The bacitracin biosynthesis operon of *Bacillus licheniformis* ATCC 10716: molecular characterization of three multi-modular peptide synthetases

Dirk Konz[1], Andrea Klens[1], Kurt Schörgendorfer[2] and Mohamed A Marahiel[1]

Background: The branched cyclic dodecylpeptide antibiotic bacitracin, produced by special strains of *Bacillus*, is synthesized nonribosomally by a large multienzyme complex composed of the three bacitracin synthetases BA1, BA2 and BA3. These enzymes activate and incorporate the constituent amino acids of bacitracin by a thiotemplate mechanism in a pathway driven by a protein template. The biochemical features of these enzymes have been studied intensively but little is known about the molecular organization of their genes.

Results: The entire bacitracin synthetase operon containing the genes *bacA–bacC* was cloned and sequenced, identifying a modular structure typical of peptide synthetases. The *bacA* gene product (BA1, 598 kDa) contains five modules, with an internal epimerization domain attached to the fourth; *bacB* encodes BA2 (297 kDa), and has two modules and a carboxy-terminal epimerization domain; *bacC* encodes BA3, five modules (723 kDa) with additional internal epimerization domains attached to the second and fourth. A carboxy-terminal putative thioesterase domain was also detected in BA3. A putative cyclization domain was found in BA1 that may be involved in thiazoline ring formation. The adenylation/thioester-binding domains of the first two BA1 modules were overproduced and the detected amino-acid specificity coincides with the first two amino acids in bacitracin. Disruption of chromosomal *bacB* resulted in a bacitracin-deficient mutant.

Conclusions: The genes encoding the bacitracin synthetases BA1, BA2 and BA3 are organized in an operon, the structure of which reflects the modular architecture expected of peptide synthetases. In addition, a putative thiazoline ring formation domain was identified in the BA1 gene.

Addresses: [1]Philipps-Universität Marburg, Fachbereich Chemie/Biochemie, Hans-Meerwein-Straße, 35032 Marburg, Germany. [2]Biochemie GmbH, 6330 Kufstein/Schaftenau, Austria.

Correspondence: Mohamed A Marahiel
E-mail: Marahiel@ps1515.chemie.uni-marburg.de

Key words: *Bacillus licheniformus*, bacitracin, DNA sequence, peptide antibiotics, peptide synthetases, thiazoline ring

Received: **30 September 1997**
Accepted: **28 October 1997**

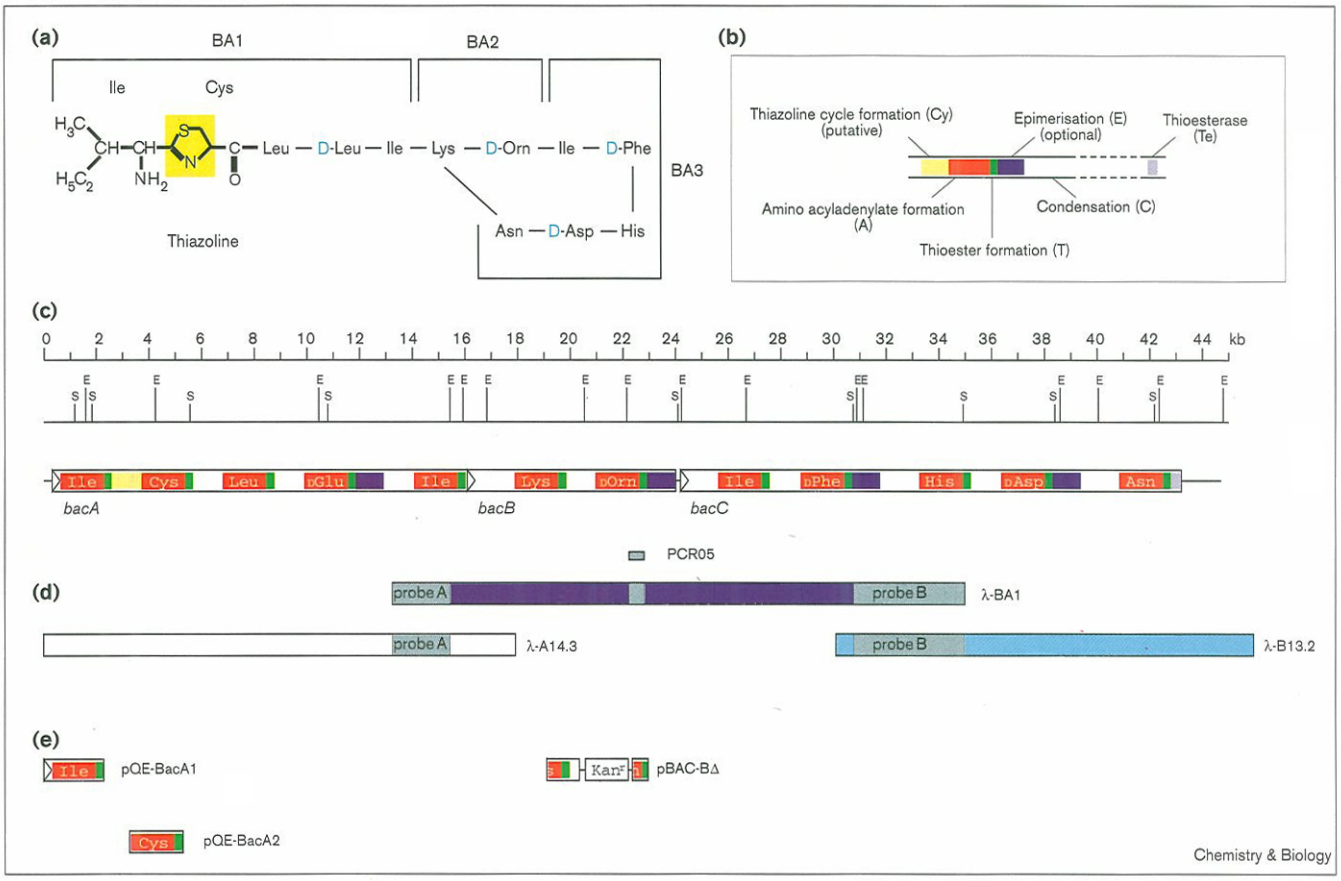
Chemistry & Biology December 1997, **4**:927–937
http://biomednet.com/elecref/1074552100400927

Introduction

Bacitracins are antibiotic polypeptides produced by certain strains of *Bacillus licheniformis* [1,2] and *Bacillus subtilis* [3]. Crudely purified bacitracin contains at least 10 distinct dodecapeptides that differ by one or two amino acids [4]. The most abundant and best characterized of these peptides is bacitracin A (Figure 1a), a branched cyclic dodecapeptide. It contains an amino-terminal linear pentapeptide moiety with an isoleucine–cysteine thiazoline condensation product and a carboxy-terminal heptapeptide ring, in which the free α-carboxy group of the carboxy-terminal Asn is bound to the ε-amino group of lysine. As well as proteinogenic amino acids, four amino acids in the D-configuration (Glu4, Orn7, Phe9 and Asp11) and the nonproteinogenic residue ornithine (Orn) are incorporated into bacitracin A. Bacitracin A is most active against Gram-positive bacteria and acts by inhibiting bacterial cell-wall biosynthesis: bacitracin A forms a tight ternary complex with C_{55}-isoprenyl pyrophosphate (IPP) and a divalent metal cation [5,6]. IPP serves as a membrane-associated carrier during the synthesis of the repeat subunits of peptidoglycan. Recycling of IPP involves its dephosphorylation to C_{55}-isoprenyl phosphate (IP) by a phosphatase at the end of each cycle; this step is blocked by the bacitracin–M^{2+}–IPP-complex (where M^{2+} is a divalent metal ion), resulting in an accumulation of uridine diphosphate-acetylmuramyl pentapeptide. This model for bacitracin action is supported by recently described mechanisms for bacitracin resistance involving the production of increased IPP [7,8]. Besides this primary mode of action, bacitracin also seems to affect membrane functions [9], the action of certain hydrolytic enzymes [10], and the biosynthesis of ubiquinone precursors [11].

