrier (thiolation) domain and a condensation domain, which is responsible for peptide-bond formation. Generally, modules are arranged colinear to the peptide sequence at the gene level. This arrangement varies from integrated structures in filamentous fungi to interacting multienzymes of varying complexity in bacteria.

The adenylation domain is probably the most important domain of each module as it recognizes and activates the appropriate residue as its acyl adenylate (using ATP to power the reaction), analogous to the activation of amino acids as aminoacyl tRNAs during ribosomal protein synthesis (Figure 1a). The genetic code, which links DNA to protein, relies on the complementarity of bases and protein–nucleic acid interactions. The nonribosomal system, in contrast, must rely on protein–protein interactions alone (Figure 1b) [1–4]. Deciphering how adenylation domains recognize their substrates is of fundamental importance to understanding and manipulating these enzyme complexes. The dissection and manipulation of the reactions performed by NRPSs has provided insight into what the nonribosomal code is and how it can be read.

Adenylation domains

The adenylation domains of NRPSs belong to a large superfamily of adenylate-forming enzymes that includes insect luciferases and acyl and aryl CoA synthetases. Despite the diversity of their origin, all of these enzymes share a homologous region of roughly 550 amino acids that contains a set of highly conserved signature sequences. The crystal structures of two family members, firefly luciferase and the phenylalanine-adenylation domain of gramicidin synthetase (GrsA), have been determined recently [5,6]. Despite these two proteins sharing only 16% sequence identity, their structures are strikingly similar, suggesting that all adenylation domains will have a similar topology. In general, these enzymes have a unique subdomain structure; the large amino-terminal and carboxy-terminal subdomains are linked by a highly conserved region. Adenylate formation involves a rotational movement of the subdomains, possibly stabilizing the adenylate structure [7].

Substrate recognition

The first step in peptide biosynthesis is selection of the substrate residue at the binding pocket of the adenylation domain. It is possible that a large part of the NRPS code operates using a 'lock and key' process dependent on the substrate specificity of the adenylation domain itself. The concept of a similar pocket architecture with conserved contact residues was proposed by Conti *et al.* [5] following the description of the phenylalanine-binding site of GrsA (see above). Modelling of 11 fairly similar binding sites in cyclosporin synthetase led Husi *et al.* [8] to suggest that only three residues are involved in determining sidechain specificity. In addition to modelling studies, conventional alignment strategies have narrowed down the specificity-determining region of this enzyme family to a stretch of about 100 amino acid residues between two highly conserved core motifs (A4

Figure 1
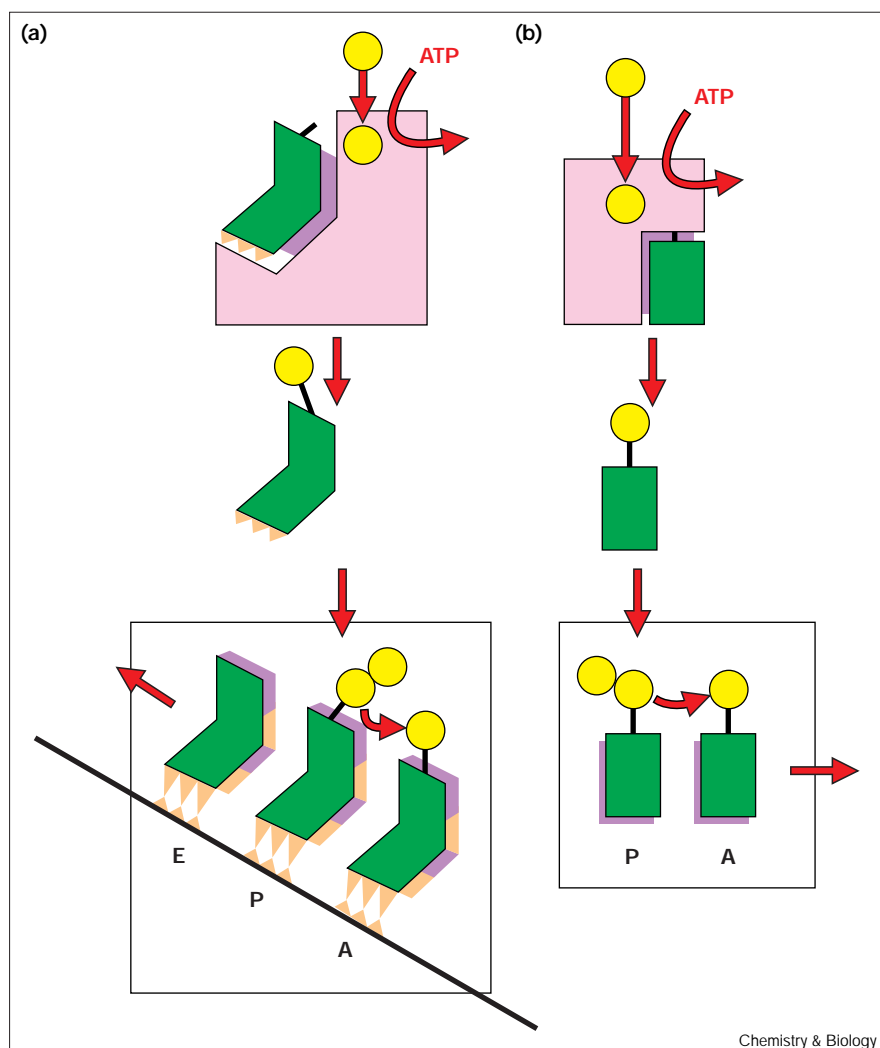
Comparison of the ribosomal and nonribosomal codes for peptide synthesis. **(a)** During ribosomal peptide synthesis, the amino acid (yellow) is activated as its aminoacyl adenylate on the aminoacyl-tRNA synthetase (pink). The cognate tRNA (green) is selected by protein–nucleic acid interactions (violet) involving the anticodon region (orange). The charged aminoacyl-tRNA then enters the ribosomal A site. (For simplicity the functions of the elongation factors have been omitted.) Selection is guided by codon–anticodon interactions with the mRNA template (shown as a black line). Peptide-bond formation occurs by peptidyltransfer to the aminoacyl residue between the P and A sites. The decharged tRNA exits the ribosome at the E site. **(b)** When a peptide is synthesized nonribosomally, the amino acid is selected at the adenylation domain of an NRPS (pink). Each adenylation domain is associated with a carrier protein domain (green; a carrier domain must be activated by addition of a 4′-phospho-pantetheine thiol group). For simplicity, covalent links between the domains are not shown. Next, the adjacent carrier domain is aminoacylated at the site of phosphopantetheinylation. Finally, peptide-bond formation is thought to proceed by simultaneous interactions of two charged carrier domains with a condensation domain, with postulated A and P sites. The process in (a) involves nucleic acid–nucleic acid interactions (orange) and protein–nucleic acid interactions (violet), whereas the process shown in (b) involves only protein–protein interactions (violet).

