

Exploitation of the Selectivity-Confering Code of Nonribosomal Peptide Synthetases for the Rational Design of Novel Peptide Antibiotics[†]

Katrin Eppelmann, Torsten Stachelhaus, and Mohamed A. Marahiel*

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

Received April 10, 2002; Revised Manuscript Received May 28, 2002

ABSTRACT: Recently, the solved crystal structure of a phenylalanine-activating adenylation (A) domain enlightened the structural basis for the specific recognition of the cognate substrate amino acid in nonribosomal peptide synthetases (NRPSs). By adding sequence comparisons and homology modeling, we successfully used this information to decipher the selectivity-confering code of NRPSs. Each codon combines the 10 amino residues of a NRPS A domain that are presumed to build up the substrate-binding pocket. In this study, the deciphered code was exploited for the first time to rationally alter the substrate specificity of whole NRPS modules in vitro and in vivo. First, the single-residue Lys239 of the L-Glu-activating initiation module C–A_{Glu}–PCP of the surfactin synthetase A was mutated to Gln239 to achieve a perfect match to the postulated L-Gln-activating binding pocket. Biochemical characterization of the mutant protein C–A_{Glu}–PCP(Lys239 → Gln) revealed the postulated alteration in substrate specificity from L-Glu to L-Gln without decrease in catalytic efficiency. Second, according to the selectivity-confering code, the binding pockets of L-Asp and L-Asn-activating A domains differs in three positions: Val299 versus Ile, His322 versus Glu, and Ile330 versus Val, respectively. Thus, the binding pocket of the recombinant A domain AspA, derived from the second module of the surfactin synthetases B, was stepwisely adapted for the recognition of L-Asn. Biochemical characterization of single, double, and triple mutants revealed that His322 represents a key position, whose mutation was sufficient to give rise to the intended selectivity-switch. Subsequently, the gene fragment encoding the single-mutant AspA(His322 → Glu) was introduced back into the surfactin biosynthetic gene cluster. The resulting *Bacillus subtilis* strain was found to produce the expected so far unknown lipoheptapeptide [Asn⁵]surfactin. This indicates that site-directed mutagenesis, guided by the selectivity-confering code of NRPS A domains, represents a powerful alternative for the genetic manipulation of NRPS biosynthetic templates and the rational design of novel peptide antibiotics.

In the past decades, many pathogenic bacteria developed potent mechanisms to withstand contemporary antibiotics (1), a circumstance that focused attention on the need for discovering and developing new antimicrobial agents. A promising strategy to approach this challenge has been the manipulation of biosynthetic pathways of certain antibiotics, which may lead to novel antibiotics with altered or enhanced bioactivities against resistant pathogens. Some classes of naturally occurring antibiotics have been found to be particularly susceptible for such approaches. Here, especially the large group of nonribosomal peptides (NRPs) has to be emphasized. A remarkable variety of pharmacological activities arises from these natural peptides, which include the β -lactam precursor ACV (leading to penicillins and cephalosporins), the immunosuppressant cyclosporin, and the

antitumor agent bleomycin, as well as vancomycin, the “antibiotic of last resort”.

Recent genetic, biochemical, and structural studies afford an insight into the biosynthesis of NRPs (2, 3), which are assembled by large multimodular protein templates, termed nonribosomal peptide synthetases (NRPSs; see Figure 1). Each module represents a functional building unit required for recognition and activation of one constituting amino acid, as well as its incorporation into the nascent peptide product. These different tasks are accomplished by a set of distinct catalytic domains: an adenylation (A) domain selects the substrate amino acids and generates the corresponding aminoacyl adenylate. Subsequently, the aminoacyl moiety is covalently tethered to a prosthetic phosphopantetheinyl (Ppant) group, which has been posttranslationally introduced into the adjacent peptidyl carrier protein (PCP) domain. Peptide-bond formation and chain translocation occur each time a peptidyl-S–Ppant donor is attacked by a monomeric aminoacyl-S–Ppant nucleophile under the catalytic control of a condensation (C) domain. The resulting peptide product can be further functionalized by the modification of incorporated monomers (epimerization or *N*-methylation) and is released by the action of a thioesterase-like (Te) domain, normally fused to the terminal module.

[†] This work was supported by the Deutsche Forschungsgemeinschaft, the Fonds der chemischen Industrie, and the Federal Ministry of Education, Science, Research and Technology.

* To whom correspondence should be addressed. Tel.: +49+6421-282-5722. Fax: +49+6421-282-2191. E-mail: marahiel@chemie.uni-marburg.de.

¹ Abbreviations: NRP(s), nonribosomal peptide(s); NRPS(s), nonribosomal peptide synthetases(s); A domain, adenylation domain; C domain, condensation domain; PCP domain, peptidyl carrier protein domain; Te domain, thioesterase-like domain; Ppant, cofactor 4'-phosphopantetheinyl.

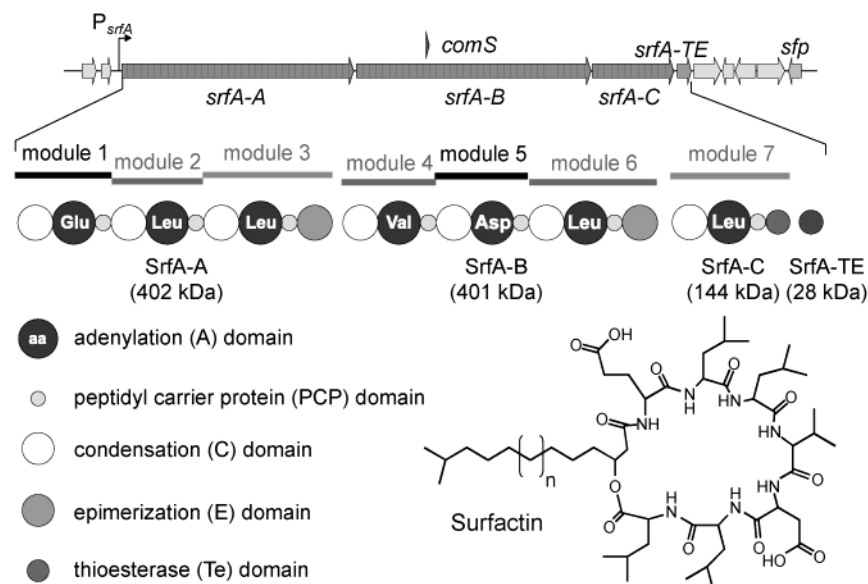


FIGURE 1: Schematic diagram of the surfactin biosynthesis operon (*srfA*) from *B. subtilis* ATCC 21332. Three NRPSs and a putative type II thioesterase are encoded by the genes *srfA-ABCTE* and act in concert for the stepwise assembly of the cyclic lipopeptide. The synthetases SrfA-ABC are composed of three, three, and one module(s), respectively, which can be further subdivided into functional domains. Substrates are recognized and adenylated by A domains and subsequently covalently tethered to the thiol group of cofactor 4'-Ppant, which has been posttranslationally introduced onto each PCP domain by the 4'-Ppant-transferase Sfp. C domains catalyze peptide-bond formation and chain translocation between the nascent peptidyl-S-Ppant intermediates and the downstream monomeric aminoacyl-S-Ppant. At positions 3 and 6, epimerization (E) domains convert L-Leu moieties into the D-isomer, and at module 7 a terminal Te domain releases the final product, surfactin.

