Methods for the preparation of (lipo)peptide synthetases and (lipo)peptides produced therewith

Leenhouts, Johannes Cornelis; Noback, Andries Michiel; Van Den Burg, Lambertus; Hamoen, Leendert Willem; Duitman, Erwin Hans; Kuipers, Oscar Paul

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 May 2001 (10.05.2001)

PCT

(10) International Publication Number
WO 01/32845 A2

(51) International Patent Classification[7]: C12N 9/00

(21) International Application Number: PCT/EP00/11237

(22) International Filing Date:
6 November 2000 (06.11.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
99203674.9 5 November 1999 (05.11.1999) EP

(71) Applicant *(for all designated States except US)*: **BIOMADE B.V.** [NL/NL]; Nijenborgh 4, NL-9747 AG Groningen (NL).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **LEENHOUTS, Cornelis, Johannes** [NL/NL]; Tussenkoelen 6, NL-9753 KX Haren (NL). **NOBACK, Michiel, Andries** [NL/NL]; Bottemaheerd 103, NL-9737 NC Groningen (NL). **VAN DEN BURG, Lambertus** [NL/NL]; Sint Marienweg 23, NL-8161 HD Epe (NL). **HAMOEN, Leendert, Willem** [NL/NL]; Moesstraat 9s, NL-9717 JT Groningen (NL). **DUITMAN, Erwin, Hans** [NL/NL]; Coendersweg 86, NL-9722 GK Groningen (NL). **KUIPERS, Oscar, Paul** [NL/NL]; Ratelaarweg 19, NL-9753 BD Haren (NL).

(74) Agent: **VAN SOMEREN, Petronella, Francisca, Hendrika, Maria**; Arnold & Siedsma, Sweelinckplein 1, NL-2517 GK The Hague (NL).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/32845 A2

(54) Title: METHODS FOR THE PREPARATION OF (LIPO)PEPTIDE SYNTHETASES AND (LIPO)PEPTIDES PRODUCED THEREWITH

(57) Abstract: The present invention relates to methods for the preparation of novel lipopeptide synthetases, to the lipopeptide synthetases thus obtained and the lipopeptides that are produced by these synthetases. The method involves modifying at least the lipid moiety and optionally also the peptide moiety of one or more known (lipo)peptide synthetase. The invention further provides DNA molecules encoding such lipopeptide synthetase, libraries containing DNA fragments encoding modules of such lipopeptide synthetases and the use of both.

METHODS FOR THE PREPARATION OF (LIPO)PEPTIDE SYNTHETASES AND (LIPO)PEPTIDES PRODUCED THEREWITH

The present invention relates to a method for the preparation of novel lipopeptide synthetases, to the lipopeptide synthetases thus obtained, and to lipopeptides that are the product of the enzymatic action of the said lipopeptide synthetases. The invention furthermore relates to a novel method for preparing lipopeptide and peptide synthetases in a controlled or semi-random manner and to the novel lipopeptide and peptide synthetases thus obtained as well as the lipopeptides and peptides that are the product of the enzymatic action of the said (lipo)peptide synthetases. In addition, the invention provides for engineered microorganisms capable of expressing the novel (lipo)peptide synthetases. The invention also relates to DNA molecules encoding the novel (lipo)peptide synthetases, to host cells harboring such DNA molecules and to DNA libraries containing (lipo)peptide synthetase module or domain encoding DNA fragments.

In the struggle for nutrients, microorganisms have a wide arsenal of chemical compounds at their disposal to inhibit competing organisms. A specific class of these compounds are the peptide antibiotics, e.g. surfactin, mycosubtilin, fengycin, iturin, gramicidin, etc. Such peptides can be linear, cyclic, or branched, contain D-, hydroxyl-, N-methylated, or other unusual amino acid constituents, and can be modified by acylation or glycosylation. A number of biological properties, among which antimicrobial, antiviral, antitumor, enzyme inhibiting and immunosuppressant activities, make them interesting from both a biotechnological and pharmaceutical point of view.

These peptides are produced by a non-ribosomal pathway according to the thiotemplate mechanism (Marahiel et al. (1997) Chem. Rev. 97, 2651-2673). In contrast to ribosomal synthesis of proteins and peptides, which is

limited by the availability of only 20 amino acids, non-ribosomal synthetases have been shown to be able to incorporate more than 300 different precursors (Kleinkauf & von Döhren (1990) Eur. J. Biochem. 192, 1-15). The peptide chains range from 2 to 48 residues.

The Gram-positive bacterium Bacillus subtilis also produces such non-ribosomally synthesized peptides, among which a variety of circular oligopeptides that contain a fatty acid, so-called lipopeptides. The lipopeptides surfactin and fengycin contain a ß-hydroxy fatty acid, whereas members of the iturin lipopeptide family, such as mycosubtilin, bacillomycin, and iturin, carry a ß-amino fatty acid modification.

For surfactin, several interesting biochemical properties have been demonstrated or suggested. Among these are its high effectiveness as biosurfactant. Furthermore it has ionophoric properties, antitumor activity, antiviral activities, membrane activity resulting in lysis of bacterial protoplasts and mycoplasma, cholesterol lowering activity, acts as a hemolytic agent, is known to lyse erythrocytes and inhibits fibrin clot formation and certain enzymes, e.g. Ca^{2+}- and Mg^{2+}-dependent cAMP-phosphodiesterase and H^+, K^+-ATPases, inhibits gamete maturation, and improves the biodegradation of oil in soil. The members of the iturin lipopeptide family exhibit strong antifungal and hemolytic activities and a limited antibacterial activity.

As already indicated above, the peptide antibiotics are synthesized by large multifunctional enzymes, the so-called peptide synthetases. Such peptide synthetases can be found in prokaryotes as well as in eukaryotes. Genetic and biochemical analyses of peptide synthetases have revealed a modular structure of these multifunctional proteins. A module is defined as a unit that catalyzes the incorporation of a specific amino acid into the peptide product. The arrangement of the modules

of a peptide synthetase is usually co-linear with the amino acid sequence of the peptide.

The modules can be further subdivided into different domains that are characterized by a set of short conserved sequence motifs and have particular functions within a module.

The core of each module is an amino acid adenylation domain that recognizes and activates a specific amino acid. The thiolation domain, located C-terminally of the adenylation domain, contains an invariant serine residue essential for the binding of a 4'-phosphopantetheine cofactor. An N-terminal condensation domain is required for the coupling of two consecutive amino acids. In addition, modules can be supplemented with domains that catalyze the modification of the activated amino acid, such as N-methylation, and epimerization from the L-configuration to the D-configuration.

To manipulate the specificity of the synthetases, the DNA fragments encoding individual modules can be exchanged using replacement recombination procedures. Additionally, shorter peptide synthetases can be produced by the introduction of a thioesterase domain and transcription termination signals at predefined positions using genetic manipulation techniques. Furthermore, novel peptide synthetases with new specificities can be obtained by introducing mutations by means of chemical agents. However, due to its random nature, this procedure is characterized by low efficiency. Other disadvantages of these methods are the relatively limited number of different peptide synthetases that can be obtained by these means and the, often, time-consuming nature of the procedures employed.

The conserved modular organization of peptide synthetases both on amino acid level and DNA level has provided the means for the creation of new peptide synthetases by genetic engineering strategies. US-5,795,738 and US-5,652,116 both describe such methods for

creating new peptide synthetases and their use in the non-ribosomal production of peptides. These methods have not been described for lipopeptides, because the manner in which peptides are linked to a fatty acid in vivo is not known. Therefore, it is not possible to modify lipopeptides in vivo in their lipid moiety. Because the specificity and activity spectrum of lipopeptides is not only achieved by differences in the amino acid composition, but also by the presence and composition of the fatty acid moiety it is highly desirable to have a means available for modifying the genetic information encoding lipopeptides such that their lipid moiety and optionally also their peptide moiety is modified.

The objective of the present invention is thus to apply methodologies that provide the means to produce very large numbers of novel lipopeptides by inducing variations in fatty acid composition and/or amino acid composition.

In the research that led to the present invention mycosubtilin was studied and it was found that this peptide antibiotic is characterized by a ß-amino fatty acid moiety linked to the circular heptapeptide Asn-Tyr-Asn-Gln-Pro-Ser-Asn, with the second, third and sixth position present in the D-configuration. Mycosubtilin belongs to the iturin family of lipopeptide antibiotics. The gene cluster from B. subtilis ATCC6633 specifying the biosynthesis of mycosubtilin was identified. The putative operon spans 38 kb and consists of four open reading frames, designated fenF, mycA, mycB, and mycC, each with strong homologies to the family of peptide synthetases. Biochemical characterization showed that the first module of MycB specifically adenylates tyrosine, as expected for mycosubtilin synthetase, and insertional mutagenesis of the operon resulted in a mycosubtilin-negative phenotype.

On the basis of this lipopeptide it was now surprisingly found that lipopeptide synthetases contain lipid synthetase modules or domains as an integral part

of the enzyme. The mycosubtilin synthetase revealed a feature that is unique for peptide synthetases as well as for fatty acid synthases, namely that the MycA subunit combines functional domains derived from peptide synthetases, amino transferases, and fatty acid synthases. MycA represents the first example of a natural hybrid between these enzyme families.

The present inventors have furthermore identified and characterized the lipopeptide synthetase operon that is responsible for the synthesis of the potent antifungal lipopeptide mycosubtilin. The overall structure of this synthetase resembles that of other peptide synthetases. However, unlike any previously known case, attached to the first module (MycA) a large N-terminal domain was identified that shows significant homology to fatty acid and polyketide synthases. This domain was found to play an unique role in the incorporation of the ß-amino fatty acid moiety into the mycosubtilin molecule. The postulated mechanism involved in this incorporation is outlined in **Fig. 4.**

According to this mechanism a condensation domain in front of the first module is sufficient to catalyze the coupling of the fatty acid to the first amino acid. In all (lipo)peptide synthetases identified so far, comprising the surfactin, fengycin, lichenysin and mycosubtilin synthetases, the first module is preceded by a condensation domain. The fatty acid is delivered as an activated thioester attached to an acyl carrier protein, a common intermediate in bacterial fatty acid synthesis, and the condensation domains catalyze the transfer of the fatty acid to the first amino acid to be incorporated into the peptide chain. Thus, the mycosubtilin operon encodes a hybrid synthetase, integrating fatty acid synthesis and peptide synthesis. The unique combination of enzymatic domains in the mycosubtilin synthetase subunit MycA provides new possibilities for the engineering of lipopeptides.

This finding is the basis for the present invention which now enables the construction of novel lipopeptides that are modified as compared to existing lipopeptides in their fatty acid composition and if desired also in their amino acid composition.

The present invention thus provides new lipopeptide synthetases by modifying the lipid moiety and optionally also the peptide moiety of one or more known (lipo)peptide synthetase, by a method comprising the steps of:

a) providing a range of DNA fragments encoding lipopeptide synthetase modules or domains thereof;

b) selecting and ligating at least two of these DNA fragments, at least one of which is specific for the lipid moiety of the lipopeptide synthetase, in order to obtain a lipopeptide synthetase encoding DNA molecule;

c) introducing the lipopeptide synthetase encoding DNA molecule in a host cell;

d) expressing the DNA molecule in the host cell.