and A5) [9]. Examining the alignment allows tentative predictions regarding the specificity of NRPS adenylation domains to be made.

Advancing this approach, Stachelhaus et al. [9] compared the phenylalanine-binding pocket of GrsA with the corresponding sequences of 159 aminoacyl and iminoacyl adenylate-forming domains of NRPSs. Their minimizing approach used only the 10 residues that line the substrate-binding pocket (on the basis of the structure of GrsA determined by Conti et al. [5]). The results were surprising — the domains were grouped into 31 clusters of defined specificity, with an as yet unsurpassed accuracy. This achievement could now permit the sequences of nonribosomal peptides to be predicted from the translated sequences of the respective synthetases themselves. It may also permit the rational alteration of adenylation domains to be used in designing new peptides. This

'contact-residue approach' attempts to define a code for amino acid selection, using a set of amino acid residues found within the adenylation domain itself, analogous to the three anticodon nucleotides on the tRNA chain used in ribosomal protein synthesis. As more sophisticated structural definitions become available, some of the domains outgrouped in the Stachelhaus et al. [9] study could be resolved.

Adenylate formation
The rate of adenylate formation can be measured using the substrate-dependent ATP–PPi exchange reaction, which determines the amount of ^{32}P-ATP formed from ^{32}P-PPi. This reaction actually measures the stability of adenylates, rather than their rate of formation. In some of these reactions, very little ^{32}P-ATP is formed, which has puzzled researchers and led to their questioning of stable adenylate intermediates in tRNA charging. In the

non-ribosomal synthesis of gramicidin S, lysine and arginine have been incorporated [10] despite their low or undetectable rates in this isotope-exchange reaction [11]. The only available comparison of actual rates of adenylate formation and reverse ATP formation (the 4-methyl-3-hydroxy-anthranilic-acid-activating enzyme of the actinomycin synthetase) revealed no obvious correlation [12]. In recent reports, the kinetic analyses have reduced this two-substrate reaction to a simple one-substrate reaction using a fixed MgATP^{2-} concentration [9,13]. The deduced binding constants and rates are, therefore, of limited value. The relevant background of isotope-exchange kinetics has been discussed by Santi *et al.* [14].

To study the possible contribution of adenylate stability to the fidelity of nonribosomal peptide synthesis, we expressed the phenylalanine-adenylate-forming domain of tyrocidine synthetase 1, and measured ATP consumption resulting from adenylate turnover by hydrolysis. Hydrolysis rates varied considerably, which implies that the stability of noncognates or analogues is one function of product control [15]. Adenylate stability alone cannot explain the failure of, for example, 2-phenylserine to be processed, as it has an activation rate (according to the ATP–PPi exchange assay) and adenylate stability comparable to that of phenylalanine.