Given the modular organization of NRPSs, primary structure, size, and complexity of a synthesized peptide are dictated by the number and organization of iterated modules and domains (Figure 1). Thus substitution, deletion, or introduction of domains or whole modules within a NRPS biosynthetic template should lead to targeted alterations within the corresponding peptide product. On the basis of this conception, various approaches had been attempted to rationally design novel peptide products. For instance, short peptides have been made *in vitro* by recombination of whole modules to build up short di- and trimodular hybrid NRPSs from scratch (4, 5). In contrast, engineered manipulation of entire NRPS biosynthetic systems was demonstrated *in vivo*, using the lipopeptide surfactin from *B. subtilis* as an example (see Figure 1). Here, predicted alterations in the amino acid composition of the lipopeptide could be accomplished by targeted replacement of minimal (A-PCP) and whole modules (C-A-PCP) (6–8). Such exchanges of modules led to the formation of the predicted peptide products, however, in most cases at the expense of product yield. This was mainly accounted to obstructions of the intermodular communication and disturbances in the tertiary structure of the hybrid NRPSs.

An approach to overcome these bottlenecks has become feasible after the structural basis for the specific recognition of the substrate amino acid in NRPSs has been enlightened by the crystal structure of the phenylalanine-activating A domain (PheA) (9). The structure revealed that only 10 residues of the 55 kDa A domain are actually involved in binding of the cognate substrate amino acid. The conserved residues Asp235 and Lys517 mediate electrostatic key interactions with the α -amino and α -carboxylate group of the substrate, while the remaining eight amino acids (Ala236, Trp239, Thr278, Ile299, Ala301, Ala322, Ile330, and Cys331) are interacting with the side-chain. By cross-comparing the

residues that line the phenylalanine-specific binding pocket of PheA with the corresponding moieties in other A domains, it had been postulated that the 10 residues have different impact on the A domain's selectivity (10). Two residues (Asp235 and Lys517) are (nearly) invariant and facilitate electrostatic key interactions with the α -amino and α -carboxylate group of the substrate. Three residues (pos. 236, 301, and 330) are only moderately variable, presumably possessing only modulating influence on the enzyme's activity and selectivity. Three residues (pos. 239, 322, and 331) are highly variable and likely responsible to mediate A domain selectivity. Two additional residues (pos. 278 and 299) are also highly variable but have been considered "wobble"-like positions. In over 40% of all known NRPS codons, these latter residues have been found to be flexible without affecting the A domain's specificity. The *in silico* studies were accompanied by the mutation of selectivity-conferring residues within the binding pocket of PheA. Several of these mutants revealed an alteration or relaxation of substrate specificity. Generalization of the specificity determinants also allowed the targeted selectivity-switch of a L-Asp-activating A domain, whose crystal structure had not yet been solved (10). All these studies led to the discovery of the selectivity-conferring code of A domains and confirmed the assumptions made about the different impact of residues, building up the corresponding substrate binding pocket.

In this study, the selectivity-conferring code of NRPSs was exploited for the first time to rationally alter the substrate selectivity of whole NRPS modules *in vitro* and *in vivo*. First, site-directed mutagenesis was used to rationally alter the substrate selectivity of the initiation module C-A_{Glu}-PCP of the surfactin synthetase complex from L-Glu to L-Gln. In a second set of experiments, a novel derivative of the lipopeptide antibiotic surfactin, containing a L-Asn

residue at position 5 in place of the native L-Asp constituent, was rationally designed in the same manner.

EXPERIMENTAL PROCEDURES

Bacterial Strains and Culture Media. Bacterial strains were grown in 2xYT medium (11), Difco sporulation medium (DSM) (12), or modified SpII medium (13). Cultures of *B. subtilis* were supplemented with 5 μ g/mL chloramphenicol and/or 15 μ g/mL neomycin (final concentrations). For *E. coli* cultures, final concentrations of 25 μ g/mL kanamycin and 100 μ g/mL ampicillin were used.

Plasmid Construction. If not indicated otherwise, NRPS gene fragments were amplified from chromosomal DNA of *B. subtilis* ATCC 21332. PCR amplification was performed using the "Expand long template PCR system" (Roche, Mannheim, Germany) following the manufacturer's protocol. Mutations (italic) as well as restriction sites (bold) for subsequent cloning were introduced with oligonucleotides, purchased from MWG-Biotech (MWG-Biotech, Ebersberg, Germany). PCR products were purified with the QIAquick-spin PCR purification kit (Qiagen, Hilden, Germany). Standard procedures were applied for all DNA manipulations (11), as well as preparation of the recombinant plasmids.

The 3.14 kb DNA fragment encoding the glutamate-activating initiation module of *srfA-A* was amplified using the oligonucleotides 5'-CA_{Glu}PCP (*Eco*RI) 5'-CTT **AGA ATT** CTT AGG GGA GGT ATG ACA-3' and 3'-CA_{Glu}PCP (*Bam*HI) 5'-TTC TCT GTG ATC **GGA TCC** CTC TGC AAG-3'. The DNA fragment was treated with the restriction enzymes *Eco*RI and *Bam*HI and subsequently ligated into the His₆-tag expression vector pQE60 (Qiagen, Hilden, Germany), previously cut in the same manner to give pCA_{Glu}-PCP. The mutant of C-A_{Glu}-PCP was constructed by site-directed mutagenesis of pCA_{Glu}PCP using the QuickChange Site-Directed Mutagenesis Kit (Qiagen, Hilden, Germany), in accordance with the manufacturer's protocol using the oligonucleotides 5'-CA_{Glu}PCP(K239Q) 5'-CCA CTT TGA TGC GTC AGT GCA GCA AAT CTT CGC GTC G-3' and 3'-CA_{Glu}PCP(K239Q) 5'-CGA CGC GAA GAT TTG CTG CAC TGA CGC ATC AAA GTG G-3'. Integrity of the both plasmids, pCA_{Glu}PCP and pCA_{Glu}PCP(K239Q), was confirmed by DNA sequencing using the ABI prism 310 Genetic Analyzer (ABI, Weiterstadt, Germany).