Like many other peptide antibiotics from *Bacillus* [12], bacitracin is produced nonribosomally by a large multienzyme complex [13–15] utilizing a thiotemplate mechanism [16–20]. In this mechanism, large multifunctional enzymes, designated peptide synthetases, activate their substrate amino acids as aminoacyl adenylates using ATP hydrolysis. These unstable intermediates are subsequently

Figure 1



(a) Primary structure of bacitracin A. Brackets embrace the residues incorporated by the three bacitracin synthetases BA1, BA2 and BA3. The thiazoline ring, a product of isoleucine and cysteine cyclization, is highlighted in yellow. **(b)** The color code of the patterns used for different domains in the remainder of the figure. **(c)** Physical organization of the *bac* operon. Three genes, *bacA*, *bacB* and *bacC* coding for the bacitracin synthetases BA1–BA3 were identified. The domain organization of the peptide synthetases is illustrated within the genes. **(d)** PCR fragment 05 used for the screening of the λ-EMBL3 genomic library and λ clones isolated in this work. Probes A and B used for chromosomal walking are marked in light blue. **(e)** Gene fragments, corresponding to AT domains, were overexpressed and biochemically analyzed. Plasmid pBAC–BΔ was used for the construction of a *bacB* gene disruption mutant. Abreviations: Orn, ornithine; E, *Eco*RI; S, *Sal*I.

transferred to a covalently enzyme-bound 4′-phosphopantetheinyl (4′-PP)-cofactor as thioesters. The peptide product is generated by the stepwise incorporation of the thioesterified amino acids in a series of amino- to carboxy-terminal-directed transpeptidations. Thorough sequence analysis of several genes encoding peptide synthetases has unveiled a modular organization principle for this class of enzymes [18,21–23]. The detected modules (around 1000 amino acids each) are responsible for the activation and incorporation of a single amino acid and thus are arranged in a colinear fashion with the product peptide [24–26]. If a biosynthetic complex consists of several peptide synthetases, the colinearity of the modules is conserved within each single enzyme [17,18]. Each module can be subdivided into several specific domains [20,22,27]. The adenylation (A) domain (550 amino acids) is the elementary domain of each module and catalyzes the specific activation of a given substrate amino acid. This module belongs to the large family of adenylate-forming enzymes that includes firefly luciferase [28] and the acetyl-CoA synthetases [22]. The adjacent thioester-binding (T) domain (80 amino acids) shares high homology with the acyl carrier protein (ACP) that is involved in fatty acid and polyketide synthesis. It contains the highly conserved sequence motif LGGx*S*I (using single-letter amino-acid code, with x as any amino acid) with an invariant serine residue (italicized), which is the 4′-PP-attachment site [29–31]. In contrast to the original model of the thiotemplate mechanism, which suggested that a single 4′-PP-cofactor would act as a 'swinging arm' to transfer the growing peptide chain from one module to the next, the presence of such a cofactor at each module of a multimodular PPS has recently been shown [32]; this finding established the 'multiple carrier model' for nonribosomal peptide synthesis. The T domain is followed by a recently reported condensation (C) domain (350 amino acids), which is thought to catalyze

peptide bond formation [33]. Within peptide synthetases with a single polypeptide chain, the C domain is always found between two A domains. If the peptide biosynthesis complex is composed of several peptide synthetases, the subunit that activates the first amino acid of the peptide lacks a C domain. The C domain is found instead at the amino-terminal end of subunits that catalyze activation and transfer of the distal amino acids of the peptide. In modules that catalyze the incorporation of *N*-methylated amino acids, an additional *N*-methylase domain (350 amino acids) is inserted between the A and T domains [24,25]. If the module introduces D-amino acids, the T domain is followed by an epimerization (E) domain (400 amino acids) [27,34]. In many bacterial systems the module that catalyzes the incorporation of the carboxy-terminal amino acid of the product peptide contains an additional putative thioesterase (Te) domain; Te domains are thought to play a role in peptide chain release or cyclization [20].

The biochemical pathway of bacitracin production has been studied intensively and the enzymes involved in bacitracin's biosynthesis were identified and characterized 20 years ago [13–15]. Bacitracin is synthesized by a large multienzyme complex, composed of three peptide synthetases BA1, BA2 and BA3. BA1 activates the five amino acids contained in the linear portion of the peptide; BA2 and BA3 activate the remaining amino acids of the heptapeptide ring (Figure 1a). In contrast to the mass of biochemical data, little is known about the bacitracin synthetase genes. Recently, the cloning and expression of the *bacB* and (partial) *bacC* genes (both are located on a 32 kb fragment of *B. licheniformis*) and expression of these genes in *Escherichia coli* has been reported [35], indicating that the genes for bacitracin synthesis are clustered in the genome of *B. licheniformis*. This proposition was supported by the isolation of a Tn917-generated bacitracin-deficient mutant with an insertion in the vicinity of the *bacB* gene [36], but to date the sequences of the genes have not been available. Given that the bacitracin-biosynthesis system is one of the best biochemically characterized systems for nonribosomal peptide synthesis, analysis of bacitracin synthetase gene sequences is of great importance. Of particular note is the fact that BA1 and BA3, on the basis of biochemical data, incorporate three D-amino acids (D-Glu4, D-Phe9 and D-Asp11) from putative internal modules, which are expected to lead to the generation of of internal E domains. In all reported prokaryotic genes of peptide synthetases to date, E domains are located exclusively at the carboxy-terminal end of the enzyme [27]. Thus, detailed sequence information could lead to new insights into the general architecture of peptide synthetases. Furthermore, the analysis of the *bacA* gene is a subject of outstanding interest from another point of view — chemical investigations of the isoleucylcysteine intermediate isolated from BA1 preparations have indicated that the thiazoline ring is formed at the dipeptide stage [37], so an unusual organization of the first two BA1 modules is to be expected.

In this study we report the cloning and sequencing of the entire *bac* operon from *B. licheniformis*. Three genes, *bacA*, *bacB* and *bacC*, coding for the three biochemically identified bacitracin synthetases were sequenced. The genes reflect the typical structure of a peptide synthetase, with five modules in *bacA*, two in *bacB* and five in *bacC*. The chromosomal arrangement of the genes is in the same order as their biochemical action. Complete sequence analysis of *bacA* and *bacC* revealed that they contain modules with internal E domains, the first example of these domains in prokaryotic peptide synthetases. In addition, a strongly modified C domain was identified between the first two A domains of *bacA;* this region shares highest homologies with proteins involved in the nonribosomal synthesis of thiazoline-containing siderophores (iron-chelating compounds) [38-41]. We conclude that this region represents a new kind of domain in peptide synthetases, and that it catalyzes the formation of a thiazoline ring. The deduced amino acid specificity of the first two BA1 modules was confirmed by cloning, heterologous expression in *E. coli* and ATP–PP_i exchange assays with purified domain proteins. Furthermore, the relevance of the cloned genes in bacitracin biosynthesis was demonstrated by constructing a bacitracin-deficient *bacB* gene-disruption mutant.

Results and discussion

Cloning and sequencing of the *bac* operon

We used a recently reported method based on the polymerase chain reaction (PCR) for the identification of peptide synthetase genes [42]. Using degenerate oligonucleotide primers derived from the core sequences 4 (Thr–Glu–Asp) and 6 (Leu–Glu–Glu–x–Ser–Ile) of peptide synthetases [27] a DNA fragment, designated PCR05, was PCR-amplified from the chromosome of *B. licheniformis* ATCC 10716. Sequencing this fragment identified the internal core sequence 5 (KIRGxRIELGEIE; using the single-letter amino-acid code, with x as any amino acid) and another highly conserved sequence motif (NGK), both of which are characteristic of peptide synthetases, thus indicating that PCR05 represents a part of a peptide synthetase gene. In order to clone the complete peptide synthetase gene, we constructed a λ-EMBL3 genomic library of *B. licheniformis* ATCC 10716 and screened this library using the PCR05 fragment as probe. Seven λ clones that hybridized with the probe were isolated. One of them, designated λ-BA1 (Figure 1d) was further investigated. The 19.5 kilobase (kb) insert of λ-BA1 was mapped and subcloned into pBluescript SK(–) using the restriction enzymes *Eco*RI and *Sal*I. By terminal sequencing of the resulting subclones, it was shown that the λ-BA1 DNA insert shares homology with peptide synthetase genes over the entire stretch of 19.5 kb. To identify the remaining part of the peptide synthetase gene, an

additional step of chromosomal walking was performed; 5′-(probe A) and 3′-(probe B) terminal restriction fragments of the λ-BA1 insert were used to screen the λ-EMBL3 genomic library (Figure 1d). Fourteen phages that reacted with probe A and 17 that hybridized with probe B were isolated. The inserts of the isolated phages were mapped and analyzed by Southern blot experiments. One clone of each set, which showed only short overlaps with λ-BA1, named λ-A14.3 (18 kb insert) and λ-B13.2 (16 kb insert; see Figure 1d), were subcloned into pBluescript SK(–) using *Eco*RI and *Sal*I. Terminal sequencing of these subclones showed that each phage contained one terminal region lacking homology to peptide synthetase genes, indicating that the complete peptide synthetase gene had been cloned. Altogether, the three λ-phages BA1, A14.3 and B13.2 span a continuous region of 46 kb. The entire nucleotide sequences of the subclones were determined, and remaining gaps between the subclones were filled by PCR amplification from chromosomal DNA. In all, 43,297 base pairs (bp) of *B. licheniformis* were sequenced, and the sequences had a typical GC-content of about 46%.