Proofreading and fidelity
In ribosomal protein synthesis, there are a number of safeguards that prevent misincorporation of an incorrect amino acid into the synthesized protein. It is assumed that discrimination between structurally similar amino acids (e.g. leucine and isoleucine) does not occur on the aminoacyl-tRNA synthetase surface (at least not with the overall fidelity of 10^{-4} for protein synthesis) [16]. This possible limitation has been resolved by the discovery of proofreading mechanisms in the ribosomal system [17]. Adenylates of misactivated amino acids can be hydrolytically removed, a process that is aided by the structure of aminoacyl-tRNA synthetase itself, possibly enhanced by binding of the correct cognate tRNA [18,19]. Mischarged tRNAs may be hydrolysed or finally discriminated at the peptidyl transferase level of the ribosome, as found for D-amino acids [20], whereas some *N*-methyl amino acids and even hydroxy acids are processed [21]. Taking corrective events into account, it has been estimated that 5–6 moles of ATP are consumed per peptide bond formed [22], much higher than the 1 mole expected. In addition, GTP is consumed during elongation, making protein synthesis an energetically expensive process.

A similar energy-consumption analysis was carried out for the NRPS system that synthesizes the β-lactam tripeptide precursor ACV (δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine). This analysis revealed a similar energy cost when the required substrates were present at very low concentrations

[23]. The optimized system, however, required only one mole of ATP per peptide bond, which is energetically less demanding than the ribosome. The energy waste observed in the unoptimized system can be ascribed to substrate misactivation because isolated activation domains do not efficiently discriminate between Aad (L-α-amino adipic acid) and valine [24]. If the activation domains cannot discriminate between Aad and valine, how then is the fidelity of ACV synthesis ensured? The key experiment used a didomain construct of the Aad adenylation domain with its cognate carrier protein. Although both Aad and valine were found to be activated at similar rates in the isotope-exchange assay, only Aad was detected as acid-stable thioester [25], implying that the carrier domain played a role in the selection of the correct substrate residue.

Does substrate misactivation actually occur during nonribosomal peptide synthesis? In the gramicidin S system, peptide synthesis is initiated when an activated phenylalanine bound to synthetase 1 is transferred to synthetase 2. In the absence of the initiation event, synthetase 2 is fully charged with four aminoacyl adenylates and their respective thioesters. Active-site titration of gramicidin S synthetase 2 in the presence of all four amino acids indeed yielded a value of 8.3 ATP/multienzyme, a value close to the expected ATP consumption for the formation of four adenylates and four thioesters (R. Kittelberger, M.P-V. and H.D., unpublished observations). Supplying single amino acids, however, demonstrated that misactivation occurred, indicated by consumption of up to 4 moles of ATP per amino acid supplied to the system. Such misactivations have been exploited *in vitro* to generate peptide analogues such as ^{3}Leu-gramicidin S in the absence of valine [26]. Does the NRPS system also have a proofreading function to counteract misacylations? It is worth considering the synthesis of cyclosporin, a nonribosomally synthesized undecapeptide, for some clues. Literally tonnes of cyclosporin have been processed industrially, and more than 30 analogues have been described. It is quite amazing, therefore, that three positions of this undecapeptide are occupied solely by leucine residues. One leucine site is less stringently controlled, and occasionally harbors isoleucine [27], however. How the NRPSs control the stringency of these selections remains to be solved.