AspA double and triple mutants were constructed by site-directed mutagenesis of plasmid pAspA(H322E) (10) using inverse PCR. For the introduction of the mutation I330V, the oligonucleotides 5'-AspA(I330V) 5'-TAC **GGG CCC** ACA GAA GCA ACG GTC GGC-3' and 3'-AspA(I330V) 5'-CTG **TGG GCC** CGT ACT CAT TGA TAA ATT CGG-3' were used. The amplified DNA fragment was terminally modified using the restriction enzyme *Apa*I. The mutation V299I was introduced using the oligonucleotides 5'-AspA-(V299I) 5'-AAC **TTC GAA** TCC TTG CGC TTG ATC ATC TTG-3' and 3'-AspA (V299I) 5'-GGA **TTC GAA** GTT CGC ATC TTT TGC AAA ACT-3'. The DNA fragment obtained was digested using the restriction enzyme *Bst*BI. The terminally modified DNA fragments were digested with *Dpn*I to get rid of the parental template DNA. After purification, DNA fragments were ligated and used for the transformation of *E. coli* XL1-blue (14). The plasmids pAspA(H322E,I330V), pAspA(H322E,V299I), and pAspA-

(H322E,V299I, I330V) could be obtained, and integrity of all constructs was confirmed by DNA sequencing.

For the construction of the *srfA-B2*-deletion plasmid, the 1.09 kb-fragment comprising the 3'-homologous region 3'SRFb was amplified using the oligonucleotides 5'-3'SRFb- (*Bam*HI) 5'-CGC **GGA TCC** AAG GAG AGC TTG ATA TTT GGC-3' and 3'-3'SRFb(*Eco*RI) 5'-CCG **GAA TTC** GCT CCA TTT TTC AGC TGT TTC-3'. The fragment was digested with *Eco*RI and *Bam*HI and ligated into pJLA503 (15) previously cut in the same manner to give pJLA/3'SRFb. The 1.10 kb-fragment comprising the 5'-homologous region 5'SRFb was amplified using the oligonucleotides 5'-5'SRFb- (*Nde*I) 5'-CGC **CAT ATG** ATC GAC GGA GAG GTT GTG C-3' and 3'-5'SRFb(*Nru*I) 5'-AGC **TCG CGA** GTC TTT CTT CTA GCA AAG CAG-3'. The DNA fragment was digested with *Nde*I and *Nru*I and ligated into pOEX5A (16) previously cut in the same manner to give pOEX/5'SRFb. Digestion of pOEX/5'SRFb with *Nde*I and *Bam*HI yielded a 1.82 kb DNA fragment (comprising 5'SRFb), which was ligated into pJLA/3'SRFb previously cut in the same manner to give pJLA/5'-3'SRFb. Finally, the resistance cassette *cat* was obtained from pDG268 (17) by digestion with *Bam*HI and *Eco*RV and ligated into pJLA/5'-3'SRFb previously cut with the endonucleases *Nru*I and *Bam*HI to give the *srfA-B2*-deletion plasmid pSRFb Δ Asp.

The 4.02 kb-fragment containing the 5' and 3' homologous regions of the aspartate-activating A domain of *srfA-B2* was amplified using the oligonucleotides 5'-homoAsp(*Cla*I) 5'-TAA **ATC GAT** GGA GGC TGC CAA GG-3' and 3'-homoAsp(*Spe*I) 5'-TAA **ACT AGT** CAG TAA ATC CGC CCA GT-3'. The DNA fragment was digested with *Cla*I and *Spe*I and ligated into pKE19 (18) previously cut in the same manner to give phomoAspA. Exploiting the intrinsic restriction sites *Eco*RI and *Pst*I, the gene fragment encoding the mutated aspartate-binding pocket was obtained from pAspA-(H322E) and ligated into phomoAspA previously cut in the same manner to give the integration plasmid phomoAspA-(H322E).

Bacillus subtilis Strain Construction. Transformation of *B. subtilis* strains was performed as described by Klein et al. using 10 ng of linearized plasmid DNA (19). The transformants were selected on solid DSM supplemented with the appropriate antibiotic(s). Gain or loss of a selectable marker was verified by replica pecking.

Transformation of *B. subtilis* ATCC 21332 (20) with the *srfA-B2* disruption plasmid pSRFb Δ Asp resulted in the chloramphenicol-resistant strain AS20 (see Figure 8). In the subsequent congression experiment, *B. subtilis* AS20 was transformed with the integration plasmid phomoAspA-(H322E) along with the helper plasmid pNEXT33A (21), yielding the chloramphenicol-sensitive, neomycin-resistant *B. subtilis* strain KE100 (see Figure 8).

Expression in Escherichia coli and Purification of Recombinant NRPS. Expression of the NRPS gene fragments and purification of recombinant proteins was carried out as described previously (22). *E. coli* strain M15[pREP4], containing the appropriate expression plasmid, was grown at 30 °C until the optical density at 600 nm (*A*₆₀₀) reached 0.7. NRPS gene expression was induced by the addition of 0.1 mM IPTG, and cells were allowed to grow for an additional 2 h. After harvesting and resuspending, cells were broken by two passages through a French Press (Amicon).

His₆-tagged proteins could be purified to apparent homogeneity by single-step Ni²⁺-affinity chromatography. Purity and concentration of the purified proteins were determined by SDS-PAGE and Bradford assay, respectively (23).

ATP-Pyrophosphate-Exchange Assay. The amino acid-dependent ATP-pyrophosphate-exchange reaction was performed as described previously (10) to determine the substrate-selectivity and catalytic activity of recombinant NRPS A domains.

Surfactin Preparation and Detection. Surfactin and derivatives of the lipopeptide were extracted from the cultured broth as described previously (7). Extracts were analyzed by HPLC/MS on a Hewlett-Packard 1100 Series instrument, using a C₈-reverse-phase column (CC250/3 Nucleosil 120-3C₈-column, Macherey & Nagel, Düren, Germany) equilibrated to 70% buffer B (buffer A, 0.05% formic acid in H₂O; buffer B, 0.045% formic acid in methanol). Samples could be separated by applying a linear gradient to 100% buffer B over 30 min (flow rate 0.3 mL min⁻¹), and were monitored at 214 nm, as well as in positive-ion mode over the *m/z* range from 900 to 1200. Fragmentation mass spectrometry was used to corroborate the Asp-to-Asn substitution at residue 5 of the engineered surfactin derivative. The ESI-MS/MS experiment was carried out on a MDS SCIEX Q-STAR Pulsar mass spectrometer, using unit resolution mode for positive product ion extraction. The ion spray voltage was set to 5 kV and collision energy to 80 kV.

RESULTS AND DISCUSSION

In this study, we report on the practical exploitation of the selectivity-conferring code of NRPS A domains in order to rationally alter for the first time the substrate selectivity of whole NRPS modules *in vitro* and *in vivo*.