Preferably, the DNA fragments encoding lipopeptide synthetase modules or domains thereof comprise DNA fragments encoding fatty acid or polyketide synthase and (lipo)peptide synthetase modules or domains thereof.

When the lipopeptide synthetase encoding DNA molecule has a configuration of modules that does not exist in nature, the method of the invention leads to new combinations of lipid modules and amino acid modules and thus to new lipopeptide synthetases and lipopeptides that are also part of this invention.

Novel lipopeptide synthetases may be generated by coupling of lipid modules to existing peptide modules, by replacing existing lipid modules by others, by replacing/ deleting/adding peptide synthetase (PS) modules or by a combination of altering lipid and PS modules.

In addition to new combinations, existing modules within a known or novel synthetase can be further modified by mutating one or more of the DNA fragments prior to expression thereof. Such mutation can be effected either prior to ligation or after the DNA molecule is assembled. Suitable mutation methods are for example error-prone PCR or site-directed mutagenesis.

Ligation of the DNA fragments encoding the modules can be very specific, for example when particular lipid modules are to be coupled to existing sequences of amino acid modules, or random. In order to obtain the highest possible amount of variation, ligation is preferably random ligation using a library of different modules.

However, according to another aspect of the invention there is provided a method in which the order of the modules in the synthetase can be directed. This is achieved by generating the DNA fragments by PCR using primers containing restriction enzyme recognition sites that generate non-palindromic sticky ends allowing ligation in a predictable manner. This novel method is not only applicable to lipopeptide synthetases but also to peptide synthetases in general.

The invention thus relates to a method for the preparation of lipopeptide or peptide synthetases, which method comprises the steps of:

a) providing a range of DNA fragments encoding lipopeptide or peptide synthetase modules or domains thereof, which DNA fragments contain at their 5'-end and 3'-end restriction enzyme recognition sites that generate non-palindromic sticky ends;

b) selecting and ligating at least two of these DNA fragments, in order to obtain a lipopeptide or peptide synthetase encoding DNA molecule;

c) introducing the lipopeptide synthetase encoding DNA molecule in a host cell;

d) expressing the DNA molecule in the host cell.

In a preferred method of the invention, ligation is effected in the presence of polyethylene glycol (PEG) and T4 DNA ligase such that linear molecules are obtained. It was found that linear molecules can be easily transformed into Bacillus cells.

Preferably the DNA sequence is introduced into naturally competent host cells. Such cells exist in a certain growth phase of particular microorganisms. This method is specifically useful in case the host cell is a Bacillus species, in particular Bacillus subtilis.

The method of the invention may further comprise the step of selecting a host cell harboring a DNA sequence encoding a lipopeptide synthetase which produces a lipopeptide having a desired activity. The production of the lipopeptide can be performed in vivo within the host cell or in vitro. A host cell can either accumulate the lipopeptide within the cell or secrete the lipopeptide in the culture medium.

The selection can thus be effected by contacting the host cells or the culture medium used for growing the host cells with an indicator microorganism. The culture medium can still contain the host cell or the host cells can be removed. As an alternative, for cells that accumulate the lipopeptide, the cells can be disrupted, for example by sonication, either in the culture medium or after being separated therefrom.

Furthermore, the invention relates to a method for the non-ribosomal preparation of novel lipopeptides, comprising the preparation of a novel lipopeptide synthetase by means of the method as described above and using this lipopeptide synthetase to produce the corresponding lipopeptide.

The lipopeptide may either be produced in the host cell in which the lipopeptide synthetase is expressed or the lipopeptide synthetase may be isolated from the host cell or the culture medium used for growing the host cell prior to using the lipopeptide synthetase for the production of the corresponding lipopeptide.

The present invention thus provides the opportunity to develop large amounts of novel lipopeptide synthetases and their corresponding lipopeptides in a relatively easy manner. In order to be able to develop novel (lipo)peptide synthetases and their resulting (lipo)peptides a range of DNA fragments encoding PS and optionally FAS (fatty acid synthase) and/or optionally PKS (polyketide synthase) modules is to be provided.

In the method of the invention it may thus be practical to first amplify large numbers of individual single module encoding DNA fragments using the polymerase chain reaction. These single modules can be obtained from natural sources, or after random mutagenesis, e.g. by error-prone PCR, of a limited number of modules. A range of such DNA fragments can for example constitute a library.

Subsequently, in the large scale process of the invention there are various ways of producing the lipopeptide synthetases.

According to a first embodiment, the fragments corresponding to modules with different specificities will be ligated in a random order in the presence of PEG and T4 DNA ligase. By this means, PEG facilitates the formation of long linear DNA molecules and prevents the formation of circular DNA molecules that are usually the resultant of enzymatic ligations (Pfeiffer & Zimmerman, NAR 11: 7853-7871 (1983)).

Another embodiment makes use of a library in which DNA fragments are contained that have at their 5'- and 3'-ends restriction enzyme recognition sites that generate non-palindromic sticky ends such that the modules can be ligated in a predetermined order. Examples of suitable restriction enzymes are SfiI, Eco31I, Esp3I and BpiI. These sites are located in the linker regions as defined in Example 2.

When DNA fragments are used that contain at their 5'-end and 3'-end restriction enzyme recognition sites that generate non-palindromic sticky ends, novel

methods of ligation are possible. A method for ligating DNA fragments encoding lipopeptide or peptide synthetase modules or domains thereof is also part of this invention and comprises the steps of:

a) providing a solid surface having coupled thereto a first DNA fragment encoding a start DNA sequence comprising at its 3'-end the sticky end of a non-palindromic restriction enzyme recognition site;

b) ligating thereto a DNA fragment encoding a lipopeptide or peptide synthetase module or domain thereof, which DNA fragment contains at its 5'-end a sticky end that is complementary to the sticky end at the 3'-end of the previously incorporated DNA fragment and at its 3'-end the other sticky end of the non-palindromic restriction enzyme recognition site;

c) removing the non-ligated DNA fragments by washing; and

d) optionally repeating steps b) and c) one or more times;

e) ligating one or more DNA stop fragments to the DNA fragment last incorporated, wherein the DNA stop fragments contain at their 5'-end a sticky end that is complementary to the sticky end at the 3'-end of the previously incorporated DNA fragment and at their 3'-end the other sticky end of the non-palindromic restriction enzyme recognition site to obtain a bound DNA molecule encoding a lipopeptide or peptide synthetase.

Alternatively, such steps are:

a) providing a solid surface having coupled thereto a first oligonucleotide encoding one of the strands of a start DNA sequence, which strand comprises at least the nucleotides complementary to a non-palindromic restriction enzyme recognition site at its 3'-end;

b) annealing to the first oligonucleotide a complementary second oligonucleotide in a manner that generates the sticky end of a non-palindromic restriction

enzyme recognition site at the 3'-end of the double strand thus obtained;

c) ligating a DNA fragment encoding a lipopeptide or peptide synthetase module or domain thereof, which DNA fragment contains at its 5'-end the sticky end of the same non-palindromic restriction enzyme recognition site as the double strand obtained in step b) at its 3'-end ;

d) removing the non-ligated DNA fragments by washing; and

e) optionally repeating steps c) and d) one or more times;

f) ligating one or more DNA fragments encoding DNA stop sequences to the DNA fragment last incorporated, wherein the DNA fragments contain at their 5'-end a sticky end that is complementary to the sticky end at the 3'-end of the previously incorporated DNA fragment and, optionally except for the last incorporated DNA stop sequence, at their 3'-end the other sticky end of the non-palindromic restriction enzyme recognition site, to obtain a bound DNA molecule encoding a lipopeptide or peptide synthetase.

The DNA molecule encoding a lipopeptide or peptide synthetase can then be used as a template in a PCR reaction to amplify the DNA molecule.

Steps c) and d) can be performed only once or be repeated as often as required.

The start DNA sequence can for example be a DNA sequence corresponding to a chromosomal fragment of the host organism, which is useful for integrating the DNA molecule into the host genome.

It is a practical embodiment of the invention when the modules of the library contain the same non-palindromic restriction enzyme recognition sites at their 5'- and 3'-ends, such that the modules can be divided in two groups, library I and II. Members of library I can not ligate to each other, but only to one of the sticky ends of library II modules and members of library II can

not ligate to each other, but only to one of the sticky ends of library I.

In **Fig. 17** this is visualized and this means that the SfiI tags of Lib1R ligates only to Lib2L and Lib2R ligates only to Lib1L. The SfiI tags are designed such that in-frame fusions of the modules are generated and such that they do not encode known proteolytic cleavage sites.

Assembly of the various modules can be performed in solution or on a solid surface. The latter allows control over the randomness of the ligations, because the composition of the module libraries is easily varied per ligation and also the number of different libraries can be varied, and control over the number of modules incorporated. After each library ligation the non-ligated modules can be removed by multiple washing steps. This process can be repeated and the number of ligations (i.e. the number of incorporated modules) can be varied at will.

A particular advantage is if the solid surface is a PCR tube. After the multiple ligations, the assembled DNA molecules (which encode multiple hybrid lipopeptide synthetases in case of a random approach) can then easily be amplified in a final PCR reaction and the PCR tube remains as a source for the hybrid lipopeptide synthetase library. The method is outlined in **Fig. 18.**

In the first step, a 5'-chromosomal fragment that flanks the chromosomal insertion site is ligated to the solid surface. After the first ligation, extensive washings are performed to remove all non-ligated fragments. This is followed by ligation of library I modules (which may contain lipid, PS or PKS modules). Each ligation is followed by washing steps to remove non-ligated modules and by ligation with a library of modules that contain SfiI tags, which fit to the previous modules. The last ligated lipid, PS or PKS module has to contain a thioesterase (TE) domain that is required for

release of the lipopeptide from the lipopeptide synthetase complex.

It is clear that the same method can be used for the assembly of peptide synthetase encoding DNA molecules. These molecules lack the FAS and PKS domains or modules.

In the above way hybrid (lipo)peptide synthetase genes or operons are generated in vitro. The diversity of the generated hybrid (lipo)peptide synthetase DNA molecules can be controlled, since the diversity of the libraries and the number of libraries can be varied, as well as the number of incorporated modules.

To the last module an antibiotic resistance marker may be ligated followed by the ligation of a 3'-chromosomal DNA fragment. The TE domain, antibiotic marker and 3'-chromosomal DNA fragment may be combined into one DNA fragment. The last step comprises a PCR reaction using primers that amplify the entire ligated product, including the 5'- and 3'-chromosomal fragments that provide the homology for the homologous recombination event, once introduced in the host strain.

The chromosomal fragments are preferably chosen such that the hybrid (lipo)peptide synthetase operon is placed under control of appropriate transcription and translation signals, e.g. those of the B.subtilis srf operon. The preferred recipient strain is preferably a natural competent B. subtilis 168 derivative devoid of all lipopeptide synthetase operons. The integration event can be directly selected through the presence of the antibiotic resistance marker in the hybrid DNA molecule. Optional is to use a B. subtilis 168 derivative with an sfp gene that is under control of an inducible promoter. This provides the possibility to identify new (lipo)peptides that have an activity, which is toxic for the host.