Aminoacylation
Once the correct substrate residue has been selected, it becomes covalently attached to an adaptor molecule — a tRNA in the case of ribosomal protein synthesis, or an activated acyl carrier domain in nonribosomal peptide synthesis (Figure 1). NRPS carrier domains have been functionally characterized at the levels of expression, cofactor transfer and aminoacylation [28]. Structurally, they resemble the acyl carrier proteins found in polyketide synthases (PKSs) in both their autonomous and integrated forms [29]. Do NRPS carrier domains resemble cognate
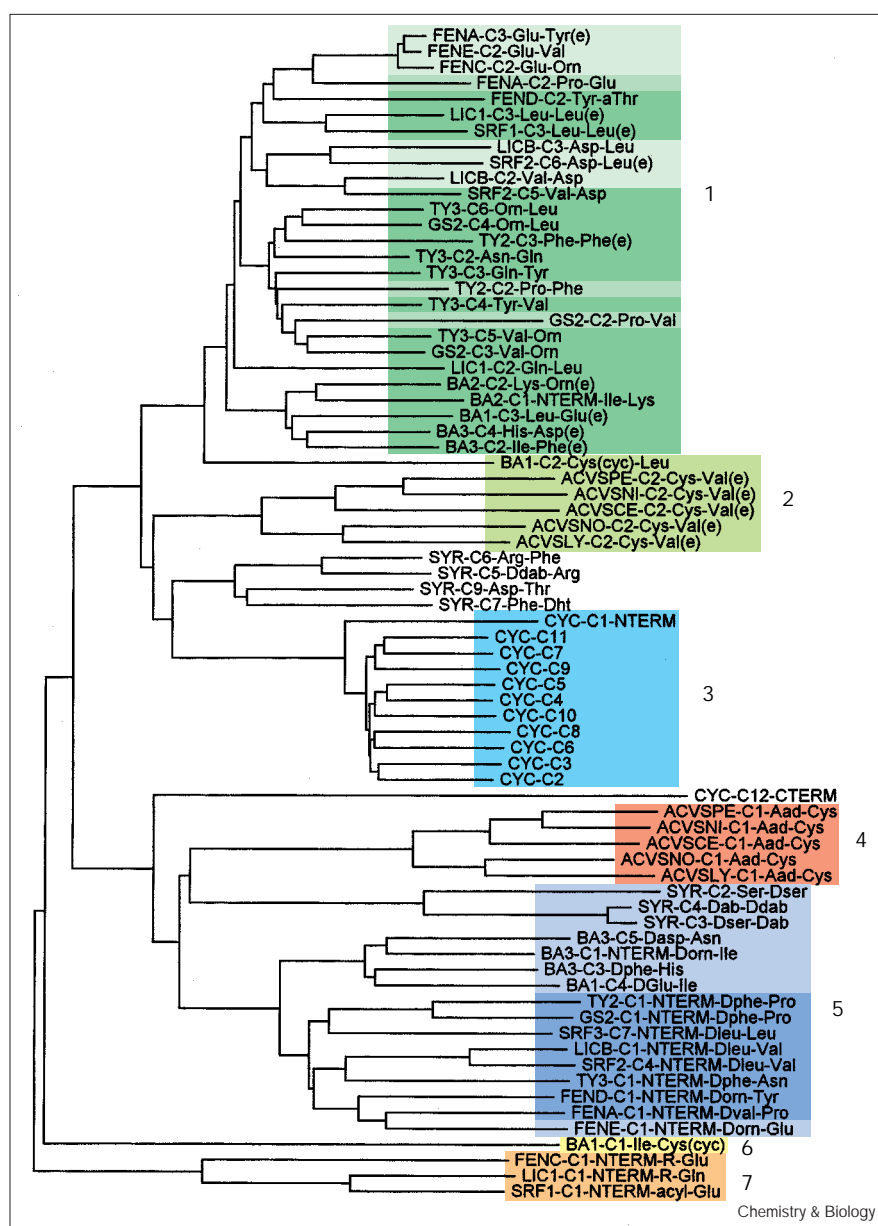
Figure 2
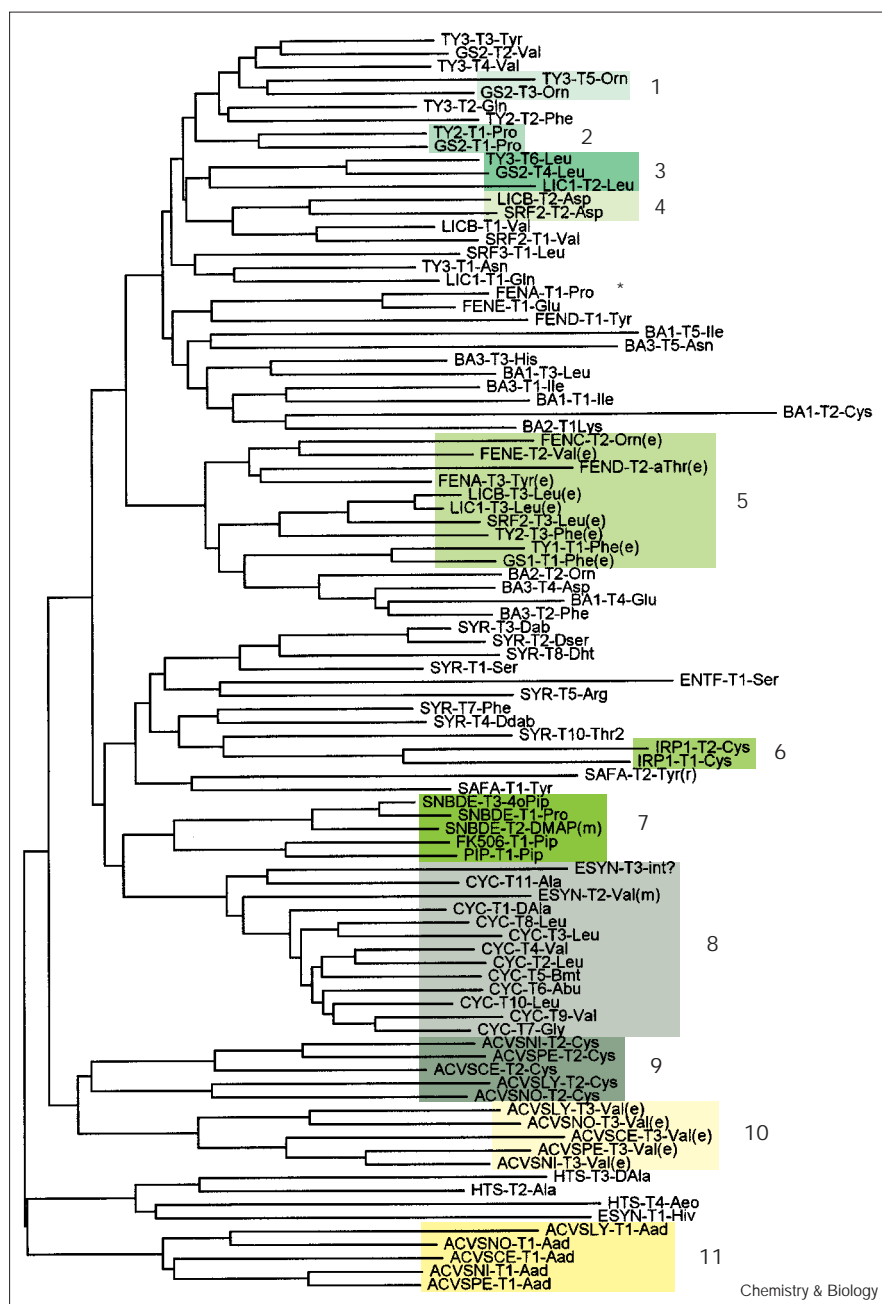
A tree construct of an alignment of NRPS carrier domain sequences. A number of clusters are seen. In the top part of the tree, small clusters (1–4) with identical acceptor specificities for ornithine, proline, leucine and asparagine are indicated. This type of residue-specific clustering is not obvious, because apparent acceptor specificities are distributed widely. In addition to proline cluster 2, an isolated proline/glycine domain (indicated by an *) is found, as well as the prominent imino acid cluster 7, closely linked to the *N*-methyl amino acid cluster 8. Carrier domains interact with modification domains, in addition to activation and condensation domains, which is reflected in some of the clustering. Carrier domains of clusters 5 and 10 are involved in epimerization reactions; their different positions in the tree could be related to their additional functions in condensation reactions (cluster 5 members) and thioester hydrolysis (cluster 10 members). All members of cluster 8 participate in the processing of *N*-methyl amino acids. It remains to be established what structural determinants promote interactions with *N*-methyltransferase domains and their respective condensation domains (see Figure 3). Cluster 11 combines the features of the first carrier domains of ACV synthetases: attaching the δ-carboxyl group of Aad, and serving the condensation domain for catalysis of a δ-carboxyl-α-amino group peptide bond.