Generation and Substrate Specificity of a C-A_{Glu}*-PCP Mutant. In the first set of experiments, we attempted to take advantage of the selectivity-conferring code of NRPS A domains in order to alter the substrate selectivity of a whole NRPS module *in vitro*. The starting point was the A domain of the initiation module C-A_{Glu}-PCP, derived from the surfactin biosynthesis complex of *B. subtilis* (24). According to the predictions of the selectivity-conferring code (10), the glutamate-specific binding pocket of SrfA-A1 was expected to differ only in the highly variable position 239 (numbering in accordance with the corresponding position in PheA) from glutamine-specific binding sites, as found, e.g., in the initiation module C-A_{Gln}-PCP of the lichenysin biosynthesis complex (25). As shown in Figure 2, glutamate-activating A domains carry a lysine residue in position 239, whose basic side chain would be perfectly situated to stabilize the substrate's acidic carboxylate moiety by electrostatic interaction. In glutamine-activating A domains, in contrast, a glutamine residue takes up the corresponding position, likely facilitating a polar interaction with the substrate's polar amide moiety. In accordance with the predictions of the selectivity-conferring code, a targeted selectivity-switch for the initiation module C-A_{Glu}-PCP should therefore be feasible, simply by affecting a single Lys239 → Gln mutation in the corresponding A domain.

To prove this assumption, the gene fragment encoding the initiation module of the surfactin biosynthetic complex was amplified from the chromosome of *B. subtilis* and cloned

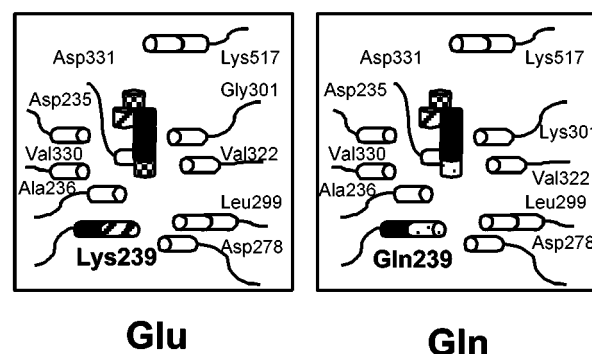


FIGURE 2: Simplified scheme of the proposed binding pockets of L-Glu (SrfA-A1) and L-Gln-activating A domains (LicA1) based on the crystal structure of PheA (10). Both binding pockets differ in only one position (239; numbering in accordance with the corresponding position in PheA). The basic Lys239 in SrfA-A1 is perfectly situated to stabilize the substrate's acidic carboxylate moiety by electrostatic interaction. In LicA1, Gln239 likely facilitates a polar interaction with the substrate's polar amide moiety. Different moieties of the substrate amino acid and certain constituents of the binding pockets are highlighted: aliphatic (gray), polar (dots), acidic (striped), and basic (checked).

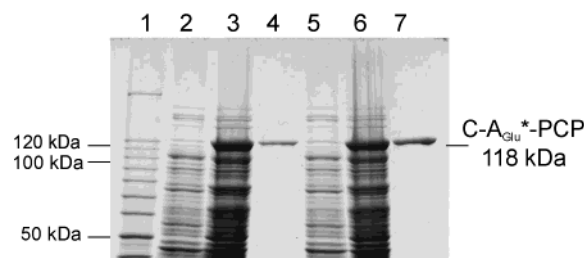


FIGURE 3: Coomassie-blue-stained SDS-polyacrylamide gel showing the production of C-A_{Glu}-PCP wild type (lane 1–3) and C-A_{Glu}-PCP (Lys279 → Gln) mutant (lane 4–6) in *E. coli* M15-[pREP4]. Lanes 1 and 4, total cellular protein before induction; lanes 2 and 5, total cellular protein after 2h induction; and lanes 3 and 6, dialyzed protein pools after purification using Ni²⁺-NTA affinity chromatography.

into the *E. coli* expression vector pQE60. Subsequently, site-directed mutagenesis was used to introduce the corresponding point mutation into the cloned gene fragment. After heterologous expression in *E. coli* M15[pREP4], both wild type and mutant protein were purified using Ni²⁺-NTA affinity chromatography (Figure 3). Substrate selectivity of the proteins was determined by amino acid-dependent ATP-pyrophosphate-exchange reactions (Figure 4). As expected, wild-type C-A_{Glu}-PCP was found to selectively activate L-glutamate, whereas the mutant protein C-A_{Glu}-PCP (Lys239 → Gln) revealed high selectivity for the corresponding amide L-glutamine.

Generation and Substrate Specificity of AspA* Mutants. Most of what has said for the initiation module of the surfactin biosynthetic system also holds true for the elongation module 5, C-A_{Asp}-PCP (Figure 1). According to the predictions of the selectivity-conferring code (10), the aspartate-specific binding pocket of SrfA-B2 differs in three positions from asparagine-specific binding sites as found, e.g., in the bacitracin synthetase C (26) (Figure 5). While the positions Val299 (vs Ile) and Ile330 (vs Val) were presumed to have only modulating impact on the enzyme's selectivity and efficiency, position His322 (vs Glu) was assumed to represent the key residue, facilitating an electrostatic or polar interaction with the carboxylate or amide

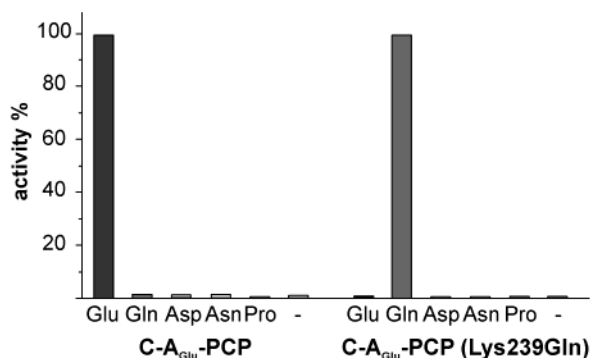


FIGURE 4: Biochemical characterization of purified C-A_{Glu}-PCP wild type and C-A_{Glu}-PCP (Lys279 → Gln) mutant using ATP-pyrophosphate-exchange reactions. The highest exchange activity observed was set to 100%. The wild-type protein C-A_{Glu}-PCP selectively activates L-Glu, whereas no activation of the corresponding amide L-Gln could be detected. In contrast, the mutant protein C-A_{Glu}-PCP (Lys279 → Gln) demonstrates highest selectivity for L-Gln but does not activate the native substrate L-Glu anymore. Neither of the proteins revealed any significant side specificity for the noncognate substrate amino acids L-Asp, L-Asn, and L-Pro.