Since the methods of the invention lead to novel (lipo)peptide synthetases and novel (lipo)peptides,

such molecules obtainable through the methods are also part of this invention. More in particular, the invention relates to (lipo)peptides having a novel structure and/or novel activity in comparison to known (lipo)peptides. Such lipopeptides are obtainable by the methods as claimed and have a module configuration that does not exist in nature.

The term "module configuration" is intended to encompass information on both the order of the modules and the character of each module. Thus, a novel configuration is also a configuration containing the same order of modules as an existing lipopeptide synthetase, but in which one or more of the modules is modified in its amino acid sequence by for example mutation of its encoding DNA fragment.

It was found that linear DNA molecules are the substrate for natural competent Bacilli. During the uptake by the cells one of the strands is degraded, whereas the other strand will be incorporated into the host's target site by means of homologous recombination. This will result in insertion of lipopeptide synthetase encoding modules in an unpredictable number and order, without the need of cloning vectors, that are frequently employed up till now for effecting transformation. The use of the linear PCR DNA molecules as a direct cloning vehicle without the extra step of incorporation of the DNA molecule in a vector makes the preferred method of the invention very efficient and quicker than the vector-based methods currently used for this purpose and as described in US-5,795,738 and US-5,652,116.

Thus, libraries of Bacillus cells harboring a large diversity of different lipopeptide synthetase encoding genes can be constructed in an efficient and quick way.

The different (lipo)peptides produced by such libraries can be screened for a wide range of antimicrobial activities using relatively simple plate assays or by means of high-throughput screening

technologies. Other activities can be screened by addition of the libraries to microtiter plates containing tumor cell lines, to select for antitumor activities, or by selection for immune modulating activities after combining the libraries with immunity system-related cell lines. Furthermore, antiviral activities can be screened for using well established methodologies.

Additional ways of producing novel lipopeptides according to the invention include the fusion of MycA (the first module of mycosubtilin synthetase comprising the domains responsible for the addition of the fatty acid) to other peptide synthetases, to turn them into lipopeptide synthetases and the use of recombinatorial procedures to modify the fatty acid composition within this particular part of the mycA gene, e.g by using recombination with known polyketide encoding DNA regions.

The invention also relates to the exploitation of the DNA uptake and recombination properties of Bacilli to create large libraries of new lipopeptide synthetase-encoding DNA fragments that will result in the synthesis and secretion of novel lipopeptides, e.g. with altered lipid moieties and/or altered amino acid composition.

In a specific embodiment the invention relates to a method for the non-ribosomal production of modified lipopeptides, which comprises amplification, by means of the polymerase chain reaction and selected pairs of oligonucleotides, and isolation of DNA fragments that encode lipopeptide synthetase modules, including those encoding the fatty acid or polyketide synthase domains, the substrate specificity of which needs not to be known; ligation of mixtures of amplified module encoding DNA fragments in the presence of polyethylene glycol (PEG) and T4 DNA ligase, to induce the formation of linearly fused modules in a random order and number; transformation of natural competent Bacillus species cells that have an inactivated surfactin synthetase and/or mycosubtilin synthetase operon (to be able to select for active new variants by means of an aspecific

screening for lysogenic activity using erythrocytes as indicators); cultivation of said transformed cells under circumstances which lead to expression of the lipopeptide synthetase and production of the corresponding lipopeptide; screening the cells to specifically select for lipopeptides having novel biological activities; and application of the resulting lipopeptides. The mycosubtilin synthetase operon is for example inactivated by means of a point mutation, deletion, or insertion of a marker gene or an inducible lethal gene in the fatty acid synthetase encoding part of the mycA gene.

In a specific embodiment of this method mutation of the module encoding DNA fragments is additionally performed, for example effected by means of site-directed mutagenesis of particular amino acid positions within the modules or domains, which positions have been selected on the basis of molecular models of said modules or domains, that have been constructed, using molecular modelling techniques, on the basis of sequence similarity with the known three-dimensional structure of the phenylalanine activating domain of Gramicidin S (Conti et al. (1997) EMBO J. 16: 4147-4183). This domain is responsible for the amino acid recognition and shows sequence similarity (20-60 %) with other amino acid activation domains. Since the activation domain has been crystalised in the presence of the substrate (phenylalanine), amino acid residues in the domain that are involved in binding of the substrate were identified. These residues will be first targets in an effort to alter substrate specificity of homologous domains.

Screening of the produced lipopeptides can be done on a selective medium, which consists of a standard solid cultivation medium that is covered with a microbial organism (bacterium, fungus or yeast) against which a novel antibiotic is searched, and for which, after growth of the constructed library, effective antibiotic producing Bacilli can be screened for by ability to lyse said microbial organism.

In a specific embodiment the lipopeptide synthetase operon present on the wild type genome of the host bacterium has been exchanged with a hybrid synthetase, which is produced by fusing the wild type synthetase operon to the fatty acid synthetase encoding part of the mycA open reading frame of the mycosubtilin operon, resulting in the coupling of a ß-amino fatty acid to the peptides produced by this operon. In addition, the fatty acid synthetase encoding part may be modified by homologous recombination with, or the addition of, modules of polyketide synthases. Such modules are for example derived from natural or modified polyketide synthetases.

In this application the term "lipid moiety" is used for the fatty acid synthase parts of a lipopeptide synthetase. "Peptide moiety" refers likewise to the peptide synthetase part thereof. Every synthase or synthetase moiety consists of "modules". Each "module" in the peptide moiety of the enzyme can consist of one or more "domains". "Peptide modules" or "domains" are specific for an amino acid, whereas "lipid modules" or "domains" are specific for a lipid in the lipopeptide. Each module or domain is encoded by a "DNA fragment". All DNA fragments that together constitute the coding sequence for the lipopeptide synthetase form the "DNA molecule". When the term "(lipo)peptide" is used it is intended to refer to both lipopeptides and peptides. A "subunit" is a polypeptide that comprises multiple modules.

The present invention (examples 2-7) and the basis therefore (example 1) will be further elucidated in the examples that follow and which are given for illustration purposes only and not to define the scope of the invention.

In the examples reference is made to the following figures:

Figure 1: (a) MALDI-TOF mass spectrum of the lipopeptide fraction of B. subtilis ATCC6633. The first

cluster (Left) corresponds to surfactin, and the second cluster (Right) corresponds to mycosubtilin. (b) MALDI-TOF mass spectrum of the lipopeptide fraction of B. subtilis ATCC6633 containing an integral mutation in mycA, the gene encoding one of the mycosubtilin subunits. Only the cluster corresponding with surfactin is present.

Figure 2: Schematic representation of the entire mycosubtilin operon comprising the ORFs fenF, mycA, mycB, and mycC. The deduced domain organizations of the different proteins specified by the operon are indicated.

Figure 3: Comparison of the highly conserved synthetase motifs of the various domains of the mycosubtilin synthetase subunits. The numbers represent the number of amino acids between the motifs in the mycosubtilin synthetase subunits.

Figure 4: Model for the initiation of iturin lipopeptide synthesis. The reaction intermediates are attached to mycA as carboxy thioesters via 4'-phosphopantetheine cofactors. The carbon chains structure of the fatty acid R is $CH_3(CH_2)_n$, where n is generally 14 (myristate).

Figure 5: Module composition of surfactin and mycosubtilin synthetases.

Figure 6: Linker region between thiolation domain (T) and condensation domain (C) or epimerization domain (E).

Figure 7: Fusion of Myc fAS to first module of srf operon.

Figure 8: Replacement of SrfA1 by Myc FAS+MycA1.

Figure 9: Coupling of DNA modules by overlapping sequences. The open and filled bars represent overlapping sequences.

Figure 10: Genetic organisation of BM04, 24, 28 and 30 in the srf locus and the phenotypic characteristics of these strains.

Figure 11: Hemolytic activities of strains BM04 (12), BM24 (4), BM28 (13), BM30 (8) and BM2823 (9) as visualised on blood agar plate.

Figure 12: Replacement of SrfB2 by MycB3 mediated by homologous recombination through SrfB1 and SrfB3 (indicated by the crosses).

Figure 13: Genetic organisation of BM04, 20, 23 and 25 in the srf locus and the phenotypic characteristics of these strains.

Figure 14: Hemolytic activities of strains BM04, 20 and 23, determined on blood agar plates.

Figure 15: Genetic organisation of BM28, 2820 and 2823 in the srf locus and the phenotypic characteristics of these strains.

Figure 16: Ligation of PCR fragments digested with SfiI that generates non-palindromic sticky ends in order to ligate the DNA fragments in a predetermined order. B can only ligate in one orientation to A and C can only ligate in one orientation to B.

Figure 17; Details of the SfiI tags of library I and II modules. The amino acids that are encoded by these tags after ligation are indicated.

Figure 18: Solid surface assembly of modules resulting in novel hybrid lipopeptide synthetases.

Figure 19: Chemical coupling of oligo's to NucleoLink™ PCR tubes and the subsequent ligation of SfiI tagged modules.

EXAMPLES

EXAMPLE 1

The mycosubtilin synthetase of Bacillus subtilis is a multifunctional hybrid between a peptide synthetase, an amino transferase and a fatty acid synthetase

1. Introduction

In this example the molecular analysis of the mycosubtilin synthetase operon is described. From this it will follow that the mycosubtilin synthetase comprises both a peptide moiety, which is responsible for the

synthesis of the peptidic part of the corresponding peptide, the antibiotic mycosubtilin, and a lipid moiety containing fatty acid synthetase module.

2. Material and methods

2.1 General methods and materials

Molecular cloning and PCR procedures were carried out using standard techniques (Birnboim, H. C. & Doly, J. (1979) Nucleic. Acids Res. **7:** 1513-1523; Ochman, H., Medhora, M. M., Garza, D. & Hart D. L. (1990) PCR protocols (Academic Press, San Diego, Calif.); and Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) Molecular cloning: a laboratory manual (2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.)).

Media for growth of B. subtilis and E. coli have been previously described (Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) Molecular cloning: a laboratory manual (2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.; Spizizen, J. (1958) Proc. Natl. Acad. Sci. **44:** 1072-1078; and Walton, R. P., and H. B. Woodruff (1949) J. Clin. Invest. **28:** 924-926.23-25). B. subtilis chromosomal DNA was purified as previously described (Venema, G., Pritchard, R. H. and Venema-Schröder, T. (1965) J. Bacteriol. **89:** 1250-1255.).

2.2 Isolation of lipopeptides

Isolation of lipopeptides produced by B. subtilis ATCC6633 was essentially performed as described by Ebata et al. (J. Antibiot. **22:** 467-472). Cells were grown in ammonium citrate/sucrose (ACS) medium at 37°C (Walton, R. P., and H. B. Woodruff (1949) J. Clin. Invest. **28:** 924-926). After 4 days of cultivation, the supernatant was collected by centrifugation and adjusted to pH 2.0 using concentrated HCl, and stirred overnight. The precipitate was collected by centrifugation and extracted twice, first with 95 % ethanol, and subsequently with 70 % ethanol.

The concentrated ethanolic extracts were mixed with 2 volumes of water and, after centrifugation, the precipitate was recrystallized 3 times and fractionated using LH-20 gel permeation chromatography (column size 2x25 cm) and eluent system C: chloroform-methanol-ethanol-water 70:30:35:15 (v/v/v/v).