and noncognate tRNA acceptors? This question has been studied by Walsh and colleagues [13], who investigated the transfer of activated valine from an adenylation domain to various carrier domains *in cis* and *in trans*. When separately expressed carrier domains were assayed for aminoacylation, only the carrier domain adjacent to the valine adenylation domain was efficiently charged. The transfer rates to other carrier domains were in the range 1–5%. This result indicates that there are indeed specific domain interactions involved in aminoacylation. But does

this specificity resemble a coding interaction like the aminoacyl-tRNA synthetase–tRNA match, or does it reflect a system-specific optimization? In other words, do, for example, valine-activation domains interact with valine-specific carrier domains, or do adjacent carrier domains interact only with their corresponding adenylation domains? An alignment of carrier domains led to a surprising clustering that correlates with function (e.g. carrier domains that interact with epimerization or *N*-methyltransferase domains, and even carrier domains that accept

Figure 3

Tree of a CLUSTAL alignment of 74 NRPS condensation domains. Seven clusters, according to function can be seen. Cluster 1 domains catalyze the formation of L–α–L–α–peptide bonds; cluster 2, peptides with cysteinyl groups in the donor position; cluster 3, peptide bonds involving N-methyl-amino acids; cluster 4, L–δ–L–α–peptide bonds; cluster 5, D–α–L–α–peptide bonds; cluster 6, cyclocondensation of cysteinyl residues in the acceptor position; and cluster 7, *N*-acylation reactions. Sidechain specificities of condensation reactions, as implied by the work of Belshaw *et al.* [34], were not evident in this gross alignment of 450 amino acid residues. In cluster 1 subclusterings of, for example, reactions involving acidic amino acids (glutamate/ aspartate) are obvious. In contrast, reactions involving condensation of imino acids (i.e. proline) are found in both clusters 1 and 5. Alignments were processed with TREECON 1.1 using the neighbor joining method, employing a Poisson correction in distance calculations [42].

certain amino acids were found clustered together [29]). A tree constructed using 94 carrier domains is shown, as an example, in Figure 2. Although the clustering needs to be analyzed in the context of structural data related to the acyl substrate binding and domain interactions (which are not yet available), it suggests that both substrate binding and domain interactions need to be taken into account when utilizing or designing carrier domains.