The *bac* operon consists of the three genes, *bacA*, *bacB* and *bacC*, coding for the bacitracin synthetases BA1, BA2 and BA3

Analysis of the subcloned nucleotide sequence revealed three huge open reading frames (ORFs) with the same direction of transcription. The first ORF, designated *bacA*, is 15,768 bp in length and codes for a protein of 5,255 amino acids and a predicted mass of 598,379 Da. By detailed alignments with the deduced amino-acid sequence, five modules typical of peptide synthetases were identified. Each of these modules consists of an A domain and a T domain. The first module lacks an amino-terminal C domain, suggesting that it might be the initiation module [33]. A strongly modified C domain, designated as a putative cyclization (Cy) domain, was identified between the first two A domains (described in detail below). The remaining A domains are connected by homologous and classical C domains. Module 4 carries an additional E domain inserted between its T and the following C domain. This organization resembles the arrangement of domains found in the first module of HC-toxin synthetase Hts1 [26], which is responsible for the incorporation of D-proline in the modified tetrapeptide HC-toxin. Taken together, these findings are in perfect agreement with the requirements for BA1 activity: induction of peptide synthesis; activation and incorporation of five amino acids; thiazoline ring formation between the first two amino acids and incorporation of a D-amino acid (D-glutamine) in the fourth position.

The second identified ORF, of 7,824 bp, is named *bacB*, and codes for a protein of 2,607 amino acids with a calculated mass of 297,428 Da. The derived amino acid sequence is composed of two AT domains, each coupled with an amino-terminal C domain. Furthermore a carboxy-terminal E domain was found, indicating that this protein might be BA2, which activates two amino acids and incorporates a D-amino acid (ornithine) from its second active site into bacitracin. ORF 3, designated *bacC*, spans a stretch of 19,080 bp, coding for a giant protein of 6,359 amino acids with a predicted mass of 722,943 Da. This protein consists of five modules each made up of C–A–T domains (see Figure 1). The second and fourth modules additionally bear an internal E domain inserted in the same fashion as described for module 4 of *bacA*. A putative thioesterase domain was found attached to the carboxy-terminal module. This organization of domains corresponds fully with the biochemical properties of BA3: activation of five amino acids; incorporation of two D-amino acids catalyzed by its second and forth active sites and release and cyclization of mature bacitracin. Taken together, the three genes *bacA*, *bacB* and *bacC* span a stretch of 42,886 bp. There are several putative start codons for *bacB*; but the ATG at bp 16,191, which is located 112 bp downstream of the *bacA* stop codon, seems to be the most probable. The *bacB* and *bacC* genes are separated by 105 bp. In other bacterial systems such as gramicidin S [43], surfactin [21] and tyrocidine [23,44] the genes encoding for the peptide synthetases were shown to be organized in operons. As judged by the short distance between the identified genes and the absence of any detectable termination signals we propose that *bacA*, *bacB* and *bacC* comprise an operon.

Analysis of the primary structure of the bacitracin peptide synthetase reveals a highly ordered domain structure that obeys the empirical rules for the molecular architecture of peptide synthetases that have been elaborated in recent years following the sequencing of several such genes [17,20,30]. In particular, the first examples of internal E domains in bacterial peptide synthetase systems, reported here, underline once more the idea that peptide synthetases can be understood as chains of several linear coupled domains that catalyze specific reactions according to their positions within these chains.

Homology of BA1 to proteins involved in the nonribosomal synthesis of siderophores

One of the most striking structural features of the polypeptide antibiotic bacitracin is the thiazoline ring formed between the first two amino-terminal amino acids (isoleucine–cysteine). Biochemical studies of the bacitracin synthetases have revealed that the heterocycle formation is catalyzed by BA1 at the dipeptide stage [37]. We characterized the *bacA* gene encoding BA1 at the molecular level and identified an unusual domain between the first two modules of this enzyme. A search of the GenBank database was performed with the BLASTX program to determine if this domain had significant homology with known protein sequences [45]. The highest percentage of similarity was

Table 1

Homology of putative cyclization domains.

	HMWP2-1	HMWP2-2	MTCY22H8.02	AngR
BA1:1-2	44%	51%	48%	36%
HMWP2-1	-	50%	51%	43%
HMWP2-2		-	48%	37%
MTCY22H8.02			-	35%
AngR				-

Percentage of homology found for the putative cyclization domains of BA1, HMWP2, MTCY22H8.02 and AngR when aligned with the clustalW program.

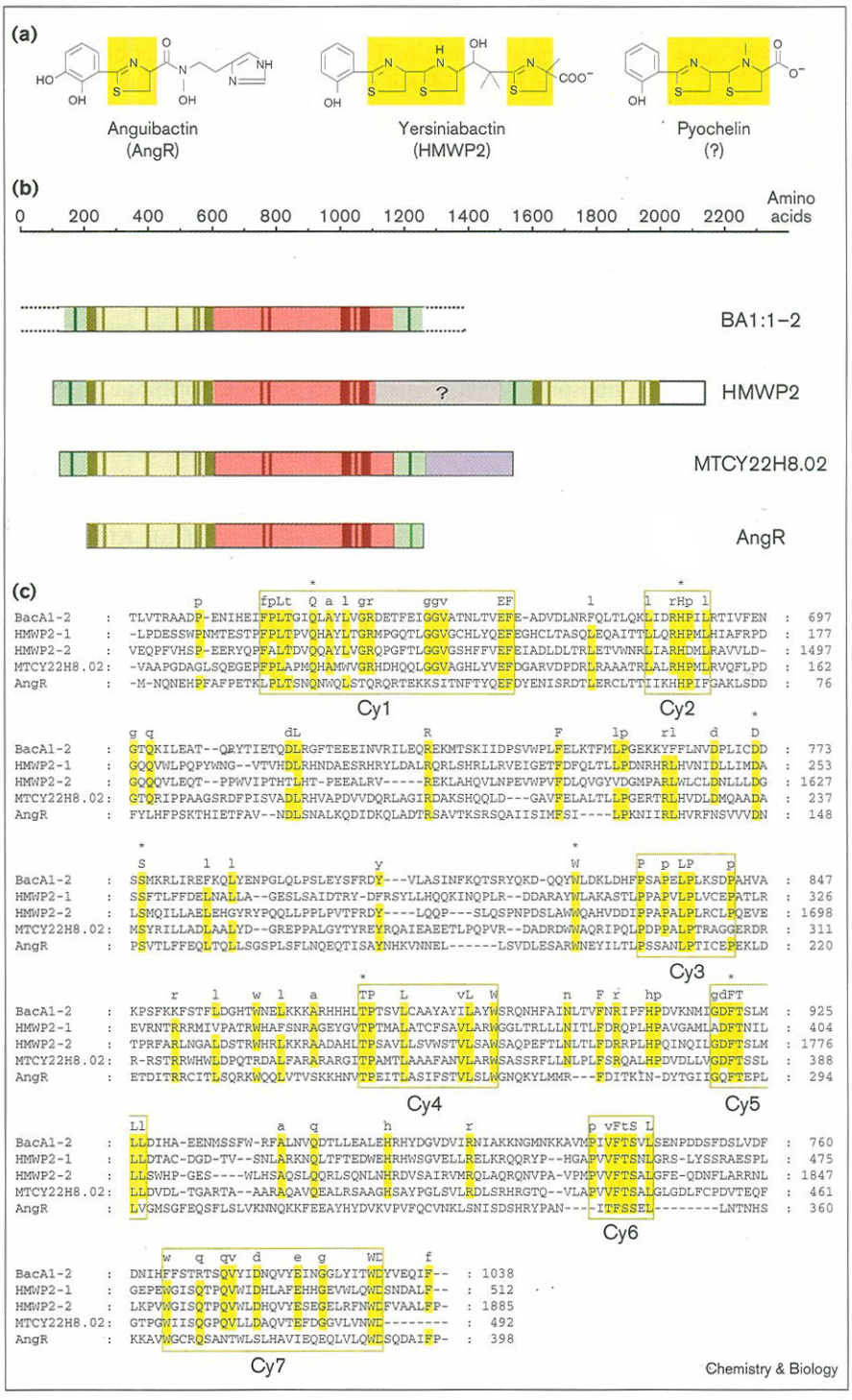
found with a group of three proteins: HMWP2 of *Yersinia enterocolitica*, MTCY22H8.02 of *Mycobacterium tuberculosis* and AngR of *Vibrio anguillarum* (Table 1) [38,39,41]. HMWP2 and AngR are reported to be involved in the nonribosomal synthesis of the thiazoline-containing siderophores yersiniabactin (HMWP2) and anguibactin (AngR), respectively [38,46]. The biological function of MTCY22H8.02 is not known, but the gene encoding it is located in a gene cluster with an aryl–AMP-ligase (MTCY22H8.01), polyketide synthetase analogues (MTCY22H8.03/04) and a peptide synthetase (MTCY22H8.05) [41]. It is possible that MTCY22H8.02 might also be involved in the nonribosomal synthesis of a heterocycle-containing catechol siderophore like yersiniabactin, anguibactin or pyochelin of *Pseudomonas aeruginosa* (Figure 2a) [47–49]. The comparative sequence analysis of BA1, HMWP2, MTCY22H8.02 and AngR revealed that our proteins have a similar domain structure. Each sequence harbors a core module containing an unusual domain comparable to the one identified in BA1 plus an A and T domain. In HMWP2, an additional domain of unknown function is inserted between the A and T domains. At the amino-terminal side of such modules, a T domain is detected in BA1, HMWP2 and MTCY22H8.02. Another Cy domain is found at the carboxyl terminus of HMWP2, and a putative Te domain is found at the carboxyl terminus of MTCY22H8.02 (Figure 2b). Taken together these findings indicate that the four enzymes catalyze similar reactions and that the newly identified type of domain might be involved in the catalysis of thiazoline ring formation, so we have designated this domain as a putative cyclization (Cy) domain.