2.3 Mass spectrometry of the lipopeptide products

Lipopeptides were analyzed using fast atom bombardment (FAB) mass spectrometry and matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry. For FAB mass spectrometry, a vacuum generator ZAB-3HF spectrometer (BEB configuration) was used. Samples were dissolved in dimethylsulfoxide/glycerol and positive ions were detected. MALDI-TOF mass spectra were recorded on a Bruker Reflex MALDI-TOF instrument with a 337 nm nitrogen laser for desorption and ionization. A saturated solution of α-cyano-4-hydroxycinnamic acid in 70 % acetonitrile/0.1 % trifluoroacetic acid (v/v) was used as matrix. Ions were accelerated with a voltage of 20 kV. The positive-ion and reflector mode was applied.

2.4 Thin layer chromatography

For the analysis of the amino acid composition, the lipopeptide fractions were separated by thin layer chromatography using silica gel 60-plates (Merck). Three eluent systems (A, B, C) were used: system A, consisting of n-butanol-acetic acid-water 4:1:1 (v/v/v), system B, consisting of chloroform-methanol-water 65:25:4 (v/v/v) and system C, consisting of chloroform-methanol-ethanol-water 70:30:35:15 (v/v/v/v). The various spots were visualized with ninhydrin, using the 4,4'-tetramethyl-diamino-diphenylmethane (TDM) assay and by charring after spraying with concentrated H_2SO_4 (28).

2.5 Amino acid analysis

For preparative isolation of the lipopeptide fractions, the corresponding spots were scratched out from the thin layer chromatograms, and the silica gel material was extracted with methanol. The purified lipopeptides were hydrolyzed using 6 N HCl in the presence of 0.2 % phenol and 0.1 % thioglycolic acid in sealed, evacuated tubes. Amino acid analysis of the hydrolysates was performed using an automated amino acid analyzer (Durrum D-500). Determination of the absolute configuration of the amino acid components was performed by derivatising the amino acids with 1-fluoro-2,4-dinitro-5-L-alanine amide (Marfey's reagent) and separation of the obtained diastereomers with reversed phase HPLC (LKB) using Hypersil ODS RP18 (Knauer; column size 250x6mm) (Marfey, P. (1984) Carlsberg Res. Commun. **49:** 591-596).

2.6 Incorporation of radioactive precursor amino acids

Radioactive precursor amino acids (1 μCi, Amersham) were added to growing cultures 24 hours after incubation. The cells were harvested 3-4 days after inoculation. HCl-precipitates of the culture supernatants were extracted with 5 ml of 95 %, and 70 % ethanol, respectively. Separation and quantification of the labeled lipopeptides were performed by TLC (eluent system A-C) and radioscanning using a Berthold Analyzer LB2832.

2.7 Identification of a new peptide synthetase operon

A genomic library of B. subtilis ATCC6633 was constructed by partial digestion of chromosomal DNA with Sau3A and ligation into the BamHI site of the vector λ-DashII (Stratagene). The DNA was packaged in vitro using GigapackII (Stratagene), according to the manufacturer's protocols, and the resulting phages were used to infect E. coli P2392 (Stratagene). Plaques were transferred to nylon membranes (Qiagen).

A mixture of oligonucleotide probes, directed against DNA regions encoding the highly conserved sequence motifs KAGGAYVP and GTTGKPKG of peptide synthetases, was used to screen the genomic DNA library, as has been previously described for screening in other microorganisms (Bernhard, F., Demel, G., Soltani, K., Döhren, H. V., & Blinov, V. (1996) DNA Seq. **6:** 319-330.). Positive λ-DashII clones were restricted with EcoRI, blotted on a nylon membrane, and hybridized at medium stringency with the oligonucleotide probes. This Southern hybridization was used to discriminate insert size and possible surfactin operon fragments. A selected λ clone was restricted with EcoRI, subcloned into pBluescript KS(+) (Stratagene) and its sequence was determined.

2.8 Sequence determination of the synthetase operon

To obtain the complete sequence of the peptide synthetase-encoding operon, inverse PCR was applied, using the Expand PCR system (Boehringer) (Ochman, H., Medhora, M. M., Garza, D. & Hart D. L. (1990) PCR protocols (Academic Press, San Diego, Calif.)). The PCR fragments were fragmented by a DNaseI treatment, subcloned into pUC18 and sequenced (Sambrook et al., supra, Yanisch-Perron, C., Vielra, J. & Messing, J. (1985) Gene **33:** 103-119).

2.9 Purification of the second amino acid-activating module

A His_6-tag fusion with the second module of the mycosubtilin synthetase was obtained by PCR cloning using primers mycB1up; CGC GGA TCC ATG TCG GTG TTT AAA AAT CAA GTA ACG, and mycB1down; GGC GTC GAC TTA GGA CGC CAG CAG TTC TTC TAT TGA G. These primers contained restriction sites for BamHI and SalI, respectively (underlined), which were used to clone this module C-terminal to a His_6-tag in the expression plasmid pQE30 (Qiagen). The resulting plasmid was transformed into E. coli DH5α

containing plasmid pREPGroESL (Cole, P. A. (1996) Structure **4:** 235-242).

For overexpression, an overnight culture, grown in LB medium at 28° C, was diluted 100 fold, and at A_{590} = 0.5, the culture was induced with 0.5 mM (final concentration) isopropyl-ß-D-thiogalactopyranoside (IPTG). The cells were harvested after three hours and disrupted using a French press. The expressed second module was purified using Ni^+-affinity chromatography. The protein was eluted from the Ni^+-column with 100 mM imidazole in a 20 mM Tris-buffer (pH 6.5). Purification was followed by SDS-PAGE.

2.10 ATP-PP_i exchange assay

To determine the substrate specificity of the purified second module of the mycosubtilin synthetase, an ATP-PP_i exchange assay was performed in the presence of different amino acids, as previously described (Turgay, K., Krause, M. & Marahiel, M. A. (1992) Mol. Microbiol. **6:** 529-546).

2.11 Mutational analysis

To obtain a stable integrational mutant of the mycosubtilin synthetase, a two-plasmid system, developed for gene disruption in Lactococcus lactis, was used (Leenhouts, K., Buist, G., Bolhuis, A., ter Berge, A., Kiel, J., Mierau, I., Dabrowska, M., Venema, G. & Kok, J. (1996) Mol. Gen. Genet. **253:** 217-224; Law, J., Buist, G., Haandrikman, A., Kok, J., Venema, G., & Leenhouts, K. (1995) J. Bacteriol. 177: 7011-7018). This system uses a temperature sensitive replication-initiation protein (RepA), delivered in trans on plasmid pVE6007 (Leenhouts et al., supra). In the SmaI restriction site of the plasmid pORI28 (Em^R), lacking RepA, an internal fragment of the second ORF of the peptide synthetase operon, was ligated.

The PCR primers used to isolate this fragment were: myc33: CCG GAA TTC GAC CAC TTT CTG TCT CTG G and

myc32; AAT TTG ATG ATA TCT ATT CTC. Both plasmids were cotransformed into B. subtilis ATCC6633 using protoplast transformation at the permissive temperature of 30°C (Chang, S. & Cohen, S. N. (1979) Mol. Gen. Genet. **168:** 111-115). Site-specific integration was achieved by selection for erythromycin resistant colonies at the restrictive temperature of 37°C. Integration was verified by PCR control.

3. Results

3.1 Lipopeptides produced by B. subtilis ATCC6633

The lipopeptide fraction obtained from the culture supernatant of B. subtilis ATCC6633 was analyzed by FAB mass spectrometry -and MALDI-TOF mass spectrometry. With FAB mass spectrometry two clusters of ions were detected; (1) at m/z = 1008.5, 1022.5 and 1036.5, and (2) at m/z = 1071, 1085, 1093 and 1107. By MALDI-TOF mass spectrometry measurements two series of ions were found; (1) at m/z = 1045.7, 1061.1 and 1075.4, and (2) at m/z = 1109.9 and 1123.9 (**Fig. 1a**). These ions indicate isoforms of surfactin (1) and mycosubtilin (2), respectively. The FAB mass spectrometry data are consistent with $[M + H]^+$-ions, while the m/z-values determined by MALDI-TOF mass spectrometry represent alkali adducts of these species.

The mass spectrometric results were corroborated by biochemical analysis. Table 1 presents the data obtained from the analysis of the amino acid composition of both lipopeptide fractions using a Durrum Analyzer and reversed phase HPLC, as described in Material and Methods.

Table 1

	Surfactin	Mycosubtilin
TLC	R_fA = 0.90	R_fA = 0.23
	R_fB = 0.65	R_fB = 0.27
	R_fC = 0.80	R_fC = 0.51

Durrum analyzer	Asx(1.0)Glx(1.0)	Asx(3.2)Glx(1.0)
	Val(1.1)	Ser(0.8)Pro(1.0)
	Leu(3.2)	Tyr(0.8)
Reverse-phase HPLC	Asx(1L)Glx(1L)	Asx(1D,2L)Glx(1L)
	Val(1L)	Ser(0.8)Pro(1.0)
HPLC	Leu(2L,2D)	Tyr(1D)

The amino acid composition of the separated lipopeptide fractions corresponds to that of surfactin and mycosubtilin. In addition, when B. subtilis ATCC6633 was grown in medium containing ^{14}C-labeled amino acids, ^{14}C-labeled L-Asp, L-Glu, L-Leu and L-Val were incorporated in the surfactin fraction, while ^{14}C-labeled L-Asp, L-Glu, L-Ser, L-Pro and L-Tyr were incorporated in the mycosubtilin fraction. These results substantiate that, under the growth conditions used, B. subtilis ATCC6633 produces two lipopeptides, surfactin and mycosubtilin.

3.2 Identification and sequence determination of a new peptide synthetase

The structural similarity of mycosubtilin with (lipo)peptide antibiotics suggests that mycosubtilin is synthesized by a member of the peptide synthetase family. Southern hybridization analyses, using oligonucleotide probes directed against the coding regions of the highly conserved KAGGAYVP and GTTGKPKG amino acid motifs of peptide synthetases, was used to search in a genomic DNA library of B. subtilis ATCC6633 for the presence of new peptide synthetases. The inserts of positive λ clones were analyzed, and a 22 kb insert, with a restriction profile different from that of the known surfactin operon, was selected. Alignments of the nucleotide sequence of the insert with other peptide synthetase genes revealed a typical modular structure, and indicated that the isolated DNA fragment encoded part of an unknown peptide synthetase. Consecutive steps of inverse PCR were

then used to complete the sequence of the peptide synthetase operon.

3.3 Specific adenylation of tyrosine by the second module

Analysis of the peptide synthetase indicated the presence of seven consecutive modules. According to the colinearity rule, the second module of the mycosubtilin synthetase would be responsible for the incorporation of tyrosine. To determine whether or not this assumption was correct, it was attempted to overexpress and purify the tyrosine-activating module. For this purpose a N-terminal $(His)_6$-tag fusion with the first 775 amino acid residues of MycB, comprising the second amino acid-activating module, was constructed. Overexpression of the 86 kDa fusion protein in E. coli resulted in the formation of inclusion bodies. Only by coexpression of the E. coli chaperons GroEL/ES, soluble fusion protein was obtained. The amino acid specificity of the purified module was determined in an ATP/PPi exchange assay. Of the 20 amino acids tested, only adenylation of L-tyrosine was detected. Since this module contains an epimerization domain, it was concluded that the second module is responsible for the incorporation of D-tyrosine, which points to a mycosubtilin producing peptide synthetase.