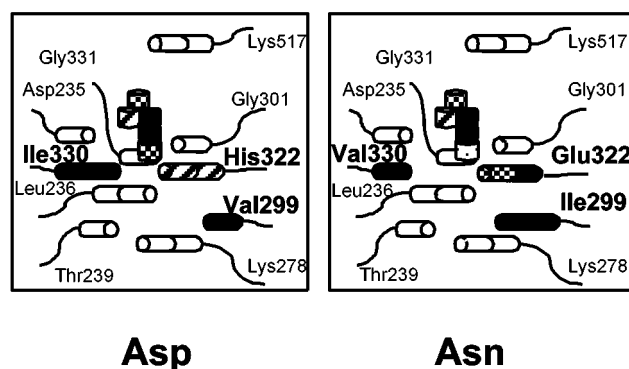


FIGURE 5: Simplified scheme of the proposed binding pockets of L-Asp (SrfA-B2) and L-Asn-activating A domains (BacC5). Both binding pockets differ in only three positions (299, 320, and 322; numbering in accordance with the corresponding position in PheA). While the residues in positions 299 (Val or Ile) and 330 (Ile or Val) are likely to have only modulating influence on the enzyme's activity, position 322 (His or Glu) presumably facilitates an electrostatic or polar interaction with the substrate's carboxylate or amide moiety, respectively. Different moieties of the substrate amino acid and certain constituents of the binding pockets are highlighted: aliphatic (gray), polar (dots), acidic (striped), and basic (checked).

moiety of the substrate's side chain, respectively. In an earlier study, we already showed that selectivity of the aspartate-activating A domain (AspA) could be altered toward asparagine by introducing a single His322 → Glu mutation (10). However, although this mutant demonstrated a high selectivity for the alternate substrate L-Asn, an about 10-fold loss in catalytic efficiency was observed. Although this still represents a rather good outcome for an experiment exclusively based on homology modeling, it was assumed that catalytic efficiency might be fully restored by total adaptation of the substrate-binding pocket to match the shape of the asparagine-recognizing A domain of BacC5 (10).

To prove this assumption, we took wild-type AspA and single-mutant AspA (His322 → Glu) and used site-directed mutagenesis to stepwisely adapt the binding pockets toward the shape of an L-Asn-recognition site. This way, the double mutants AspA (His322 → Glu, Ile330 → Val) and AspA

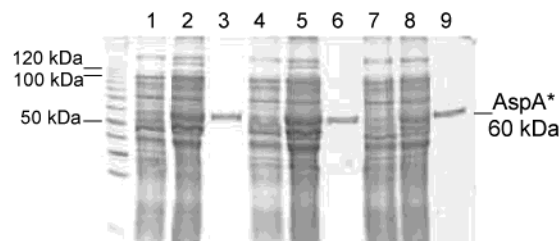


FIGURE 6: Coomassie-blue-stained SDS-polyacrylamide gel showing exemplary the production of AspA wild type (lane 1–3), AspA (His322 → Glu) (lane 4–6), and AspA (His322 → Glu, Ile330 → Val) in *E. coli* M15[pREP4], as well as subsequent purification using Ni²⁺-NTA affinity chromatography. Lanes 1, 4, and 7, total cellular protein before induction; lanes 2, 5, and 8, total cellular protein after 2h induction; and lanes 3, 6, and 9, dialyzed protein pools after purification.

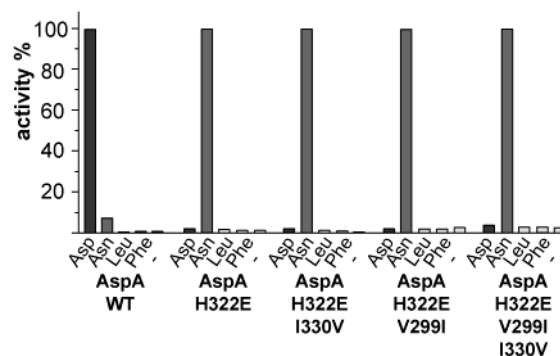


FIGURE 7: Biochemical characterization of the purified AspA wild type protein, as well as single, double, and triple mutants using ATP-pyrophosphate-exchange reactions. The highest exchange activity observed was set to 100%, respectively. The wild type protein AspA selectively activates the cognate amino acid L-Asp, whereas no activation was found for the corresponding amide L-Asn. The mutant proteins AspA(His322 → Glu), AspA(His322 → Glu, Ile330 → Val), AspA(His322 → Glu, Val299 → Ile), and AspA(His322 → Glu, Val299 → Ile, Ile330 → Val), in contrast, specifically activate L-Asn but did not recognize the original substrate L-Asp. Neither of the proteins revealed any significant side selectivity for the noncognate amino acids L-Leu and L-Phe.

(His322 → Glu, Val299 → Ile), as well as the triple mutant AspA (His322 → Glu, Ile330 → Val, Val299 → Ile), were constructed. After heterologous expression in *E. coli* M15[pREP4], all proteins could be purified to apparent homogeneity by single-step Ni²⁺-NTA chromatography (Figure 6). Subsequently, substrate-selectivity was determined by amino acid-dependent ATP-pyrophosphate-exchange reactions (Figure 7 and Table 1). For the wild-type enzyme AspA, this analysis revealed a high selectivity for the activation of L-Asp. In contrast, no activation of L-Asn could be observed. On the other hand, all mutants (single, double and triple) revealed a strict selectivity for L-Asn and no side-specificity for the original substrate L-Asp.

Catalytic Efficiency of the Constructed Mutants. Biochemical analysis of C-A_{Glu}-PCP (Lys279 → Gln) revealed that selectivity could be switched toward L-Gln. However, catalytic efficiency of the mutated NRPS module remained unaffected (wild type and mutant: $k_{cat}/K_M \approx 10 \text{ mM}^{-1} \text{ min}^{-1}$; Table 1). For the wild-type enzyme AspA, the same analysis revealed a catalytic efficiency of $k_{cat}/K_M = 2 \text{ mM}^{-1} \text{ min}^{-1}$. All mutant proteins AspA* (single, double, and triple), however, suffered the same about 10-fold loss in catalytic efficiency ($k_{cat}/K_M = 0.22 \text{ mM}^{-1} \text{ min}^{-1}$) in comparison with the wild-type enzyme (Table 1). Surprisingly, introduction

Table 1: Catalytic Efficiency of the NRPS A Domains (Wild Type and Mutants)^a

protein	substrate	K_M (mM)	k_{cat} (min ⁻¹)	k_{cat}/K_M (mM ⁻¹ min ⁻¹)
AspA (wt)	L-Asp	2.6	5.2	2
	L-Asn	nd	nd	nd
AspA (H322E)	L-Asp	nd	nd	nd
	L-Asn	15.5	3.4	0.22
AspA (H322E, I330V)	L-Asp	nd	nd	nd
	L-Asn	12.5	2.7	0.22
AspA (H322E, V299I)	L-Asp	nd	nd	nd
	L-Asn	20.6	4.6	0.22
AspA (H322E, V299I, I330V)	L-Asp	nd	nd	nd
	L-Asn	27.5	6.3	0.23
C-A _{Glu} -PCP (wt)	L-Glu	2.9	27.3	9.4
	L-Gln	nd	nd	nd
C-A _{Glu} -PCP (K279Q)	L-Glu	nd	nd	nd
	L-Gln	2.9	28.5	9.8

^a nd = not determined, due to low adenylation activity for the given substrate.

of further mutations, to match the shape of an asparagine-specific binding pocket, did not improve the enzymes' catalytic activity.