By aligning the five homologous Cy domains, seven conserved regions Cy1–7 (Figure 2c) were identified. When the Cy domain was compared with C domains of peptide synthetases (25–33% homology), some residues were identified that are conserved in both types of domains (Figure 2b) whereas others are unique to each domain. The most noticeable difference between the two types is the reduced conservation of the HHIxDGS signature-sequence motif in Cy domains. In particular, the histidine residues, which are thought to play a key role in the transpeptidation catalysis [33], are missing from Cy domains, although the aspartic acid and serine residues are strongly conserved in both types of domain. These data, in combination with the fact that no additional sequence other than the identified Cy domain is located between the first two modules of BA1, led us to the suggestion that the identified Cy domain catalyzes both peptide bond formation and thiazoline ring formation. Thus, analysis of the conserved residues in C and Cy domains, by site-directed mutagenesis, could provide a deeper insight into the molecular mechanisms of transpeptidation catalysis by peptide synthetases. An enzyme complex involved in the post-translational formation of thiazole and oxazole rings in the ribosomally synthesized polypeptide antibiotic microcin B17 of *E. coli* has recently been reported [50]. This complex, composed of the three enzymes McbB–McbD, introduces four thiazole and four oxazole rings into premicrocin B17 involving the sidechains of serine and cysteine residues. In a sequence alignment of McbB–McbD with the Cy domains identified in this report, no significant homologies were detected. These data indicate that both systems catalyze similar reactions but utilize alternative mechanisms of heterocycle formation.

Amplification, expression and biochemical characterization of the first two BA1 AT domains

DNA fragments corresponding to the AT domains of the first two BA1-modules were amplified by PCR out of total DNA from *B. licheniformis* ATCC 10716. The first module fragment *bacA*1–AT (1,946 bp) was amplified with the native amino terminus of BA1 and the carboxy-terminal end was chosen 75 residues downstream of the core sequence 6 (MGGxS) [18]. For the second module *bacA*2–AT (1,856 bp) an amino-terminal primer located 98 residues in front of core sequence 1 (LKtGA) and a carboxy-terminal primer 65 residues behind core sequence 6 (IGGDS) were used [18,23]. The DNA fragments obtained were cloned into the pQE expression vector as described in Materials and methods. Expression in *E. coli* yielded His^6-tagged recombinant domain proteins, designated BacA1–AT and BacA2–AT, which were found to be completely soluble (Figure 3). The observed electrophoretic mobility on SDS–PAGE was in good agreement with the calculated mass of 73,164 Da for BacA1–AT and 70,461 Da for BacA2–AT. Purification of these proteins was performed by Ni^{2+}-chelate affinity chromatography. The resulting protein fractions were applied to amino acid-dependent ATP–PP_i exchange assays using all 20 proteinogenic amino acids plus ornithine and one control without amino acids. The results of these experiments are presented in Figure 4, and clearly reveal that BacA1–AT specifically activates isoleucine, whereas BacA2–AT activates cysteine. If the highest activation for each domain protein is defined as 100%, the measured background level of the control reaction without amino acids was found to be

Figure 2

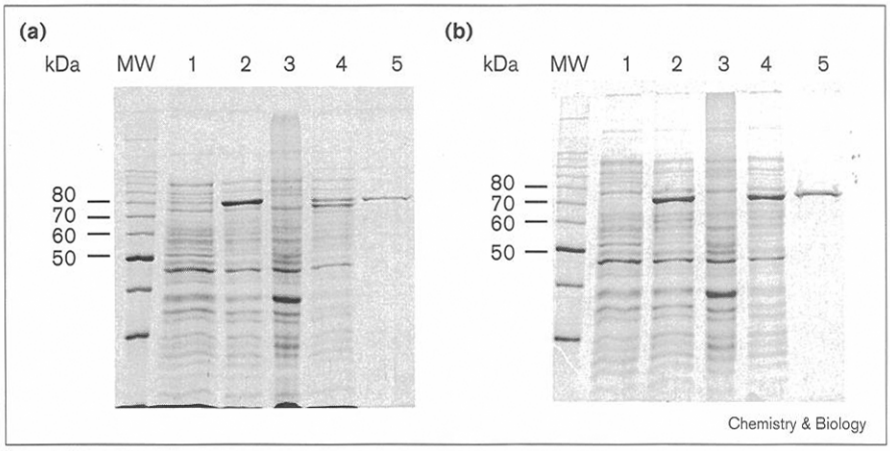


(a) Primary structures of the nonribosomally produced siderophores anguibactin [47], yersiniabactin [48] and pyochelin, with their possible synthetic proteins in parentheses (where known) [49]. **(b)** Comparative alignment of BA1:1–2, HMWP2, MTCY22H8.02 and AngR. The colored regions indicate the identified domains: green, thioester-binding (T) domain; yellow, putative-cyclization (Cy) domain; red, adenylation (A) domain; grey, not identified; light purple, thioesterase (Te) domain. Dark lines in the different domains indicate the position of highly conserved sequence motifs [27]. **(c)** Alignments of amino-acid sequences of the cyclization (Cy) putative domains of BA1:1–2 (residues 621–1038) of *B. licheniformis* with HMWP2-1 (residues 100–512), HMWP2-2 (residues 1481–1885) of *Y. enterocolitica* [38], MTCY22H8.02 (residues 85–492) of *M. tuberculosis* [41] and AngR (residues 1–398) of *V. anguillarum* [39] performed with the clustalW program. Identical amino acids and amino acids that match four of the five templates are indicated by yellow shading. The boxed sequences represent the seven highly conserved signature sequences (C1–C7). Residues that are also conserved in C domains are marked by stars above the sequence.

< 1% for both samples. With BacA1–AT, a small side preference for leucine (4.5%), phenylalanine (4%) and valine (3%) was detected, whereas the activities for the remaining amino acids were in the same range as background. In the assay with BacA2–AT, no comparable alternative activation was identified. The specific activation of isoleucine and cysteine strongly supports the statement that *bacA* encodes for the bacitracin synthetase BA1 and it can be assumed that *bacB* and *bacC* represent the structural genes encoding BA2 and BA3.

Figure 3

The first two BA1 AT domains. Coomassie Brilliant Blue stained SDS–polyacrylamide gels of *E. coli* overexpressed AT domains **(a)** BacA1–AT and **(b)** BacA2–AT. Lane 1, whole cell extract before induction; lane 2, whole cell extract after 2 h induction with IPTG; lane 3, pellet; lane 4, supernatant after lysis; lane 5 purified protein.