3.4 Insertional mutagenesis results in a mycosubtilin-negative phenotype

To unambiguously prove that the mycosubtilin operon was cloned, the operon was genetically disrupted and the effect on mycosubtilin production was examined. As the competence development of B. subtilis ATCC6633 was insufficient to follow the protocol for natural transformation, protoplast transformation using a two plasmid system initially developed for gene disruption in Lactococcus lactis (Leenhouts et al., supra; Law et al., supra) was used. A 2 kb internal fragment of the second

ORF of the operon was cloned into the integration plasmid pORI28 (Em^R), which lacks repA, encoding the replication initiation protein. Protoplasts of B. subtilis ATCC6633 were transformed subsequently with plasmid pVE6007 (Cm^R), containing the gene for a temperature sensitive RepA protein, and pORI28, containing 2 kb of ORF2. At the permissive temperature (30°C) RepA is stable, and both plasmids replicate. However, at the restrictive temperature (37°C) RepA is inactive, and pORI28 can only be stably maintained when integrated in the genome of the cell.

As integration occurs preferentially via homologous recombination, selection for erythromycin resistant colonies at 37°C results in B. subtilis ATCC6633 mutants with pORI28 inserted into the second ORF of the peptide synthetase. PCR control using primers flanking the expected integration site of pORI28 indicated that the plasmid had integrated at the expected locus (data not shown). Analysis using MALDI-TOF mass spectrometry revealed that such mutants failed to produce mycosubtilin, since only mass peaks of surfactin were visible (**Fig. 1b**). This shows that the mycosubtilin synthetase was identified.

3.5 Analysis of the mycosubtilin synthetase gene cluster

Sequence analysis revealed the presence of a cluster of four open reading frames (ORF), designated fenF, mycA, mycB, and mycC, respectively (**Fig. 2**). The cluster spans about 38 kb, and is flanked by two putative transcriptional terminators located 685 base pairs upstream of the first ORF (fenF), and 71 base pairs downstream of the last ORF (mycC). The intergenic regions between the four ORFs, each containing a putative ribosomal binding site, varied between 35 and 100 bp.

The first ORF of the mycosubtilin operon, fenF, encodes a protein of 45.2 kDa, with 50 % similarity to malonyl-CoA transacylases. The protein is 94 % identical to FenF, encoded by a gene located upstream of the

fengycin synthetase operon of B. subtilis F29-3 (Chen, C. L., Chang, L. K., Chang, Y. S., Liu, S. T. & Tshen, J. S. (1995) Mol. Gen. Genet. **248:** 121-125). MycA, mycB and mycC encode proteins with estimated molecular masses of 449.3 kDa, 612.3 kDa and 297.9 kDa, respectively. They show strong similarity with members of the peptide synthetase family and display the ordered assembly of conserved condensation, adenylation and thiolation domains characteristic for such multienzymes. The most conserved amino acid motifs, characteristic for the different domains commonly found in peptide synthetases are indicated in **Fig. 3**. As shown in **Fig. 2**, a total of seven amino acid-activating modules can be distinguished; one in MycA, four in MycB and two in MycC. Modules two, three and six contain an epimerization domain, indicating that the activated amino acids are converted into the D-configuration. The number of modules and the location of epimerization domains correspond to the number of amino acids and the position of D-amino acids in the peptide moiety of mycosubtilin. The last domain of this multienzyme system is a thioesterase domain, which is presumably required for release and possibly for cyclization of the synthesized lipopeptide molecule.

MycA exhibits a remarkable complexity. The C-terminally located amino acid-module is preceded by several domains with homology to proteins involved in the synthesis of fatty acids and polyketides. The conserved motifs, characteristic for these domains, were also identified in these domains.

As indicated in **Fig. 2**, four different domains could be distinguished. The first domain (AL) shows 49 % similarity to long-chain fatty acid CoA-ligases and contains a putative ATP binding box. The second and fourth domains (ACP) show 45 % similarity to acyl carrier proteins and both contain a putative 4'-phosphopantetheine binding box. The third domain (KS) is 53 % similar to β-ketoacyl synthetases. These fatty acid synthase domains are followed by a domain (AMT) with 54 %

similarity to glutamate-1-semialdehyde aminotransferases, which contains a typical pyridoxal phosphate binding box. The fatty acid synthase and amino transferase domains are connected to the first module of this peptide synthetase via an extra condensation and thiolation domain.

4. Conclusion

B. subtilis strain ATCC6633 produces a potent anti-fungal lipopeptide, mycosubtilin. In this the identification of a peptide synthetase operon specifying the synthesis of this antibiotic was described. The number of modules of the synthetase, and the position of epimerization domains correspond to the amino acid sequence of mycosubtilin. The predicted amino acid specificity of the second module was confirmed in vitro. Genetic disruption of the operon abolished mycosubtilin production.

The mycosubtilin operon containing 38 kb of DNA consists of four ORFs; fenF, mycA, mycB and mycC. Although the overall structure of the synthetase resembles that of other peptide synthetases, mycosubtilin synthetase exhibits a number of important deviations. Most strikingly is the large N-terminal multifunctional part of MycA, with domains that show homology to fatty acid and polyketide synthetases. These domains play a role in the incorporation of the β-amino fatty acid moiety into the mycosubtilin molecule. Based on the function of similar domains in fatty acid-, polyketide- and peptide synthetases as well as in amino transferases, the synthesis of mycosubtilin and other iturin lipopeptides is according to the model, schematically presented in **Fig. 4**. In the first step (**Fig. 4**, step I), the acyl CoA-ligase domain couples coenzyme A (CoA) to a long-chain fatty acid, for example myristate, in an ATP dependent reaction. The activated fatty acid is then transferred to the 4-phosphopantetheine cofactor of the first acyl carrier domain. In addition, a malonyl-CoA is attached to the 4-phosphopantetheine cofactor of the

second acyl carrier domain, a reaction catalyzed by the malonyl-CoA transacylase, encoded by fenF. In these early steps of synthesis asparagine is activated and coupled to the 4-phosphopantetheine cofactor of the first module. The second step is the condensation of the malonyl and acyl thioesters (**Fig. 4**, step II), catalyzed by the ß-ketoacyl synthetase domain, resulting in a ß-ketoacyl thioester. In contrast to normal fatty acid synthesis, this ß-ketoacyl thioester is not converted into a fatty acid by subsequent reduction and dehydration reactions, but in a ß-amino fatty acid, by a transamination reaction, catalyzed by the domain homologous to amino transferases (**Fig. 4**, step III). Subsequently, the ß-amino fatty acid is transferred to a thiolation domain, catalyzed by the preceding condensation domain. The ß-amino fatty acid is then coupled to the activated asparagine thioester, catalyzed by the condensation domain preceding the first module of the peptide synthetase (**Fig. 4**, step IV). In subsequent condensation reactions the mycosubtilin molecule is then synthesized, analogous to other non-ribosomal peptide synthetase reactions (**Fig. 4**, step V).

It was thus shown that the mycosubtilin operon encodes a hybrid synthetase, integrating fatty acid synthesis and peptide synthesis. The unique combination of enzymatic domains in MycA is the basis of the present invention in which several possibilities are provided to engineer lipopeptides. Modifications of the peptide moiety of lipopeptides has been described before. However, the research that led to the present invention, namely that the fatty acid synthetase is an integral part of the lipopeptide synthetase enzyme, now also offers the potential to modify the fatty acid moiety of lipopeptides in general and mycosubtilin in particular. In addition, MycA might be fused to other peptide synthetases, to turn them into lipopeptide synthetases.

EXAMPLE 2

Developing novel lipopeptide synthetases

1. Introduction

Novel lipopeptide synthetases may be generated by coupling of lipid modules to existing peptide modules, by replacing existing lipid modules by others, by replacing/ deleting/adding peptide synthetase (PS) modules or by a combination of altering lipid and PS modules.

The successful fusion or exchange of modules requires:

(i) Definition of DNA fragments, indicated as modules, which enable the activation of a compound (e.g. an amino acid or a lipid chain) and the coupling of this activated compound to the activated compound of the next module. For the PS modules it was established that a CAT configuration, meaning that a module starts with a condensation domain, followed by an adenylation and thiolation domain is the configuration that constitutes the functional module in vivo. The module composition of the surfactin and mycosubtilin operons is shown in **Fig. 5**. The CAT configuration is in agreement with the in vitro experiments of Belshaw et al. (1999. Science 284: 486-489). In this application the MycA lipid module (FAS module) is defined as consisting of the FenF gene plus AL, ACP, KS, ACP and AMT domains of MycA plus the condensation and thiolation domain that directly follow the AMT domain (see **Figs. 2 and 5**).

(ii) The identification of regions (linker regions) that allow fusion of modules such that functional lipopeptide synthetases are created. In the surfactin and mycosubtilin modules linker regions were defined between the thiolation domains and the condensation domains (**Fig. 6**). In the DNA region that encodes the indicated linker region of 40 amino acids in **Fig. 6**, restriction enzyme recognition sites can be inserted or can replace part of the sequence. The presence of artificial restriction enzyme sites enables

the in vitro coupling of DNA fragments that encode modules. The in vitro coupled modules yield functional lipopeptide synthetases after integration into the chromosome of B. subtilis by homologous recombination. Similar linker regions are present between PS modules of other (lipo)peptide operons and these linker regions provide a means to couple modules in a controlled or random way. This is supported by the recent experiments of Mootz et al. (2000. Proc. Natl. Acad. Sci USA 97: 5848-5853), who used similar linker regions between tyrocidine PS modules of Bacillus brevis to insert a restriction enzyme site and fused in this way a number of modules which were successfully expressed in E. coli. In in vitro experiments it was shown that the hybrid enzymes were functional.

A. Addition of lipid modules

A method of developing novel lipopeptide synthetases consists of coupling lipid modules to existing peptide modules. For example the addition of the MycA fatty acid synthase (FAS) module by in frame fusion to the gramicidin S operon of Bacillus brevis, thereby adding a fatty acid chain to the normally non-lipomodified non-ribosomally synthesized gramicidin S. Yet another example is the in frame fusion of the MycA FAS module to the existing PS modules of the surfactin operon in a B. subtilis strain producing surfactin. This fusion changes it from a surfactin with a β-hydroxy fatty acid into one with a β-amino fatty acid. Two types of this particular fusion were obtained:

(i) one in which Myc FAS is fused in frame to the first PS module (SrfA1) of the srf operon (**Fig. 7**);

(ii) one in which Myc FAS including the first PS module of the myc operon, MycA1, is fused in frame to the second PS module (SrfA2) of the srf operon (**Fig. 8**).

These fusions are described below.