Peptide-bond formation
Peptide-bond formation requires the adjacent positioning of one aminoacyl residue and one peptidyl residue (or two aminoacyl residues during initiation). The ribosomal peptidyl transferase is a complex structure that involves both rRNA and ribosomal proteins. The NRPS condensation domain, in contrast, is composed of about 500 amino-acid residues, and contains regions that are homologous to chloramphenicol acetyltransferases and dehydrolipoamide acyltransferases [30]. The condensation domain of a system that forms the dipeptide D-Phe–Pro has been functionally identified (including analysis of point mutations in the proposed active-site region) [31,32].

In order to further analyze protein synthesis, methods for directly aminoacylating (i.e. activating) amino acids (and other types of substrate residues) have been developed;

these methods circumvent the normally required adenylation step. In order to expand the scope of residues (beyond the 20 naturally occuring amino acids) that could be incorporated into proteins ribosomally, Schultz and colleagues (reviewed in [21]) prepared a series of *N*-protected cyanomethyl esters, which provide, upon photodeprotection, unlimited access to aminoacyl-tRNAs. More than 120 analogues have been synthesized so far. After adaptation to a spare codon these analogues can be incorporated into a peptide *in vitro*, if compatible with the L–L stereochemistry of peptidyltransferase (i.e., excluding D-configured residues). In the NRPS system, Walsh and colleagues [33] took advantage of the *holo*-NRPS synthase, which activates NRPS carrier domains by adding the 4′-phosphopantetheine moiety of coenzyme A (CoA). Because acyl-CoA derivatives have been found to be substrates for this class of enzymes in charging PKS carrier domains, they expected and found that CoA derivatives of amino acids were functional in an NRPS system [34]. As aminoacyl-CoA esters are unstable (which led, in early experiments with sulfur esters, to polypeptide formation by aminolysis [35]), Belshaw *et al*. [34] used a photolabile protection group. Directly aminoacylating carrier proteins provided the first experimental evidence that NRPS condensation domains could have an editing function at the peptide-bond-forming step. Belshaw *et al*. [34] used a D-Phe module to initiate synthesis with a module activating the imino acid proline to investigate the possible specificities of the condensation domain (at the donor and acceptor positions). Because the D-Phe module contained an epimerization domain, L-amino acids introduced were found to be epimerized and stereospecificity remains an open question. The system did process an L-amino acid as an acceptor, although with a 10-fold decreased rate compared with the natural imino acid acceptor proline. These results imply that condensation domains could exhibit control over residue selection, perhaps with respect to the wider variety of acyl substrates used by NRPSs: D-amino acids, *N*-methylated amino acids, hydroxy acids and so on.

Is there evidence to support the idea of different classes of condensation domains from the available data? A simple alignment of 74 condensation domains does indeed show a clustering according to the types of reactions catalyzed: L–L condensations, D–L condensations, condensations involving *N*-methylated amino acids, *N*-acylations, reactions involving cysteinyl-peptides as donors and δ–L–α–L condensations (Figure 3). So far, classifications of condensation domains have been limited, and compilations of the condensation domain motifs available [1–4] are clearly outdated. Although the clustering approach does not relate structural data to function, it does clearly indicate that care must be taken in how NRPS domains are combined in engineering approaches. The engineering of peptide synthetases as exemplified in the pioneering work of Marahiel and colleagues [36,37] successful employed the respective didomain constructs of adenylate and adjacent carrier domains. Condensation reactions to be engineered were all of the α–L–α–L type. The reduced rate of product formation might indicate either a donor–acceptor substrate adaptation of condensation domains, or a general lack of domain interaction efficiency of the engineered systems. The importance of interactions between domains (both covalently linked and not) is convincingly demonstrated by the aminoacylation of carrier domains *in trans* [13], and how the carboxy-terminal thioesterase domain of the surfactin NRPS functions when expressed separately from the rest of the NRPS [38]. Clearly, one of the main aims of current research is to characterize the domain interactions that facilitate the fascinating organization of multistep condensations in biosynthesis [39].