Disruption of the *bacB* gene

A *bacB* gene-disruption mutant was constructed in order to confirm the role of the *bac* gene products in the biosynthesis of bacitracin. A special strategy for genetic modification in *B. licheniformis* had to be performed, because *B. licheniformis* lacks natural competence, and the cells are also unable to perform of integrative recombination when transformed through the protoplast method. We therefore constructed the integrative plasmid pBAC–BD which bears a temperature-sensitive origin of replication (REPts) [51] and an internal fragment of the *bacB* gene with an inserted kanamycin-resistance cassette (Figure 5) [52]. This plasmid was transformed into *B. licheniformis* protoplasts at a permissive temperature below 32°C and kanamycin-resistant (Kan^r) colonies were isolated. A heat-shock treatment (48°C) of these colonies in the presence of kanamycin was performed in order to inactivate REPts and to force an integrative recombination event. Ten Kan^r resistant colonies were obtained which, when applied to a microbiological growth-inhibition test with *Micrococcus flavus*, showed no bacitracin production. One of these colonies, designated *B. licheniformis* AK1, was further investigated. Southern blot analysis using the kanamycin-resistance cassette and the internal *bacB* fragment BAC–BE (Figure 5) as probes demonstrated that the resistance cassette had integrated into the chromosomal *bacB* gene by a double-crossover recombination event (data not shown). To confirm the results of the microbiological assay culture, broth from *B. licheniformis* strains ATCC 10716 and AK1 was extracted with butanol, and analyzed by high performance liquid chromatography (HPLC; Figure 6). A prominent elution peak at 95 min was detected for the wild-type extract but was completely absent from the AK1 extract. Using mass spectrometric analysis of the fraction corresponding to the 95 min peak of the wild type extract, a mass of 1,421 was measured, which correlates perfectly with the calculated mass of 1,421.8 Da for bacitracin A. These results clearly

Figure 4

Results of amino acid-dependent ATP–PP_i exchange experiments with the purified AT domains BacA1–AT and BacA2–AT. Assays were performed with all 20 proteinogenic amino acids plus ornithine and one control without amino acids.

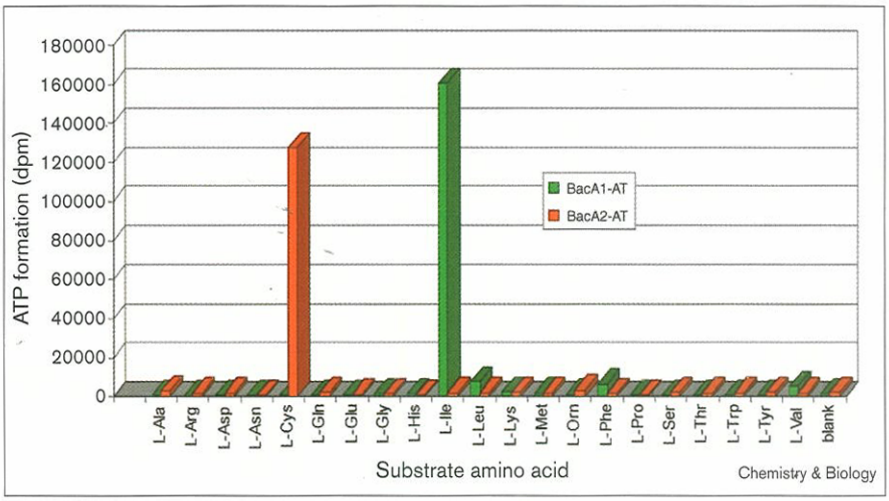
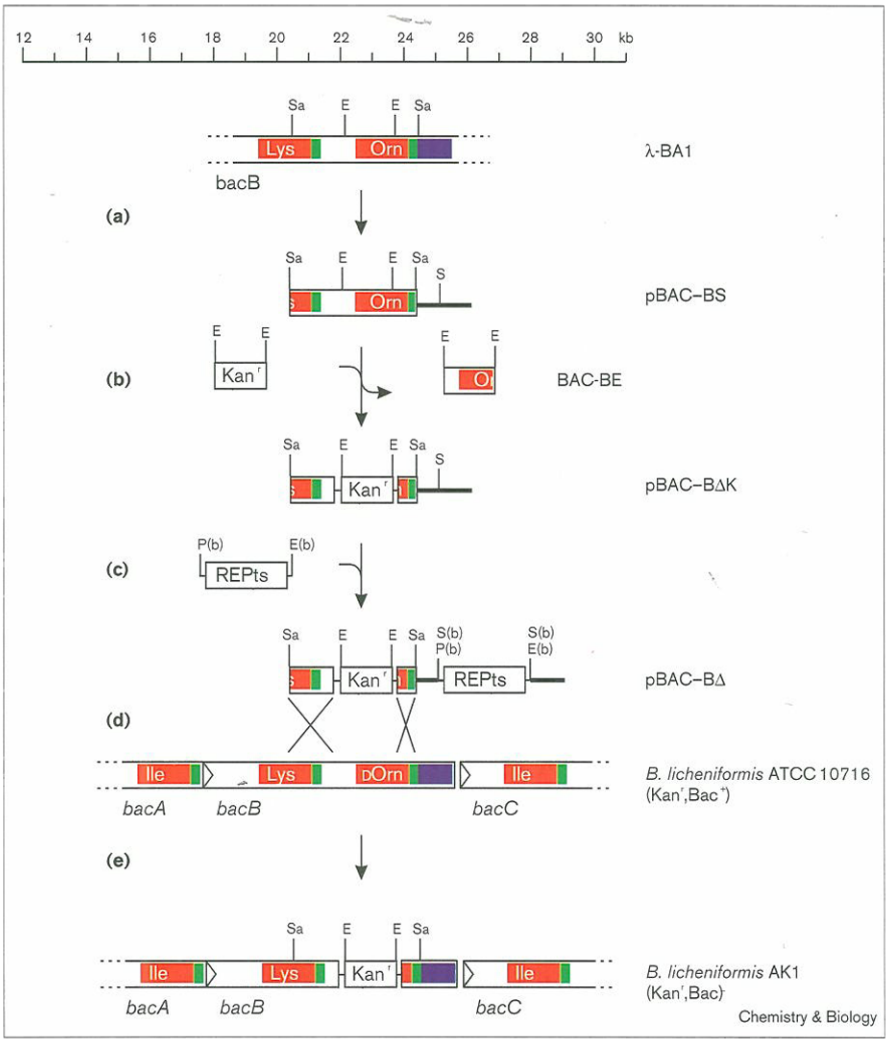
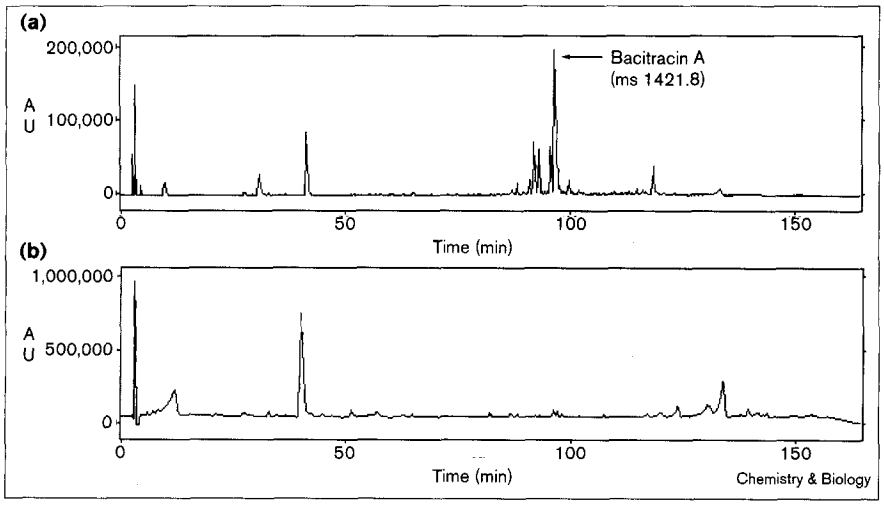


Figure 5



Construction of a *bacB* gene-disruption mutant. **(a)** An internal 3,897 bp *Sac*II (Sa) fragment from the *bacB* gene isolated from λ-BA1 was cloned into pBluescipt SK(–). **(b)** The resulting plasmid pBAC–BS was linearized with *Eco*RI (E) and the vector-containing portion was ligated with a kanamycin-resistance cassette from pDG783 [52] giving the plasmid pBAC–BΔK. **(c)** This plasmid was digested with *Sal*I (S), blunted and ligated with a blunt-ended DNA fragment containing the temperature-sensitive origin of replication (REPts) derived from from pLTV1 [51], resulting in the plasmid pBAC–BΔ. **(d)** The pBAC–BΔ plasmid was then used to transform *B. licheniformis* ATCC 10716 protoplasts, which were screened after regeneration for kanamycin-resistant colonies. **(e)** Inactivation of REPts by a heat-shift treatment to 48°C forced an integrative recombination into the chromosomal *bacB* gene, which yielded a bacitracin-deficient AK1 mutant. Abbreviations: P, PstI; (b), blunt ends; Kan^r, kanamycin-resistant; Bac$^{+/-}$, bacitracin producing/not producing.