Insertion of the Myc FAS module by means of integration of coupled PCR fragments

The preferred method to insert the Myc FAS module is by integration of linear PCR DNA of the Myc FAS module that is coupled by PCR at the 5'- and 3'-ends to chromosomal DNA that flanks the chromosomal insertion site. This coupled linear DNA molecule is then taken up by a B. subtilis strain that shows natural competence (van Sinderen D et al. (1994) Mol. Microbiol. 11: 695-703), which subsequently integrates it by homologous recombination. The preferred recipient strain is a B. subtilis 168 derivative with an active sfp gene (encoding a cofactor that is required for active FAS, PS and PKS modules), which thus produces surfactin.

In order to detect the insertion of the FAS or FAS + MycA1 module, first the chromosomal insertion site was genetically labeled by the insertion of marker genes, an antibiotic resistance gene (a kanamycin resistance gene [Km]) and the E. coli β-galactosidase gene (lacZ). This event inactivated the surfactin operon, since it uncouples the srf promoter and ribosome binding site from the srf operon. The intended fusion site is the ATG start codon of the surfactin operon. The primers used for the amplification of the DNA fragments are listed in Table 2.

The templates used in the PCR reactions were plasmids pDG792 (Guerout-Fleury AM et al. (1995) Gene 167: 335-6), pMutin2 (Vagner V, Dervyn E & Ehrlich SD (1998) Microbiology 144: 3097-3104), chromosomal DNA of B. subtilis 168 and ATCC6633 for the Km^r gene, lacZ, Srf5' (containing the srf promoter), SrfA1 and the Myc FAS or Myc FAS + MycA1 module, respectively. Alternatively, the PCR fragments were obtained by using plasmid DNA of the library as described in example 4 as template. High fidelity PCR enzymes were used for amplification of all the fragments (e.g. Pfu of Stratagene, Expand or Expand High Fidelity of Roche). The DNA fragments were amplified under standard PCR conditions according to the guidelines provided by the suppliers of the enzymes with a maximum

of 20 cycles. The amplified DNA fragments were generated such that overlapping sequences of 15-30 bp were generated with fragments to which a fragments had to be fused (**Fig. 9**).

Prior to the fusions all PCR fragments were purified using the High Pure PCR Product Purification Kit (Roche) followed by a polyethylene glycol (PEG) precipitation (PEG8000, end concentration 10% + 10 mM $MgCl_2$) to remove all residual primers. The DNA was resuspended in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8).

The fusion of the fragments was done in PCR reactions. In this method, two or more DNA fragments are coupled in a modified PCR procedure. These overlapping sequences typically have a melting temperature (Tm) between 50°C and 65°C. They serve as primers in the DNA synthesis. The coupling between the fragments is effected by the DNA polymerase that is used (e.g. Herculase of Stratagene, rTth of Perkin Elmer, Expand Long Template PCR system or Expand 20kbPlus PCR system of Roche). The fragments to be coupled (about 1 μg of each) are mixed together with an appropriate amount of PCR buffer, deoxyribonucleotides (dNTP's), and DNA polymerase in a PCR tube, and this mix is incubated in a thermocycler. In a typical coupling experiment, the program will look as follows:
Preliminary incubation at 94°C for 2 minutes (initial denaturation) followed by 10 cycles with (1, denaturation) 30 sec. at 94°C, (2, annealing) 30 sec. at 50°C and (3, extension) 30 sec. to 1 min. per kilobasepair at 68°C. The exact nature of the program is dependent on the length of the fragments to be coupled and the Tm of the overlapping sequences.

First, the coupled DNA of Srf5'-Km-LacZ-SrfA1 was introduced into the naturally competent host strain B. subtilis BM04, a 168 derivative that produces surfactin. The transformants were selected on standard transformation medium containing the selective antibiotic

and 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (Xgal). The latter is a chromogenic substrate for LacZ, which converts it into a dye that stains bacterial colonies blue.

A Km resistant, blue transformant that contained the Km-LacZ cassette at the correct position was labeled BM24 (**Fig. 10**). Blood agar plates were used to determine whether the bacterial colonies produced hemolytic activities (Mulligan CN et al. (1984) J. Ferment. Technol. 62: 311-314). BM24 did not show hemolytic activity on such plates (**Fig. 11**), whereas strain BM04 does, indicating that the surfactin operon is inactive in BM24.

Subsequently, the two different coupled fragments containing the Myc FAS module, Srf5'-FAS-SrfA1 and Srf5'-FAS+MycA1-SrfA2, were used to transform strain BM24. To select a transformant in which the Km-LacZ cassette had been replaced, use was made of co-transformation. For this purpose a PCR fragment encoding a chloramphenicol resistance marker (Cm) was fused at its 5'- and 3'-end, to B. subtilis 168 chromosomal PCR fragments of the yhjM gene region using the PCR coupling technique described above (for primer sequences see Table 2).

This Cm coupled fragment was used in the two types of transformations together with each of the FAS coupled fragments. Initially, chloramphenicol resistant colonies were selected on agar plates that also contained Xgal. Then, white colonies were screened for the loss of the Km marker and, subsequently, for the ability to produce a hemolytic activity on blood agar plates (**Fig. 11**).

Using this procedure strain BM28 was identified that contains an in frame fusion of the Myc FAS module with the first module of the srf operon (SrfA1) (**Fig. 10**), which was confirmed by sequencing of the chromosomal DNA of strain BM28 in this particular region. The hemolytic activity produced by strain BM28 differs from that produced by strain BM24. BM28 shows a higher

hemolytic activity, which may be attributed by the alteration of the type of fatty acid bond (conversion from β-hydroxy to β-amino bond).

In the same way strain BM30 was identified that contains an in frame fusion of the Myc FAS + MycA1 module with the second module of the srf operon (SrfA2) (**Fig. 10**). Also BM30 had a hemolytic activity that differs from BM24 (**Fig. 11**).

These examples show that coupled PCR fragments can be used directly in transformations of B. subtilis 168 derivatives to make new combinations of lipid and PS modules in vivo, and that these new combinations result in active lipopeptide synthetases that produce new lipopeptides.

B. Replacement of PS modules

Similar to the insertion of the Myc FAS module by integration of coupled PCR fragments, PS modules can be exchanged. To test this approach the SrfB2 module was exchanged by the MycB3 module (**Fig. 12**). In a first step the SrfB2 module was replaced by the Km-LacZ cassette. Primers required to obtain the Km and LacZ DNA fragment are listed in Table 2. The primers to amplify the SrfB1, MycB3 and SrfB3 modules are listed in Table 3. The SrfB1-Km-LacZ-SrfB3 and SrfB1-MycB3-SrfB3 fragments were coupled in PCR reactions as described above.

Strain BM04 was transformed with the SrfB1-Km-LacZ-SrfB3 fragment and this resulted in strain BM20, which is Km resistant, blue on agar plates with Xgal and has no hemolytic activity on blood agar plates (**Figs. 13 and 14**). Strain BM20 was then used as a recipient for the SrfB1-MycB3-SrfB3 DNA molecules. Selection (by co-transformation) and screening was as described before. This procedure yielded strain BM23 that contains an in frame fusion of SrfB1-MycB3-SrfB3 (**Fig. 13**), which was confirmed by sequencing of the chromosomal DNA in this gene region. Strain BM23 showed less hemolytic activity on blood agar plates then BM04 (**Fig. 14**).

The exchange of the SrfB2 module was also effectuated in strain BM28, similar as was done for BM04. This yielded first strain BM2820 in which the Km-LacZ cassette replaced the SrfB2 module, thereby inactivating the lipopeptide synthetase operon (**Fig. 15**). In the next step, strain BM2823 was obtained. The incorporation of the MycB3 module at the SrfB2 resulted again in the production of a hemolytic activity (**Figs. 11 and 15**).

The above described examples also show that coupled PCR fragments can be used directly in transformations of B. subtilis 168 derivatives to make new combinations of lipid and PS modules in vivo, and that these new combinations result in active lipopeptide synthetases that produce new lipopeptides.

Insertion of the Myc FAS and MycB3 modules by means of replacement vectors

An alternative approach to construct the above-described strains is to clone the PCR fragments first in integration vectors, e.g. a pORI type vector (e.g. pORI28; Leenhouts K et al. (1998) Methods in Cell Science 20: 35-50). Such a vector is not capable of being autonomously replicated in the host organism but can be used to insert DNA fragments at a specific chromosomal location via homologous recombination. DNA primers should be used in this case that generate restriction enzyme recognition sites at the 5'- and 3'-ends of the PCR fragment to allow insertion of said fragment in the replacement vector. After construction and isolation of the integration vector, linearized vector DNA can be introduced into B. subtilis by natural competence. Selection of the strains is as described above.

In the examples described thus far use is made of a B. subtilis 168 derivative as the host organism to make new combinations of FAS and PS modules. This bacterial strain shows natural competence, if the recipient strain does not possess natural competence use is made of replacement vectors of the pORI280-type

(Leenhouts, K. et al. (1996) Mol. Gen. Genet. 253: 217-224.). The PCR fragments are inserted into this vector similar as described above. The replacement vectors are introduced into the recipient strain using a standard protoplast transformation protocol (e.g. Chang S & Cohen SN. (1979) Mol. Gen. Genet. 168: 111-115, or any other known protocol suitable for the organism used). The desired strains are selected in a two-step procedure to integrate the desired fragment into the selected chromosomal position (Leenhouts K et al. (1998) Methods in Cell Science 20: 35-50). Screening of transformants as above.

EXAMPLE 3

Development of novel lipopeptides by replacement of the MycA lipid module

A method of developing novel lipopeptide synthetases consists of replacing the MycA FAS module of the mycosubtilin operon with other FAS modules and/or polyketide synthase modules (PKS). The other FAS and PKS modules are selected on the basis of homologies and may originate from Bacilli or other organisms.

Examples from B. subtilis are homologues to long chain fatty acid synthase, such as yhfL, lcfA and yngI. Examples from B. subtilis are homologous to polyketide synthases, such as pksJ, pksP and pksK. Examples of Streptomyces lividans are the eryA PKS genes (Xue Q et al. (1999) PNAS 96: 11740-11745).

All FAS modules are obtained by PCR and introduced in replacement vectors or used in direct transformations similar as in example 2. The replacement vector in this case containing as flanking sequences that provide the homology for the replacement integration event, fragments of the mycosubtilin operon and or surfactin operon such that an in frame fusion will occur after the integration event. Selection and screening of the transformants as in example 2.

EXAMPLE 4

Construction of fatty acid and polyketide synthase and peptide synthetase libraries

1. Introduction

In order to be able to develop novel lipopeptide synthetases and their resulting lipopeptides a range of DNA fragments encoding FAS, PS and PKS modules is to be provided. This can best be done in a library. This example describes the construction of such a library.

2. Selection and isolation of PS, FAS and PKS modules

On the basis of available sequence information in the databases FAS, PS or PKS modules can be obtained from various organisms by PCR amplification using high fidelity polymerases (e.g.: Pfu and Herculase of Stratagene; Pwo, Expand High Fidelity and Expand Long Template PCR System of Roche or rTth of Perkin Elmer). The primers, which are specific for the individual modules, may include at their 5'-ends restriction enzyme recognition sites or overlapping sequences with other modules in order to enable in vitro coupling of modules (see examples 2 and 6). The PS module and MycA FAS boundaries were defined in example 2 (**Fig. 5**). The boundaries of PKS modules were defined by Gokhale et al. (1999. Science 284: 482-485).