Future directions

Understanding how NRPSs recognize and activate individual residues for incorporation into peptides is crucial for us to be able to engineer substrate-binding pockets of adenylation domains, and reprogram NRPSs. The nonribosomal code described by Stachelhaus *et al*. [9] could turn out to be useful in some cases. Understanding the ribosomal code has allowed, for example, the substrate specificity of a Phe-tRNA synthetase to be relaxed enough to permit ring-substituted analogues to be incorporated into proteins [40]. It would be wise, however, to remember that residue incorporation probably relies on a number of additional interactions. A similar attempt to change the binding pocket of a tRNA synthetase from glutamine to glutamate was successful in terms of amino-acid binding, but both the aminoacylation and proofreading parameters were affected [41]. When developing 'rules' for engineering NRPS adenylation domains, the influences of additional domains will have to be considered.

References
1. Konz, D. & Marahiel, M.A. (1999). How do peptide synthetases generate structural diversity? *Chem. Biol.* **6**, R39-R48.
2. Marahiel, M.A., Stachelhaus, T. & Mootz, H.D. (1997). Modular peptide synthetases involved in non-ribosomal peptide synthesis. *Chem. Rev.* **97**, 2651-2673.
3. von Döhren, H., Keller, U., Vater, J. & Zocher, R. (1997). Multifunctional peptide synthetase. *Chem. Rev.* **97**, 2675-2705.
4. Kleinkauf, H. & von Döhren, H. (1996). A nonribosomal system of peptide biosynthesis. *Eur. J. Biochem.* **236**, 335-351.
5. Conti, E., Stachelhaus, T., Marahiel, M.A. & Brick, P. (1997). Structural basis for the activation of phenylalanine in the nonribosomal biosynthesis of gramicidin S. *EMBO J.* **16**, 4174-4183.
6. 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.
7. Dieckmann, R., Pavela-Vrancic, M., Kleinkauf, H. & von Döhren, H. (1999). Probing the domain structure and ligand-induced conformational changes by limited proteolysis of tyrocidine synthetase 1. *J. Mol. Biol.* **288**, 129-140.
8. Husi, H., Schörgendörfer, K., Stempfer, G., Taylor, P. & Walkinshaw, M.D. (1997). Prediction of substrate-specific pockets in the cyclosporin synthetase. *FEBS Lett.* **414**, 532-536.

9. Stachelhaus, T., Mootz, H.D. & Marahiel, M.A. (1999). The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol.* **6**, 493-505.

10. Leenders, F. (1999). Investigation of the gramicidin S synthetase multienzyme complex for *Bacillus brevis* ATCC 999 [PhD thesis]. Technical University, Berlin.

11. Kleinkauf, H. & von Döhren, H. (1982). A survey of enzymatic biosynthesis of peptide antibiotics. In *Trends in Antibiotic Research.* pp. 220-232. Japan Antibiotics Research Association.

12. Keller, U., Kleinkauf, H. & Zocher, R. (1984). 4-methyl-3-hydroxyanthranilic acid activating enzyme from actinomycin-producing *Streptomyces chrysomallus. Biochemistry* **23**, 1479-1484.

13. Weinreb, P.H., Quadri, L.E., Walsh, C.T. & Zuber, P. (1998). Stoichiometry and specificity of *in vitro* phosphopantetheinylation and aminoacylation of the valine-activating module of surfactin synthetase. *Biochemistry* **37**, 1575-1584.

14. Santi, D.V., Webster, R.W. & Cleland, W.W. (1974). Kinetics of aminoacyl-tRNA synthetases catalyzed ATP–PPi exchange. *Methods Enzymol.* **29**, 620-642.

15. Pavela-Vrancic, M., Dieckmann, R., von Döhren, H. & Kleinkauf, H. (1999). Editing of non-cognate amino acyl adenylates by peptide synthetases. *Biochem. J.,* **342**, 715-719..

16. Pauling, L. (1957). Festschrift für Prof. Arthur Stoll, pp. 597-602, Birkhäuser Basel.

17. Arnez, J.G. & Moras, D. (1997). Structural and functional considerations of the aminoacylation reaction. *Trends Biochem. Sci.* **22**, 211-216.

18. Lin, L., Hale, S.P. & Schimmel, P. (1996). Aminoacylation error correction. *Nature* **384**, 33-34.

19. Carter, C.W., Jr. (1993). Cognition, mechanism and evolutionary relationships in aminoacyl-tRNA synthetases. *Annu. Rev. Biochem.* **62**, 715-748.

20. Yamane, T., Miller, D.L. & Hopfield, J. (1981). Discrimination between D- and L-tyrosyl transfer ribonucleic acids in peptide chain elongation. *Biochemistry* **20**, 7059-7064.