Apparently, the residues constituting the substrate-binding pocket (here especially, three to five key positions) determine mainly the enzyme's selectivity, whereas the catalytic efficiency is also affected by the A domain's entire tertiary structure and proteinogenic vicinity. This hypothesis is supported by sequence comparison, which revealed a robust similarity of 60% between the L-Glu and L-Gln-activating A domains of SrfA-A1 and LicA1 (25). In contrast, similarity between AspA (SrfA-B2) and the asparagine-activating A domains of BacC5 (26) and TycC1 (22) amounts to only 40%. This value is rather inferior, since it is comparable to the overall similarity found among A domains with various amino acid selectivities. The structural features that promote catalytic efficiency are yet to be discovered. Nonetheless, our results indicate that protein evolutionary approaches to increase the catalytic efficiency of A domains should not be reduced to the enzyme's active site but should also take into consideration the entire proteinogenic environment of 550 amino acids.

Rational Design of the Novel Lipoheptapeptide Variant [Asn⁵]Surfactin. A second set of experiments focused on exploitation of the selectivity-conferring code of NRPS A domains for the in vivo production of novel peptide antibiotics. The model system of choice was the surfactin biosynthetic gene cluster from *B. subtilis*. Several *Bacillus* strains have been known to produce a vast variety of lipoheptapeptide derivatives, summarized in two families designated surfactins (Glu¹) and lichenysins (Gln¹). Apart from the nature of the starter amino acids, derivatives of both classes show only variations in the hydrophobic L-amino acid residues situated at positions 2, 4, and 7 (8, 27), whereas the D-amino acids located at position 3 and 6, as well as the acidic L-Asp residue (position 5), are strictly conserved among all the members of the families. Apparently, certain positions within the lipoheptapeptide are more promiscuous and susceptible for alteration. Consequently, previous attempts to rationally design novel surfactin derivatives have focused on the variable positions 1, 2, and 7. Most variations have been affected at position 7, where the substitution of

NRPS minimal modules (A-PCP) led to the production of the unprecedented derivatives [Orn⁷], [Cys⁷], and [Phe⁷]-surfactin (6). Analogously, the nonproteinogenic amino acid L-Orn was integrated at position 2 (7). Unfortunately, although all the genetically engineered *Bacillus* strains produced the expected surfactin derivatives, all minimal module swaps were on the expense of product yield. This was mainly accounted to obstructions of the intermodular communication and disturbances in the tertiary structure of the hybrid NRPS. Only a whole module swap in the surfactin biosynthetic operon, exchanging the highly homologous modules SrfA-A1 (C-A_{Glu}-PCP) and LicA1 (C-A_{Gln}-PCP), resulted in a high-level lipopeptide production in *B. subtilis* (8).

An approach to overcome these bottlenecks and to minimize perturbations to the tertiary structure of a NRPS protein template could be the change-of-selectivity mutagenesis guided by the selectivity-conferring code of NRPS A domains. To challenge the potentialities of this technique, we attempted the first alteration of the strictly conserved amino acid residue L-Asp at position 5 within the lipoheptapeptide surfactin. In a previous study, it has been found that exchange of the minimal module at position 5 of the surfactin biosynthesis complex did not lead to the production of the expected lipoheptapeptide (Schneider, Ph.D. thesis). On the basis of the results of the preceding experiments, introduction of the single His322 → Glu point mutation into the A domain of SrfA-B2 should lead to an alteration of the substrate selectivity from L-Asp to L-Asn. According to the presented model for the template-driven, nonribosomal assembly of NRPs, this alternate amino acid residue should be subsequently incorporated into the lipoheptapeptide antibiotic, leading to the formation of the up to this point unknown derivative [Asn⁵]surfactin.

Reprogramming of the surfactin biosynthetic operon was accomplished by a two-step method using consecutive gene disruption and replacement monitored by a selectable marker (outlined in Figure 8) (28). For the first step, the disruption plasmid pSRFbΔAsp was constructed, carrying the resistance marker *cat* in place of the gene fragment that encodes the aspartate-specific A domain of module 5 (SrfA-B2). Transformation of the surfactin producer strain *B. subtilis* ATCC 21332 (20) with the linearized disruption plasmid resulted in several chloramphenicol resistant transformants. Correct disruption of the *srfA-B2* locus was verified by Southern hybridization (data not shown). The resulting strain, harboring the *cat* gene within the *srfA-B2* site, was designated AS20. Next, reconstitution of *srfA-B2* and introduction of the mutated gene fragment, encoding the asparagine-specific A domain AspA (His322 → Glu), was achieved by a conjugation experiment. *B. subtilis* AS20 was transformed in parallel with the integration plasmid phomoAsp(H322E) and the helper plasmid pNEXT33A (Figure 8) (21). The latter carries the selectable resistance marker *neo* and integrates into the *metD* locus within the chromosome of *B. subtilis*. Neomycin-resistant transformants were selected and screened for chloramphenicol sensitivity. Southern hybridization and DNA sequencing were used to confirm the correct integration of the mutated gene fragment (data not shown). The resulting *Bacillus* strain harboring the single mutation His322 → Glu within the *srfA-B2* site was designated KE100. SDS-PAGE analysis revealed production of all three surfactin synthetases

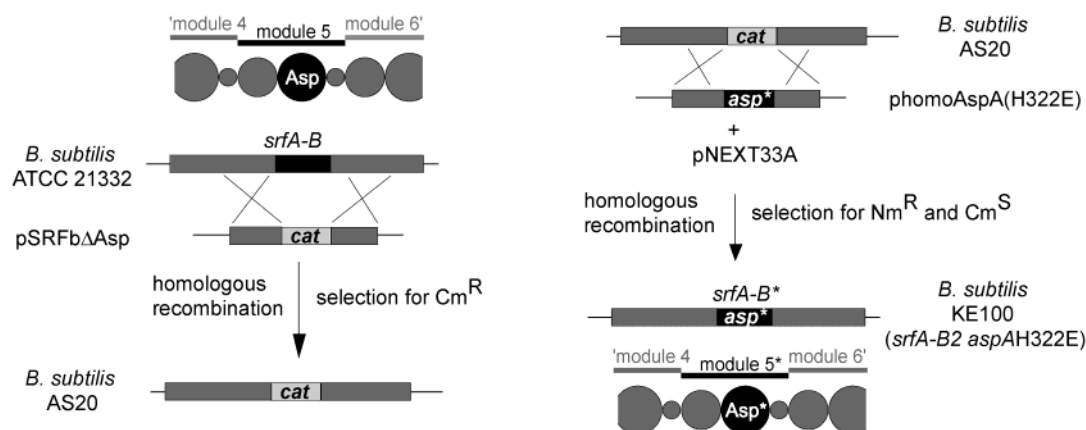


FIGURE 8: Simplified scheme of the marker exchange replacement experiment in *B. subtilis* ATCC 21332. Construction of the disruption strain AS20 (left) was accomplished by transformation, followed by homologous recombination with the linearized deletion plasmid pSRFbΔAsp. Subsequently, reconstitution of the *srfA-B2* locus (right) was achieved by congression using the mutated integration plasmid phomoAspA(H322E) in combination with the helper plasmid pNEXT33A. Integration of the mutated region *aspA*(H322E) occurred by marker exchange homologous recombination and yielded *B. subtilis* strain KE100, carrying in single His322 → Glu point mutation in the A domain of module 5 of the surfactin biosynthetic complex.