A small library consisting of SrfA1, A2, SrfB1, B2, B3 and MycB3 PS modules and MycA FAS and MycA FAS + MycA1 modules was made using the primers listed in Table 4.

PCR reactions were performed with high fidelity enzymes as described before under standard conditions as recommended by the supplier of the enzymes (with a maximum of 20 cycles). The PCR fragments were purified by using the High Pure PCR Product Purification Kit of Roche followed by a polyethylene glycol (PEG) precipitation (PEG8000, end concentration 10% + 10 mM $MgCl_2$). The DNA

was resuspended in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8). Fragments containing specific restriction enzyme recognition sites at their 5'-ends were digested under standard conditions, followed by purification as described above and the fragments were stored in TE buffer at -20°C.

It is optional to clone the PCR fragments first into commercially available E. coli cloning vectors (e.g. pUC18, pCR4Blunt-TOPO), transform E. coli with the ligation mixtures, and determine the nucleotide sequences prior to their use.

The library thus obtained provides the substrate for creating novel combinations of modules by means of the methodology described below (example 6). In addition, PCR/cloning of additional PS modules and FAS and PKS modules from Bacilli or other organisms can increase the diversity of the library as described in examples 3 and 5.

EXAMPLE 5

Developing new lipopeptide synthetases by means of mutation

1. Introduction

Another possibility for developing novel lipopeptide synthetases and their corresponding lipopeptides is mutation of one or more of the DNA fragments encoding the modules. This example describes how such mutation can be effected.

2. Materials and methods

2.1 General methods and materials

Molecular cloning and PCR procedures were carried out using standard techniques (see Example 1).

2.2 Mutagenesis procedures

Site-directed mutations in the modules were introduced using the QuickChangeTM Site-Directed Mutagenesis Kit (Stratagene). For the introduction of

random mutations error-prone PCR was used. Therefore, PCR reactions are carried out using Taq polymerase (Boehringer) in the presence of a reduced concentration of dATP (dGTP:dATP ratio of 5:1) and 0.5 mM $MnCl_2$, essentially as described by Leung et al. (Technique 1:11-15, 1989). Under these conditions, a mutation frequency of approximately 1×10^{-2} is feasible. Mutagenesis efficiency and accuracy was confirmed by DNA sequence determination as described in Example 1.

2.3 Molecular modeling and selection of amino acid positions

The structural model of the Gramidicin S Phe activating domain has been determined at 1.9 Å resolution (Conti et al. (1997) EMBO J.16: 4174-4183). The sequence identities between the different peptide synthetase domains are roughly between 30 and 60 %. Previously, homology-based modelling techniques have been successfully applied for the construction of three dimensional models of proteins based on similar or even lower levels of sequence conservation (Mosimann et al. (1995) Proteins 23: 301-317). The construction of models based on sequence similarity levels of this order has been shown to be quite accurate for the core regions, whereas some errors are to be expected in surface located regions. The regions in the activating domain that are responsible for amino acid binding can be regarded to form part of the core of the protein. Indeed, comparison of the crystal structures of the Gramidicin A Phe activating domain and firefly luciferase (16 % sequence identity), indicated a high degree of structural conservation, in particular in the substrate binding region (Conti et al. (1997) EMBO J. 16: 4174-4183). The structure of the Gramicidin A Phe activating domain has been obtained with the substrate (phenylalanine) in the active site.

For molecular modeling and database searches, the program WHAT IF was used (Vriend, G. (1990) J. Mol.

Graphics 8: 52-56). By this means, amino acids positions and substitutions suitable for modification of the substrate specificities (amino acid recognition) of peptide synthetase (Gramicidin S Phe activating domain and models constructed on the basis of homology) domains were predicted.

2.4 Purification and characterization of isolated modules

Isolation and characterization of individual modules were performed as described in Example 1.

3. Results

Homology-based molecular modeling was used to build models for the different activation domains of the surfactin and mycosubtilin synthetases. The amino acid positions that were homologous to those involved in binding of the phenylalanine substrate by Gramicidin S Phe activating domain were identified and mutated.

Peptide synthetase domains that have been altered by means of directed or random mutagenesis were expressed and produced in Escherichia coli, and the effects of the mutations on substrate specificity were determined.

EXAMPLE 6

Preparation of novel lipopeptides synthetase encoding DNA molecules

1. Introduction

Based on the DNA fragments contained in a library various methods can be applied to develop hybrid DNA molecules encoding novel lipopeptide or peptide synthetases. First, the order of the DNA fragments within a DNA molecule can be varied. Second, the amount of DNA fragments within a hybrid DNA molecule can be different than in existing peptide synthetases. This variation can be achieved by randomly ligating the DNA fragments to form a DNA molecule.

After transformation of the hybrid DNA molecules to a suitable host and expression of the DNA molecule, the expressed lipopeptide synthetases can be used to produce lipopeptides, either in vivo in the host cell or in vitro. The produced lipopeptides can subsequently be analyzed for their activity (cf. example 7). The number of different (lipo)peptides that can be obtained by creating different combinations of the modules is potentially enormous.

In this example all modules are obtained using PCR primers that have at the 5'-ends and 3'-ends restriction enzyme recognition sites that generate non-palindromic sticky ends such that the modules can be ligated in a predetermined order (**Fig. 16**). Examples of suitable restriction enzymes are SfiI, Eco31I, Esp3I and BpiI. These sites are located in the linker regions as defined under example 2.

The modules of the library in example 4 contain SfiI sites at their 5'- and 3'-ends (SfiI tags), such that the modules can be divided in two groups, library I and II. Members of library I can not ligate to each other, but only to one of the sticky ends of library II modules and members of library II can not ligate to each other, but only to one of the sticky ends of library I. In **Fig. 17** this is visualized and this means that the SfiI tags of Lib1R ligates only to Lib2L and Lib2R ligates only to Lib1L. The SfiI tags are designed such that in frame fusions of the modules are generated and such that they do not encode known proteolytic cleavage sites.

2.1 Assembly in solution

The concept of ligating modules in a predetermined order through the use of non-palindromic sticky-ends was analyzed using the SrfB1, MycB3 and SrfB3 modules. The SrfB1 module was generated using the SrfB1-U2 and SrfB1-D3 primers (Table 3). Consequently, this module contains only a SfiI tag (Lib1R) at its 3'-end.

The MycB3 module was generated using the primers indicated in Table 4, which result in Lib2 SfiI tags at both sides of the module. Finally, the SrfB3 module was obtained using the SrfB3-U and SrfB3-D2 primers (Table 3). PCR reactions were carried out with high fidelity polymerases as described above with a maximum of 20 cycles.

The modules were digested with SfiI and after purification approximately 1 μg of each module was used in a ligation reaction. A typical ligation was done at room temperature for 2 to 16 hours in a volume of 20 to 30 μl with 1 μl T4 ligase. At this point it is optional to include PEG during the ligation. The inclusion of PEG in the ligation mixture keeps the DNA in a linear form and may be advantageous, especially in ligations that contain more than three different modules. The procedure for the generation of linear DNA molecules using ligation in the presence of PEG was described by Pfeiffer & Zimmerman ((1983) Nucl. Acids. Res. 11: 7853-7871).

After ligation the DNA was purified by PEG precipitation and the full length product was amplified to yield sufficient coupled product for transformation. The primers used in this PCR reaction were SrfB1-U4 (5'-TGCTGTATCATGCGATGCTTGATCC-3') and SrfB3-D4 (5'-TCGTCCATTTCAAGGTCTTCGGCG-3'). The reaction conditions were similar as described before with a maximum of 20 cycli.

The linear SrfB1-MycB3-SrfB3 DNA molecules thus obtained were transformed to competent BM20 cells (**Fig. 13**). Selection and screening of transformants was as described before. This yielded strain BM25 (**Fig. 13**) that was Km sensitive, had lost LacZ (white colony on agar plates with Xgal) and showed hemolytic activity similar to BM23 on blood agar plates (**Fig. 14**). The in frame insertion of the MycB3 module in strain BM25 was confirmed by sequencing the chromosomal DNA in this particular region.

Taken together, these results demonstrate that SfiI tags can be inserted in the linker regions indicated in **Fig. 6** and that the non-palindromic sticky ends can be used to assemble modules in a predetermined order, and that this results in novel active lipopeptide synthetases in vivo using a B. subtilis 168 derivative.

2.2 Assembly using a solid surface

Although random ligations of modules in solution are feasible, the method benefits considerably by using a solid surface.

The coupling of DNA fragments to a solid surface was tested using Nunc NucleoLink PCR tubes (alternatively, Streptavidin-coated PCR tubes of Roche may be used). Oligonucleotides containing a 5'-phosphate (P) group are chemically coupled to these PCR tubes (a 5'-Biotin group in case of the Steptavidin-coated tubes, in this case the coupling is not chemically but by affinity). Such an oligonucleotide was designed, Lib1R (5'-P-TTTTTTTTTTGGGGGCGGCC GCCCGGGCCGGGG), and 200 ng of it was chemically coupled to the PCR tube according to the instructions of the supplier. After washing steps with SHT buffer (0.4 M NaOH, 0.25% Tween 20) according to the protocol, additional washing steps (6 times) were done with 200 µl prewarmed (50°C) TNT buffer (100 mM Tris-HCl, pH 7.5; 150 mM NaCl; 0.1% Tween 20).

A second oligonucleotide (Lib1-D: 5'-CGGCCCGGGCGGCCGCCCCC), complementary to Lib1R, was subsequently annealed to the covalently attached Lib1R oligonucleotide (**Fig. 19**). The annealing was done in PCR buffer with 500 ng Lib1-D in a total volume of 100 µl. The mixture was incubated in a thermocycler for 2 min at 94°C, 30 min 68°C and 30 min at 50°C. Oligonucleotides that were not annealed were washed away with 6 times 200 µl prewarmed (50°C) TNT buffer. By the annealing of Lib1-D to Lib1R a SfiI-Lib1 sticky end was generated (**Fig. 19**). In the next step a SfiI digested Lib2 module was ligated to this sticky end. As a Lib2 module the Km gene

was used that had been generated using the primers KMR01 **atactaaggccgggggggcc**tgataggtaagattataccg (in bold the SfiI Lib2L tag) and KMR02 **taatgtaggccgcggcggccgc**gggacccctatctagcgaac (in bold the SfiI Lib2R tag). After digestion of the purified Km module with SfiI, 1-5 μg purified DNA was ligated to the oligonucleotides on the surface of the PCR tube in standard ligation buffer with 2.5 μl T4 ligase (Roche) in a total volume of 50 μl at room temperature for 2 to 16 hours (optional is to include PEG as described before). Non-ligated fragments were washed away with TNT buffer as described above.

Subsequently, a SfiI digested Lib1 module was ligated to the Km module. The SrfB3 module, obtained by PCR amplification using primers SrfB3-U and SrfB3-D (Table 3), was ligated to the Km module in the same way as the previous ligation. After ligation all non-ligated modules were removed by six washes with 200 μl TNT buffer. The ligated product could be obtained in two ways:

(i) by liberating the ligation product from the PCR tube using one of the restriction endonuclease enzymes that have a recognition site in the annealed Lib1 oligonucleotides (**Fig. 19**; in this case there are sites for NotI, SmaI and SrfI, but any other recognition site can be incorporated);

(ii) by PCR amplification. In this case the PCR tubes are first washed 6 times with 200 μl SHT buffer in order render the ligated product single stranded, followed by an additional 6 washes with TNT buffer. The PCR amplification was done using high fidelity PCR enzymes for 20 cycli with standard conditions.