Figure 6



HPLC analysis of butanolic culture broth extracts from **(a)** *B. licheniformis* ATCC 10716 and **(b)** *B. licheniformis* AK1. The bacitracin A peak and the result of the mass spectroscopy analysis of this peak are indicated.

indicate that *bacB* is located in the bacitracin-specific transcriptional unit.

Significance

This report presents the molecular characterization of the bacitracin synthetase genes. Bacitracin is a branched cyclic dodecylpeptide antibiotic ({Ile–Cys}$_{thiazoline}$–Leu–D-Glu–Ile–{Lys–D-Orn–Ile–D-Phe–His–D-Asp–Asn}$_{cyclic}$), produced in *B. licheniformis* ATCC 10716 via a nonribosomal pathway by a large multienzyme complex which contains three peptide synthetases, BA1, BA2 and BA3 [13–15]. The genes encoding the three bacitracin synthetases were cloned and sequenced, and were found to be organized in an operon. A modular structure typical of peptide synthetases [18] was detected, within the identified genes, and the domains identified match the categories of biochemical activity required for bacitracin synthesis. BA1 is encoded by *bacA*, the first gene of the identified operon, and is thought to activate and incorporate the five amino-terminal amino acids into bacitracin ({isoleucine–cysteine}$_{thiazoline}$–leucine–D-glutamine–isoleucine). BA1 is comprised of five modules, which include adenylation (A) and thioesterification (T) modules. These two modules are linked by putative condensation (C) domains, which are thought to catalyze the transpeptidation of the activated amino acids, thus promoting peptide chain growth. Additionally, an internal epimerization (E) domain was found attached to the fourth module. BA2 is encoded by *bacB*, and is responsible for the activation and incorporation of two amino acids (lysine–D-ornithine). Two modules and a carboxy-terminal E domain were identified in BA2. The deduced amino-acid specificity of the first two BA1-modules was confirmed by cloning, heterologous expression in *E. coli* and ATP–PP$_i$ exchange assays with purified AT domain proteins. Furthermore, we demonstrated the relevance of the cloned genes in bacitracin biosynthesis by constructing a bacitracin-deficient *bacB* gene-disruption mutant. Like BA1, the bacitracin synthetase BA3 is also responsible for incorporating five amino acids (isoleucine–D-phenylalanine–histidine–D-aspartate–asparagine) into bacitracin. The *bacC* gene encoding BA3 contains five modules with two internal E domains in the expected locations. Furthermore a putative carboxy-terminal thioesterase domain was detected, as has been reported for other bacterial peptide synthetases [17,20]. Besides these classical peptide synthetase domains, a strongly modified C domain was found between the first two modules of BA1. This domain shares the highest homology (35–50%) with proteins involved in the biosynthesis of thiazoline-containing siderophores (bacterially produced iron-chelating compounds), such as yersiniabactin and anguibactin [38,39]. These findings, in combination with the fact that bacitracin also contains a thiazoline ring between its first two amino acids, strongly support the idea that the modified C domain could be responsible for thiazoline ring formation.

Materials and methods

Bacterial strains and growth conditions

Bacillus licheniformis ATCC 10716 was maintained on agar plates of DSM sporulation medium and grown in rich medium (2xYT) at 37°C for DNA isolation procedures. *Escherichia coli* XL1 Blue (Stratagene, Heidelberg, Germany) was used for preparation of recombinant plasmids. *E. coli* XL1 Blue MRA(P2) was used as the host for the λ-EMBL3 genomic library and preparations of λ-phages. Overexpression of recombinant proteins was carried out in *E. coli* M15 (Qiagen, Hilden, Germany). For the construction of a gene disruption mutant, derivatives of the plasmids pDG783 [52] and pLTV1[51] were used. *Micrococcus flavus* ATCC 10240 was used as the test strain for microbiological bacitracin assays.

Transformation of E. coli and DNA manipulations

Standard genetic techniques for DNA manipulations *in vitro*, and cloning and the transformation of competent *E. coli* cells were used [53]. Total DNA from *B. licheniformis* was obtained using a lysozyme treatment and phenol–chloroform extraction [54]. For the construction of a λ-EMBL3 genomic library (Stratagene), chromosomal DNA of *B. licheniformis* ATCC 10716 was partially digested with *Sau*3AI and size-fractionated in 15–40% (w/v) NaCl gradients. Fractions containing DNA fragments of 13–22 kb were pooled and ligated to λ-EMBL3-arms digested with *Bam*HI. *In vitro* packaging was performed with Gigapack III Gold (Stratagene). For Southern blots and plaque filter hybridizations, the ECL random prime-labelling and detection system (Amersham/Buchler, Braunschweig, Germany) and positively charged nylon membranes (Amersham/Buchler) were used according to the manufacturer's protocol.

Identification, cloning and sequencing of the bac *operon*

For the identification of the bacitracin synthetase genes, a PCR method previously described [42] was employed, using degenerate primers derived from core sequences 4 and 6 of peptide synthetases [27]. The sequences of the degenerate oligonucleotides used were as follows (nucleotides in parentheses are degenerate; I, inosine): oligo TGD: 5′-T(AT)(CT) CGI ACI GGI GA(TC) (CT)(TG)I G(TG)I CG-3′ and LGG: 5′-A(TA)I GA(GA) (TG)(CG)I CCI CCI (GA)(GA)(GC) I(AC)(AG) AA(GA) AA-3′. All primers used in this study were synthesized by MWG Biotech (Ebersberg, Germany). Variation of the annealing temperature over a range between 40°C to 60°C yielded a set of different fragments which were isolated and sequenced. The derived amino-acid sequence of PCR fragment 05 (Figure 1d) showed high homology with other peptide synthetases and was used as a probe in the subsequent screening of the *B. licheniformis* genomic λ-EMBL3-library.

In all, around 8000 plaques were screened with approximately 500 plaques per agar plate (8.5 cm in diameter), and the phage clone λ-BA1 (Figure 1d) was isolated. The insert DNA was mapped and subcloned in pBluescript SK(–) using the restriction enzymes *Sal*I and *Eco*RI. Sequencing of the double-stranded pBluescript plasmids was performed with standard universal/SK primers and walking primers derived from the determined sequences. Small gaps between the subclones were filled by PCR amplification and sequencing of the product or by sequencing of the λ-DNA template. In a second round of screening, using 5′- and 3′-terminal fragments of the λ-BA1-insert as probes, the phage clones λ-A14.3 and λ13.2 were isolated among others. The inserts of these phages were also subcloned and sequenced in the same way as λ-BA1.

Sequencing reactions were carried out using the chain-termination method [55] with dye-labeled dideoxy terminators from the PRISM Ready Reaction DyeDeoxy Terminator Cycle Sequencing Kit with AmpliTaq FS polymerase (Applied Biosystems) according to the manufacturer's protocol, and analyzed on the ABI 310 genetic analyzer.

Amplification and cloning of BA1 AT domains

PCR amplification of the *bacA* AT domains (BacA1–AT and BacA2–AT) was used to generate the terminal restriction sites for the

subsequent cloning and was performed with the Long Range PCR System (Boehringer, Mannheim, Germany) following the manufacturer's protocol. The sequences of the 5′-modified oligonucleotides used were as follows (italics, modified sequences; bold, restriction sites): oligo 5′-Ile1AT(*NcoI*): 5′-*TAT* ***ACC ATG*** GTT GCT AAA CAT TCA TT-3′, 3′-Ile1AT(*BglII*): 5′-*TAA* ***CAG ATC T***GC CAA CCA AAT AAG C-3′, 5′-Cys2AT(*NcoI*): 5′-*TAT* ***ACC ATG*** GTT TCA ACC ATC CAA ATG-3′ and 3′-Cys2AT(*BglII*): 5′-*TAA* ***CAG ATC*** TTT GTT GGG CAG GG-3′. The pQE60 His^6-tag fusion vector (Qiagen) digested with *NcoI*/*BglII* was used to clone the AT-domain fragments. The fusion sites between the vector and insert were confirmed by sequencing with the following primers: oligo 5′-promotor: 5′-GGC GTA TCA CGA GGC CC and 3′-terminator: 5′-ACG CCC GGC GGC AAC CG-3′.