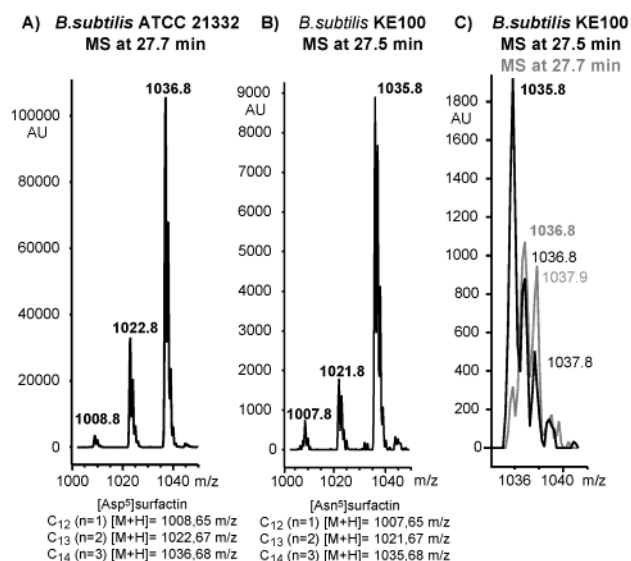


FIGURE 9: LC/MS analysis of extracts derived from the wild-type strain *B. subtilis* ATCC 21332 (A) and the mutant strain *B. subtilis* KE100 (B). The spectra clearly show production of the expected surfactin derivatives [Asp⁵]surfactin and [Asn⁵]surfactin for the wild type and *srfA-B2* mutant strain, respectively. Ion extraction of the mutant spectra (C) also revealed formation of small amounts of wild-type surfactin (gray), in addition to the novel derivative [Asn⁵]surfactin (black) in *B. subtilis* KE100.

SrfA-AB**C* in significant amounts at the transition from logarithmic to stationary phase of growth, and thereby reconstitution of the surfactin biosynthetic system (data not shown).

To examine the influence of a single-point mutation within the approximately 10.8 kb *srfA-B* gene on lipopeptide biosynthesis, we extracted the surfactin derivatives produced by the wild type and the mutant strain KE100 and analyzed them by HPLC/MS (Figure 9). An extract of the *srfA-B2* disruption strain AS20 served as negative control, on account of the incapacity to produce surfactin (data not shown). The lipopeptide obtained from wild type yielded peaks at 27.7 min with *m/z* 1008.8, 1022.8, and 1036.8, as expected from the known structure with slight variations in the length of its β -hydroxy fatty acid moiety (Figure 1; *n* = 1–3). Comparison with the data obtained for the mutant strain

KE100 indicates the replacement of the original Asp⁵ residue and formation of the novel [Asn⁵]surfactin derivative (at 27.5 min, (C₁₂) [M + H] = 1007.8 *m/z*, (C₁₃) [M + H] = 1021.8 *m/z*, (C₁₄) [M + H] = 1035.8 *m/z*). The fatty acid patterns between wild type and mutant revealed no significant differences (Figure 9), and in both cases the C₁₄ derivative (*n* = 3; Figure 1) was the most abundant product.

Since both lipopeptide antibiotics differ only by one mass unit, a MS/MS experiment was carried out to additionally prove the nature of the amino acid residue at position 5 of the corresponding surfactin derivative. The y-series obtained for both wild type and mutant peptide were as follows: Product ion selected for the wild type: (C₁₄) [M + K] = 1074.63 *m/z*. Minus Leu⁷: calcd, 961.55 *m/z*; measured, 961.52 *m/z*. Minus Leu⁶: 848.47 and 848.43 *m/z*. Minus Asp⁵: 733.44 and 733.41 *m/z*. Minus Val⁴: 634.37 and 634.35 *m/z*. Product ion selected for the mutant: (C₁₄) [M + K] = 1073.60 *m/z*. Minus Leu⁷: calcd, 960.52 *m/z*; measured, 960.54 *m/z*. Minus Leu⁶: 847.43 and 847.48 *m/z*. Minus Asp⁵: 733.39 and 733.42 *m/z*. Minus Val⁴: 634.32 and 634.34 *m/z*. Comparison of the y-series, as well as the y-H₂O and y-NH₃ series (data not shown), clearly corroborated the presence of an Asn⁵ residue in various fragments of the novel surfactin derivative. Thus, as expected, alteration of the substrate selectivity of module 5 using site-directed mutagenesis led to the incorporation of an asparagine residue at position 5 of the lipopeptide.

Surprisingly, however, a significant amount of wild type [Asp⁵]surfactin (Figure 9c: at 27.7 min, (C₁₄) [M + H] = 1036.8 *m/z*) could be detected in the extracts of *B. subtilis* KE100, accompanying the signals of the novel [Asn⁵]surfactin derivative (Figure 9c: at 27.5 min, (C₁₄) [M + H] = 1035.8 *m/z*). A possible explanation for this observation could be the desamidation of L-Asn residues, which has been described for various proteins and peptides (29). In our case, desamidation would lead to the conversion of the expected [Asn⁵]derivative into wild type [Asp⁵]surfactin. Another reason could be the editing barrier, provided by the C domain of module 5. Although A domains represent the main selectivity-conferring “gate-keeper” of NRPSs, acceptor-site selectivity has been demonstrated for the accompanied C

domains (30). Taking this into account, the C domain of SrfA-B2 may still prefer the original substrate L-Asp and that way promote the formation of wild type [Asp⁵]surfactin. To investigate whether appearance of [Asp⁵]surfactin was due to an instability of the [Asn⁵]-derivative or caused by a possible Asp-selectivity of the corresponding C domain of SrfA-B2, we examined the time-dependent emergence of wild-type surfactin. To this end, mutant *B. subtilis* KE100 was allowed to grow in modified SpII medium, and at certain time points (12, 24, 48, and 72 h) samples were taken, extracted, and analyzed by HPLC/MS. The experiment revealed a significant shift in time of the [Asn⁵]-to-[Asp⁵]surfactin ratio in favor of wild type surfactin (data not shown). This result clearly points to instability and slow desamidation of the [Asn⁵]-derivative, rather than system inherent selectivity.