21. Mendel, D., Cornish, V.W. & Schultz, P.G. (1995). Site-directed mutagenesis with an expanded genetic code. *Annu. Rev. Biophys. Biomol. Struct.* **24**, 435-462.

22. Cramer, F., Englisch, U., Freist, W. & Sternbach, H. (1991). Aminoacylation of tRNAs as critical step of protein biosynthesis. *Biochimie* **73**, 1027-1035.

23. Kallow, W., von Döhren, H. & Kleinkauf, H.(1998). Penicillin biosynthesis: energy requirement for tripeptide precursor formation by δ-L-α-aminoadipyl-L-cysteinyl-D-valine synthetase from *Acremonium chrysogenum. Biochemistry* **37**, 5947-5952.

24. Etchegaray, A., Dieckmann, R., Kennedy, J., Turner, G. & von Döhren, H. (1997). ACV synthetase: expression of amino acid activating domains of the *Penicillium chrysogenum* enzyme in *Aspergillus nidulans. Biochem. Biophys. Res. Commun.* **237**, 166-169.

25. Schwecke, T. (1993). Studies on the biosynthesis of the β-lactam precursor δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine in *Streptomyces clavuligerus* [PhD thesis]. Technical University Berlin.

26. Altmann, M., Koischwitz, H., Salnikow, J. & Kleinkauf, H. (1978). (3,3′-Leu)-Gramicidin S formation by gramicidin S synthetase. *FEBS Lett.* **93**, 247-250.

27. Kleinkauf, H. & von Döhren, H. (1997). Applications of peptide synthetases in the synthesis of peptide analogues. *Acta Biochim. Polonia* **44**, 839-848.

28. 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.

29. Dieckmann, R. & von Döhren, H. (1997). Structural model of acyl carrier domains in integrated biosynthetic systems forming peptides, polyketides and fatty acids based on analogy to the *E. coli* acyl carrier protein. In *Developments in Industrial Microbiology* (Baltz, R.H., Hegeman, G.D. and Skatrud, P.L., eds), pp. 79-87, Society for Industrial Microbiology, Fairfax, USA.

30. De Crécy-Lagard, V., Marlière, P. & Saurin, W. (1995). Multienzymatic non ribosomal peptide biosynthesis: identification of the functional domains catalysing peptide elongation and epimerization. *C.R. Acad. Sci. Paris* **318**, 927-936.

31. Stein, T. & Vater, J. (1996). Amino acid activation and polymerization at modular multienzymes in nonribosomal peptide biosynthesis. *Amino Acids* **10**, 201-222.

32. Stachelhaus, T., Mootz, H.D., Bergendahl, V. & Marahiel, M.A. (1998). Peptide bond formation in nonribosomal peptide biosynthesis: catalytic role of the condensation domain. *J. Biol. Chem.* **273**, 22773-22781.

33. Walsh, C.T., Gehring, A.M., Weinreb, P.H., Luis, E.N. & Flugel, R.S. (1997). Post-translational modification of polyketide and nonribosomal peptide synthetases. *Curr. Opin. Chem. Biol.* **1**, 309-315

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

35. Wieland, T. (1988). Sulfur in biomimetic peptide syntheses. In *The Roots of Modern Biochemistry.* (Kleinkauf, H., von Döhren, H. & Jaenicke, L., eds) pp. 213-222, de Gruyter, Berlin.

36. 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.

37. Schneider, A., Stachelhaus, T. & Marahiel, M.A. (1998). Targeted alteration of the substrate specificity of peptide synthetases by rational module swapping. *Mol. Gen. Genet.* **257**, 308-318.

38. Guenzi, E., Galli, G., Grgurina, I., Pace, E., Ferranti, P. & Grandi, G. (1998). Coordinate transcription and physical linkage of domains in surfactin synthetase are not essential for proper assembly and activity of the multienzyme complex. *J. Biol. Chem.* **273**, 14403-14410.

39. Gokhale, R.S., Tsuji, S.Y., Cane, D.E. & Khosla, C. (1999). Dissecting and exploiting intermodular communication in polyketide synthases. *Science* **284**, 482-485.

40. Agou, F., Quevillon, S., Kerjan, P. & Mirande, M. (1998). Switching the amino acid specificity of an aminoacyl-tRNA synthetase. *Biochemistry* **37**, 11309-11314.

41. Ibba, M. & Hennecke, H. (1995). Relaxing the substrate specificity of an aminoacyl-tRNA synthetase allows *in vitro* and *in vivo* synthesis of proteins containing unnatural amino acids. *FEBS Lett.* **364**, 272-275.

42. Van de Peer, Y. & de Wachter, R. (1994) TREECON for windows: a software package for the construction and drawing of evolutionary trees for the microsoft windows environment. *Comput. Appl. Biosci (CABIOS)* **10**, 569-570.