Expression, purification and ATP–PP$_i$-exchange of the His6-tagged AT domains

The pQE60-derivatives containing the amplified AT domains (designated pQE-BacB1AT and pQE-BacB2AT; Figure 1e) were transformed in *E. coli* strain M15(pREP4) and expressed as previously described [27], except that cells were induced with 1 mM IPTG at A_{600} 0.7 and allowed to grow for an additional 2 h before harvesting. Purification of the His^6-tagged proteins (designated BacA1–AT and BacA2–AT), was basically carried out as described [27] using Ni^{2+}-affinity chromatography, but cells were broken by three passages through a French Press. Expression and purification was checked by Coomassie Brilliant Blue-stained SDS–polyacrylamide gels (Figure 3) and protein concentration was measured using the Bradford method [56]. The amino-acid-dependent ATP–PP$_i$-exchange experiments were performed as described previously [27].

Construction of a bacB *gene disruption mutant*

An internal 3,897 bp *Sac*II DNA fragment of the *bacB* gene, isolated from λ-BA1, was cloned in pBluescript SK(–)ΔE, where the *Eco*RI restriction site had been deleted by digestion with *SmaI*/*EcoRV*, followed by re-ligation of the vector. This plasmid pBAC–BS (Figure 5), was digested with *Eco*RI, the internal fragment BAC–BE (1,368 bp) was separated and the fragment containing the vector portion (5,469 bp) was ligated with an *Eco*RI-digested kanamycin-resistance cassette (KAN^r) from pDG783 [52]. The resulting plasmid pBAC-BΔK was linearized with *SalI*, blunted and ligated with a blunt cut *PstI*/*Eco*RI DNA fragment carrying the temperature-sensitive origin of replication (REPts) from pLTV1 [51]. The product plasmid pBAC–BΔ was transformed in *B. licheniformis* protoplasts as described [57] with modified regeneration conditions according to Gray and Chang [58] and kanamycin as selection marker. In order to inactivate the REPts and to force integration into the chromosome, the isolated Kan^r transformants were subjected to a two-stage dilution and heat-shift treatment (48°C) [59]. Small discs of Whatmann paper (5 mm in diameter) were then soaked with cultures of heat-stable Kan^r colonies and were placed on freshly prepared *Micrococcus flavus* ATCC 10240 test plates [15]. After incubation overnight at 37°C, bacitracin-deficient (and kanamycin-resistant) mutants could be identified by a missing growth-inhibition zone of *M. flavus*. The mutant *B. licheniformis* AK1 (Kan^r,Bac$^-$) was isolated and the integration of the kanamycin-cassette into the chromosomal *bacB* gene was verified by Southern blots, PCR amplification of the chromosomal region corresponding to the *bacB* gene, and sequencing of the PCR product (data not shown).

Analysis of bacitracin production

M20 medium (200 ml) [60] was inoculated with *B. licheniformis* and incubated with shaking for 30 h at 37°C. The cells were harvested by centrifugation and the supernatant was extracted with 2 volumes of *n*-butanol. The organic phase was evaporated to complete dryness and resuspended in 50% (v/v) methanol / 0.1 % (v/v) trifluoroacetic acid. This suspension was analyzed by HPLC for its bacitracin content. Reverse-phase HPLC analysis was performed by using a Sephasil TM C^{18}-column (5 µm; 250 x 4 mm; Pharmacia LKD Biotechnology, Freiburg, Germany) and a 0–63% (w/v) gradient of acetonitrile / 0.1% (v/v) trifluoroacetic acid as mobile phase. Bacitracin was detected by absorption at 214 nm and 253 nm. HPLC fractions corresponding to absorption peaks were analyzed by mass spectrometry or used in a biological assay. For this assay the fractions were evaporated to complete dryness, resuspended in water and loaded on small discs of Whatmann paper (0.5 mm in diameter) placed on freshly prepared *M. flavus* test plates [15]. After incubation overnight at 37°C, *M. flavus* growth-inhibition zones were measuredx.

Mass spectrometric analysis was performed by 'Abt. f. Gaschromatographie u. Massenspektrometrie, FB Chemie, Phillipps-Universität Marburg. Ionization of samples was performed by electrospray ionization (Hewlett Packard 59987A; ion source: Analytica of Bradford). Spectra were recorded using a Hewlett Packard 5989B mass spectrometer.

Accession numbers

The nucleotide sequences from *B. licheniformis* ATCC 10716 described in this paper have been submitted to the GenBank database under the accession number AF007865 for the region containing the peptide synthetases *bacA*, *bacB* and *bacC*.

Acknowledgements

We are grateful to Inge Schüler and Gabriele Schimpf-Weiland for technical assistance and to Philipp J. Youngman for plasmid pLTV1. We also thank Paul Weinreb and Luis Quadri for assistance in some of the presented experiments. This work was supported by Biochemie GmbH, 6330 Kufstein/Schaftenau, Austria and by the Deutsche Forschungsgesellschaft.

References

1. Haavik, H.I. & Froyshov, O. (1975). Function of peptide antibiotics in producer organisms. *Nature* **254**, 79-82.
2. Johnson, B.A., Anker, H. & Meleney, F.L. (1945). Bacitracin: a new antibiotic produced by a member of the *B. subtilis* group. *Science* **102**, 376-377.
3. Azevedo, E.C., Rios, E.M., Fukushima, K. & Campos-Takaki, G.M. (1993). Bacitracin production by a new strain of *Bacillus subtilis*. Extraction, purification, and characterization. *Appl. Biochem. Biotechnol.* **42**, 1-7.
4. Morris, M. (1994). Primary structural confirmation of components of the bacitracin complex. *Biol. Mass Spectrom.* **23**, 61-70.
5. Stone, K.J. & Strominger, J.L. (1971). Mechanism of action of bacitracin: complexation with metal ions and C_{55}-isoprenyl pyrophosphate. *Proc. Natl Acad. Sci. USA* **68**, 3223-3227.
6. Storm, D.R. & Strominger, J.L. (1973). Complex formation between bacitracin peptides and isoprenyl pyrophosphates. *J. Biol. Chem.* **248**, 3940-3945.
7. Cain, B.D., J, N.P., W, E., Nick, H.S. & Allen, C.M. (1993). Amplification of the *bacA* gene confers bacitracin resistance to *Escherichia coli*. *J. Bacteriol.* **175**, 3784-3789.
8. Pollock, T.W., Thorne, L., Yamazaki, M., Mikolajczak, M.J. & Armentrout, R.W. (1994). Mechanism of bacitracin resistance in Gram-negative bacteria that synthesize exopolysaccarides. *J. Bacteriol.* **176**, 6229-6237.
9. Rieber, M.T., Imaeda, T. & Cesari, I.M. (1969). Bacitracin action on membranes of *Mycobacteria*. *J. Gen. Microbiol.* **55**, 155-159.
10. Makinen, K.K. (1972). Inhibition of bacitracin of some hydrolytic enzymes. *Int. J. Pept. Protein Res.* **4**, 21-28.
11. Schechter, N., Momose, K. & Rudney, H. (1972). The effect of bacitracin and Ca^{++} on the formation of polyprenylphosphates and their incorporation into ubiquinone precursors in mammalian and bacterial systems. *Biochem. Biophys. Res. Commun.* **48**, 833-839.
12. Kleinkauf, H. & von Dohren, H. (1987). Biosynthesis of peptide antibiotics. *Annu. Rev. Microbiol.* **41**, 259-289.
13. Ishihara, H., Endo, Y., Abe, S. & Shimura, K. (1975). The presence of 4′-phosphopantheine in the bacitracin synthetase. *FEBS Lett.* **50**, 43-46.
14. Frøyshov, Ø. (1975). Enzyme-bound intermediates in the biosynthesis of bacitracin. *Eur. J. Biochem.* **59**, 201-206.
15. Rieder, H., Heinrich, G., Breuker, E., Simlot, M.M. & Pfaender, P. (1975). Bacitracin synthetase. *Meth. Enzymol.* **43**, 548-559.
16. Lipmann, F. (1980). Bacterial production of antibiotic polypeptides by thiol-linked synthesis on protein templates. *Adv. Microb. Physiol.* **21**, 227-266.