CONCLUSIONS

Up to now, prospects for the genetic manipulation of NRPS biosynthetic templates and the rational design of novel peptide antibiotics were restricted to the substitution of individual modules and catalytic domains. Obviously, these extensive manipulations frequently led to significant interferences in the tertiary structure—and thus productivity—of the corresponding hybrid NRPSs. A novel strategy to overcome these bottlenecks has become available with the change-of-selectivity mutagenesis, minimizing the perturbations to the overall structure of the biosynthetic template to few amino acid substitutions. In this study, we present the first example for the practical exploitation of the selectivity-conferring code for the rational design of novel peptide antibiotics. Despite of the successful outcome, an important limitation of this novel approach has to emphasize. All examples, which had been described so far for the successful alteration of A domain selectivity, were limited to maximal three point mutations within the enzyme's substrate binding pocket and restricted to relatively slight variations in the substrate's shape and/or polarity. Thus, applicability of this approach may be limited to the generation of similar derivatives or structurally focused combinatorial libraries of a given peptide antibiotic. On the other hand, NRPSs are relatively promiscuous enzymes that usually yield the formation of a heterogeneous mixture of derivatives of a certain peptide antibiotic. This circumstance frequently gets in the way with an efficient fermentative production of a pharmacologically important peptide antibiotic, since normally only one of these compounds possesses the desired bioactivity. In this connection, the change-of-selectivity mutagenesis could be advantageous and useful to remove side-specificities of NRPS A domains, thereby suppressing undesired byproducts and facilitating the fermentative high-yield production of a certain NRP.

No negative influence on the growing behavior of the [Asn⁵]surfactin producer *B. subtilis* KE100 was observed compared to the wild-type strain ATCC 21332. However, general concerns during the in vivo production of novel hybrid antibiotics are the export of the generated products, as well as the auto-immunity of the producing organism. Recently, the self-resistance-conferring mechanism against surfactin could be enlightened in *B. subtilis* (31). The gene-product of the identified gene, *yerP*, represents the first member of the resistance, nodulation, and cell division

(RND) family of proton motive force-dependent efflux pumps in Gram-positive bacteria. By deletion of *yerP*, surfactin susceptible mutants were obtained. Although the transport of surfactin has not been demonstrated, the efflux of the inherent metabolite catalyzed by YerP was postulated. Even if the specificity of YerP has not been determined, the protein's ability to possibly export newly designed lipopeptides has great implications for the genetic manipulation of the surfactin biosynthetic pathway in *B. subtilis*.

ACKNOWLEDGMENT

We are indebted to Jürgen J. May and Axel Schneider for the construction of pCA_{Glu}PCP and *Bacillus subtilis* strain AS20, respectively. We thank Veit Bergendahl for his help during HPLC/MS analysis. We also acknowledge the excellent practical work of Inge Schüller, Stephan Grünwald, and Michael Maue.

REFERENCES

- Walsh, C. (2000) *Nature* 406, 775–781.
- Schwarzer, D., and Marahiel, M. A. (2001) *Naturwissenschaften* 88, 93–101.
- Doekel, S., and Marahiel, M. A. (2001) *Metab. Eng.* 6, 64–77.
- Doekel, S., and Marahiel, M. A. (2000) *Chem. Biol.* 7, 373–384.
- Mootz, H. D., Schwarzer, D., and Marahiel, M. A. (2000) *Proc. Natl. Acad. Sci. U.S.A.* 97, 5848–5853.
- Stachelhaus, T., Schneider, A., and Marahiel, M. A. (1995) *Science* 269, 69–72.
- Schneider, A., Stachelhaus, T., and Marahiel, M. A. (1998) *Mol. Gen. Genet.* 257, 308–318.
- Yakimov, M. M., Giuliano, L., Timmis, K. N., and Golyshin, P. N. (2000) *J. Mol. Microbiol. Biotechnol.* 2, 217–224.
- Conti, E., Stachelhaus, T., Marahiel, M. A., and Brick, P. (1997) *EMBO J.* 16, 4174–4183.
- Stachelhaus, T., Mootz, H. D., and Marahiel, M. A. (1999) *Chem. Biol.* 6, 493–505.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Nakano, M. M., Marahiel, M. A., and Zuber, P. (1988) *J. Bacteriol.* 170, 5662–5668.
- Cutting, S. M., and Vander Horn, P. B. (1990) in *Molecular Biological Methods for Bacillus* (Harwood, C. R., and Cutting, S. M., Eds.) p 34, John Wiley and Sons, Chichester, U.K.
- Bullock, W. O., Fernandez, J. M., and Short, J. M. (1987) *Biotechniques* 5, 376–379.
- Schauder, B., Blöcker, H., Frank, R., and McCarthy, J. E. G. (1987) *Gene* 52, 279–283.
- Herbst, B., Kneip, S., and Brehmer, E. (1994) *Gene* 151, 137–142.
- Antoniewski, C., Savelli, B., and Stragier, P. (1990) *J. Bacteriol.* 172, 86–93.
- Eppelmann, K., Doekel, S., and Marahiel, M. A. (2001) *J. Biol. Chem.* 276, 34824–34831.
- Klein, C., Kaletta, C., Schnell, N., and Entian, K. D. (1992) *Appl. Environ. Microbiol.* 58, 132–142.
- Cooper, D. G., Macdonald, C. R., Duff, S. J. B., and Kosaric, N. (1981) *Appl. Environ. Microbiol.* 42, 408–412.
- Itaya, M., and Tanaka, T. (1990) *Mol. Gen. Genet.* 223, 268–272.
- Mootz, H. D., and Marahiel, M. A. (1997) *J. Bacteriol.* 179, 6843–6850.
- Laemmli, U. K. (1970) *Nature* 227, 680–685.
- Cosmina, P., Rodriguez, F., de Ferra, F., Grandi, G., Perego, M., Venema, G., and van Sinderen, D. (1993) *Mol. Microbiol.* 8, 821–831.
- Konz, D., Doekel, S., and Marahiel, M. A. (1999) *J. Bacteriol.* 181, 133–140.
- Konz, D., Klens, A., Schörgendorfer, K., and Marahiel, M. A. (1997) *Chem. Biol.* 4, 927–937.

27. Peypoux, F., Bonmatin, J. M., and Wallach, J. (1999) *Appl. Microbiol. Biotechnol.* 51, 553–563.
28. Stachelhaus, T., Schneider, A., and Marahiel, M. A. (1996) *Biochem. Pharmacol.* 52, 177–186.
29. Bischoff, R., and Kolbe, H. V. (1994) *J. Chromatogr., B* 662, 261–278.
30. Belshaw, P. J., Walsh, C. T., and Stachelhaus, T. (1999) *Science* 284, 486–489.
31. Tsuge, K., Ohata, Y., and Shoda, M. (2001) *Antimicrob. Agents Chemother.* 45, 3566–3573.

BI0259406