17. Marahiel, M. A. (1997). Protein templates for the biosynthesis of peptide antibiotics. *Chem. Biol.* **4**, 561-567.
18. 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.
19. Kleinkauf, H. & Von Döhren, H. (1996). A nonribosomal system of peptide biosynthesis. *Eur. J. Biochem.* **236**, 335-351.
20. Marahiel, M.A., Stachelhaus, T. & Mootz, H.D. (1997). Modular peptide synthetases involved in non-ribosomal peptide synthesis. *Chem. Rev.* **97**, 2651-2674.
21. Cosmina, P., *et al.* & van Sinderen, D. (1993). Sequence and analysis of the genetic locus responsible for surfactin synthesis in *Bacillus subtilis*. *Mol. Microbiol.* **8**, 821-831.
22. 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.
23. 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.***179**, 6843-6850.
24. 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.
25. Weber, G., Schörgendorfer, K., Schneider-Scherzer, E. & Leitner, E. (1994). The peptide synthetase catalyzing cyclosporine production in *Tolypocladium niveum* is encoded by a giant 45.8-kilobase open reading frame. *Curr. Genet.*. **26**, 120-125.
26. 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.
27. 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.
28. 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.
29. Schlumbohm, W., *et al.* & Wittman-Liebold, B. (1991). An active serine is involved in covalent substrate amino acid binding at each reaction center of gramicidin S synthetase. *J. Biol. Chem.* **266**, 23135-23141.
30. Stachelhaus, T., Schneider, A. & Marahiel, M.A. (1996). Engineered biosynthesis of peptide antibiotics. *Biochem. Pharmacol.* **52**, 177-186.
31. Lambalot, R.H., *et al.* & CT Walsh (1996). A new enzyme superfamily - the phosphopantetheinyl transferases. *Chem. Biol.* **3**, 923-936.
32. Stein, T., *et al.* & Morris, H.R. (1996). The multiple carrier model of nonribosomal peptide biosynthesis at modular multienzymatic templates. *J. Biol. Chem.* **271**, 15428-15435.
33. De Crecy-Lagard, V., Marliere, P. & Saurin, W. (1995). Multienzymatic non ribosomal peptide biosynthesis: identification of the functional domains catalysing peptide elongation and epimerisation. *C R Acad. Sci. III* **318**, 927-936.
34. Stein, T., Kluge, B., Vater, J., Franke, P., Otto, A. & Wittmann-Liebold, B. (1995). Gramicidin S synthetase 1 (phenylalanine racemase), a prototype of amino acid racemases containing the cofactor 4'-phosphopantetheine. *Biochemistry* **34**, 4633-4642.
35. Ishihara, H., Hara, N. & Iwabuchi, T. (1989). Molecular cloning and expression in *Escherichia coli* of the *Bacillus licheniformis* bacitracin synthetase 2 gene. *J. Bacteriol.* **171**, 1705-1711.
36. Herzog-Velikonja, B., Podlesek, Z. & Grabnar, M. (1994). Isolation and characterization of Tn917-generated bacitracin deficient mutants of *Bacillus licheniformis*. *FEMS Microbiol. Lett.* **121**, 147-152.
37. Ishihara, H. & Shimura, K. (1979). Thiazoline ring formation in bacitracin biosynthesis. *FEBS Lett.* **99**, 109-112.
38. Guilvout, I., Mercereau-Pauijalon, O., Bonnefoy, S., Pugsley, A.P. & Carniel, E. (1993). High-molecular-weight protein 2 of *Yersinia enterocolitica* is homologous to AngR of *Vibrio anguillarum* and belongs to a family of proteins in nonribosomal peptide synthesis. *J. Bacteriol.* **175**, 5488-5504.
39. Tolmasky, M.E., Actis, L.A. & Crosa, J.H. (1993). A single amino acid change in AngR, a protein encoded by pJM1-like virulence plasmids, results in hyperproduction of anguibactin. *Infect. Immun.* **61**, 3228-3233.
40. Serino, L., Reimmann, C., Visca, P., Beyeler, M., Chiesa, V.D. & Haas, D. (1997). Biosynthesis of pyochelin and dihydroaeruginoic acid requires the iron-regulated pchDCBA operon in *Pseudomonas aeruginosa*. *J. Bacteriol.* **179**, 248-257.
41. Philipp, W.J., *et al.*, & Jacobs, E.R., Jr. (1996). An integrated map of the genome of the tubercle bacillus, *Mycobacterium tuberculosis* H37Rv, and comparison with *Mycobacterium leprae*. *Proc. Natl Acad. Sci. USA* **93**, 3132-3137.
42. Turgay, K. & Marahiel, M.A. (1994). A general approach for identifying and cloning peptide synthetase genes. *Pept. Res.* **7**, 238-241.
43. 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.
44. Mittenhuber, G., Weckermann, R. & Marahiel, M.A. (1989). Gene cluster containing the genes for tyrocidine synthetases 1 and 2 from *Bacillus brevis*: evidence for an operon. *J. Bacteriol.* **171**, 4881-4887.
45. Gish, W. & States, D.J. (1993). Identification of protein coding regions by database similarity search. *Nat. Genet.* **3**, 266-272.
46. Bearden, S.W., Fetherston, J.D. & Perry, R.D. (1997). Genetic organization of the yersiniabactin biosynthetic region and construction of avirulent mutants in *Yersinia pestis*. *Infect. Immun.* **65**, 1659-1668.
47. Jalal, M.A.F., Hossian, M.B., van der Helm, D., Sanders-Loehr, J., Actis, L.A. & Crosa, J.H. (1989). Structure of anguibactin, a unique plasmid-related bacterial siderophore from the fish pathogen *Vibrio anguillarum*. *J. Am. Chem. Soc.* **111**, 292-296.
48. Chambers, C.E., McIntyre, D.D., Mouck, M. & Sokol, P.A. (1996). Physical and structural characterization of yersiniophore, a siderophore produced by clinical isolates of *Yersinia enterocolitica*. *Biometals* **9**, 157-167.
49. Cox, C.D. (1981). Pyochelin: novel structure of an iron-chelating growth promotor for *Pseudomonas aeruginosa*. *Proc. Natl Acad. Sci. USA* **78**, .4256-4260.
50. Li, Y.M., Milne, J.C., Madison, L.L., Kolter, R. & Walsh, C.T. (1996). From peptide precursors to oxazole and thiazole-containing peptide antibiotics: microcin B17 synthase. *Science* **274**, 1188-1193.
51. Youngman, P., Poth, H., Green, B., York, K., Olmedo, G. & Smith, K. (1989). Methods for genetic manipulations, cloning and functional analysis of sporulation genes in *Bacillus subtilis*. In *Regulation of Prokaryotic Development*. (Smith, I., Slepecky, R. & Setlow, P., eds.), pp. 65-87, American Society for Microbiology, Washington, DC.
52. Guerout-Fleury, A.M., Shazand, K., Frandsen, N. & Stragier, P. (1995). Antibiotic-resistance cassettes for *Bacillus subtilis*. *Gene* **167**, 335-336.
53. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
54. Cutting, S.M. & Vander Horn, P.B. (1990). Genetic analysis. In *Molecular Biological Methods for Bacillus*. (Harwood, C.R. & Cutting, S.M., eds.), pp.27-74, John Wiley & Sons Ltd., New York.
55. Sanger, F., Miklen, S. & Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl Acad. Sci. USA* **74**, 5463-5467.
56. Bradford, M.M. (1976). A rapid and sensitve method for the quantitation of microgramm quantities using the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254.
57. Quax, W. (1990). *Bacillus licheniformis* protoplast transformation. In *Molecular Biological Methods for Bacillus*. (Harwood, C.R. & Cutting, S.M., eds.), p. 152, John Wiley & Sons Ltd., New York.
58. Gray, O. & Chang, S. (1981). Molecular cloning and expression of *Bacillus licheniformis* beta-lactamase gene in *Escherichia coli* and *Bacillus subtilis*. *J. Bacteriol.* **145**, 422-428.
59. Youngman, P. (1990). Use of transposons and integrational vectors. In *Molecular Biological Methods for Bacillus*. (Harwood, C.R. & Cutting, S.M., eds.), pp.221-266, John Wiley & Sons Ltd., New York.
60. Haavik, H.I. (1974). Studies on the formation of bacitracin by *Bacillus licheniformis*: effect of glucose. *J. Gen. Microbiol.* **81**, 383-390.