In both cases the desired ligation product could be obtained. The PCR method having the advantage that the PCR tube remains as a backup of the generated hybrid DNA molecules.

This method of the ligation of modules to the surface of PCR tubes (**Fig. 18 and 19**) opens the

possibility to generate an enormous variety of new hybrid lipopeptide synthetases that can be integrated in the chromosome of a bacterial host and subsequently expressed. The expressed hybrid lipopeptide synthetases can then be used to produce novel lipopeptides, either in vivo in the host cell or in vitro. The produced lipopeptides can subsequently be analyzed for their activity as described in example 7. The method is also applicable for the generation of peptide synthetases that lack FAS and PKS modules and domains.

EXAMPLE 7

Screening of the library for colonies that express a (lipo)peptide synthetase producing a lipopeptide having a desired biological activity

1. Introduction

For the selection of novel (lipo)peptides with altered specificity efficient screening methodologies are of great importance. To improve the efficiency direct screening methods for activity towards the viruses or cells against which activity is sought has been developed. Indicator microorganisms (e.g. pathogenic bacteria and fungi) are selected, that are not, or only moderately sensitive to known (lipo)peptides. Furthermore, special emphasis is given to the use of microorganisms that are known to be involved in the contamination of food and feed.

After transferring Bacillus subtilis transformed with a library of mutated and reconstituted (lipo)peptide synthetase genes to plates containing such indicator-microorganisms, those B. subtilis variants expressing altered (lipo)peptides with antibiotic activity against the indicator-organism, can be selected by their ability of halo formation, caused by lysis of the indicator strain. Availability of such screening methods enables the use of directed evolution strategies to select for variants with altered or improved properties.

Moreover, novel (lipo)peptides can be selected for enzyme inhibiting properties by means of high-throughput screening technologies. Interesting enzymes to be screened for inhibitors include matrix-metalloproteases or other disease-related enzymes.

2. Materials and methods

2.1 Use of erythrocytes

Competent B. subtilis cells that have an inactivated (lipo)peptide synthetase encoding DNA fragment on their genome were transformed with the libraries obtained (see Example 6). A vast majority of the transformed cells will have a recombined synthetase DNA fragment. To select for active variants, the cells can be transferred to agar plates that contain erythrocytes, that are very sensitive indicators for most (lipo)peptides. Lysis of erythrocytes occurs due to production of active (lipo)peptide variants, which can be seen as the formation of a clearing zone around the colonies.

2.2 Use of specific indicator organisms

Active variants, identified as those that lyse erythrocytes, are transferred to agar plates that contain microbials (bacteria or yeasts) against which improved antibiotics are needed. Those variants that lyse the indicators cells will be identified in more detail. Alternatively, the transformation mixtures can be transferred to the plates containing the indicator organisms directly, without the need of preselection on erythrocyte lysing activity.

2.3 Use of specific cell lines

Active variants, identified as those that lyse erythrocytes, are transferred into microtiter plates in which, e.g. tumor, cell lines are propagated. By this means, (lipo)peptides with anti tumor cell activity can be identified.

3. Results

By using erythrocytes as indicator, several novel (lipo)peptides have been obtained, after transforming competent B. subtilis cells in which the (lipo)peptide synthetase encoding DNA fragment was inactivated, with PEG ligated modules from the constructed library. These novel peptides are the result of the unique combination of the DNA uptake and recombination properties of the recipient Bacillus subtilis cells and the use of randomly-ligated linear (lipo)peptide module encoding DNA fragments.

Table 3

Primer sequences for construction of BM04, 20, 23 and 25

DNA fragment	primer name	primer sequence 5'→ 3'
SrfB1	SrfB1-U2	ATGAGCAAAAAATCGATTCAAAAGG
	SrfB1-D3	TAATGTAGGCCCCCCCGGCCGCCAAATACAGTGCCAGTTCTTGAATAGTCG
SrfB3	SrfB3-U	ATACTAAGGCCGCCGCGGCCGAGGCGTCTGCTGTTTCAGG
	SrfB3-D2	TTTTAAATTCTCCTCAAGCATG
SrfB1	SrfB1-U2	ATGAGCAAAAAATCGATTCAAAAGG
	SrfB1-D2	CAGGAGCTGGTCTAATTCGTCCAAATACAGTGCCAGTTCTTG
MycB3	MycB3-U	ATACTAAGGCCGGGGGGGCCGACGAATTAGACCAGCTCCTG
	MycB3-D	TAATGTAGGCCGCGGCGGCCGCGATTGCCTGTGCGAGCTGTTC
SrfB3	SrfB3-U2	GAACAGCTCGCACAGGCAATCGAGGCGTCTGCTGTTTCAGG
	SrfB3-D2	TTTTAAATTCTCCTCAAGCATG

Table 2

Primer sequences for construction of BM24, 28 and 30

DNA fragment	primer name	primer sequence 5'→ 3'
Srf5'	Srf5'-U	CCAGGCATTCTGATGATGAGGTCCGC
	Srf5'-D	TAATGTAGGCCCCCCCGGCCGCTCTTTATAAAGCAGTGAACATGTGCA ACGC
Km	KMR01	ATACTAAGGCCGGGGGGGCCTGATAGGTAAGATTATACCG
	KMR03	GGGACCCCTATCTAGCGAAC
LacZ	KMR04	GTTCGCTAGATAGGGGTCCCCTACACAGCCCAGTCCAGAC
	LACZ-D	TAATGTAGGCCGCGGCGGCCGCGGAAGGAAATGATGACCTCG
SrfA1	SrfA1-U2	ATACTAAGGCCGCCGCGGCCGAAATAACTTTTTACCCTTTAACGGATGCAC
	SrfA1-D2	TCTCAACGTTCAGCACGTCCTGGC
FAS	FAS-U3	ATACTAAGGCCGGGGGGGCCATTAGGGGAGGTATGACAATATGAATAAT CTTGCCTTTTTATTTCCGGG
	FAS-D	CTCAACAGGCTTCATCGGGGGATACTTGAC
SrfA1	SrfA1-U1	GTCAAGTATCCCCCGATGAAGCCTGTTGAGGAAATAACTTTTTACCCTTTA ACGGATGCAC
	SrfA1-D2	TCTCAACGTTCAGCACGTCCTGGC
FAS + MycA1	FAS-U3	ATACTAAGGCCGGGGGGGCCATTAGGGGAGGTATGACAATATGAATAAT CTTGCCTTTTTATTTCCGGG
	MycA1-D1	CGCCATGTTAGCCAACTGTTCAATAGTAAG
SrfA2	SrfA2-U1	CTTACTATTGAACAGTTGGCTAACATGGCGGATCACAGAGAAAGCAAAG CTTTTGCG
	SrfA2-D1	AATCACCTGTGCCAAGCCCTCAAC
yhjM5'	yhjM-UF	TGTGCCTGCTCCATATCCTAAG
	yhjM-UR	TAATGTAGGCCCCCCCGGCCATGGATGGTCGATACTCGGAAC
Cm	CAT-F	ATACTAAGGCCGGGGGGGCCCCGGCAATAGTTACCCTTATTATC
	CAT-R	TAATGTAGGCCGCGGCGGCCACTAACGGGGGCAGGTTAGTGAC
yhjM3'	yhjM-DF	ATACTAAGGCCGCCGCGGCCTCATCACTCAGCGGACATCACG
	yhjM-DR	TCCGAATACCTCTGATAGCCAC

Table 4

Primer sequences for construction of a library

DNA fragment	primer name	primer sequence 5'→ 3'
SrfA1	SrfA1-U2	ATACTAAGGCCGCCGCGGCCGAAATAACTTTTTACCCTTTAACGGATGCAC
	SrfA1-D3	TAATGTAGGCCCCCCCGGCCGCTCTCAACGTTCAGCACGTCCTGGC
SrfA2	SrfA2-U2	ATACTAAGGCCGGGGGGGCCGATCACAGAGAAAGCAAAGCTTTTGCG
	SrfA2-D2	TAATGTAGGCCGCGGCGGCCGCAATCACCTGTGCCAAGCCCTCAAC
SrfB1	SrfB1-U	ATACTAAGGCCGCCGCGGCCATGAGCAAAAAATCGATTCAAAAGG
	SrfB1-D3	TAATGTAGGCCCCCCCGGCCGCCAAATACAGTGCCAGTTCTTGAATAGTCG
SrfB2	SrfB2-U2	ATACTAAGGCCGGGGGGGCCGAAGAGAACGAAAGCAAGGAGGAGCAG
	SrfB2-D3	TAATGTAGGCCGCGGCGGCCGCCATATAAGCCGCCAGCTGGCGAACTGTC G
SrfB3	SrfB3-U	ATACTAAGGCCGCCGCGGCCGAGGCGTCTGCTGTTTCAGG
	SrfB3-D	TAATGTAGGCCCCCCCGGCCGCTTTTAAATTCTCCTCAAGCATG
MycB3	MycB3-U	ATACTAAGGCCGGGGGGGCCGACGAATTAGACCAGCTCCTG
	MycB3-D	TAATGTAGGCCGCGGCGGCCGCGATTGCCTGTGCGAGCTGTTC
FAS	FAS-U3	ATACTAAGGCCGGGGGGGCCATTAGGGGAGGTATGACAATATGAATAAT CTTGCCTTTTTATTTCCGGG
	FAS-D2	TAATGTAGGCCGCGGCGGCCGCCTCAACAGGCTTCATCGGGGGATACTTG AC
FAS + MycA1	FAS-U3	ATACTAAGGCCGGGGGGGCCATTAGGGGAGGTATGACAATATGAATAAT CTTGCCTTTTTATTTCCGGG
	MycA1-D2	TAATGTAGGCCGCGGCGGCCGCCGCCATGTTAGCCAACTGTTCAATAGTA AG

Underlined sequences are *Sfi*I tags.

CLAIMS

1. Method for the preparation of novel lipopeptide synthetases by modifying the lipid moiety and optionally also the peptide moiety of one or more known (lipo)peptide synthetase, which method comprises the steps of:

a) providing a range of DNA fragments encoding lipopeptide synthetase modules or domains thereof;

b) selecting and ligating at least two of these DNA fragments, at least one of which is specific for the lipid moiety of the lipopeptide synthetase, in order to obtain a lipopeptide synthetase encoding DNA molecule;

c) introducing the lipopeptide synthetase encoding DNA molecule in a host cell;

d) expressing the DNA molecule in the host cell.

2. Method as claimed in claim 1, wherein the lipopeptide synthetase encoded by the DNA molecule has a module configuration that does not exist in nature.

3. Method as claimed in claim 1 or 2, wherein the DNA fragments encoding lipopeptide synthetase modules or domains thereof comprise DNA fragments encoding fatty acid or polyketide synthase and (lipo)peptide synthetase modules or domains thereof.

4. Method as claimed in claims 1-3, wherein one or more of the DNA fragments are mutated prior to expression thereof.

5. Method as claimed in claim 4, wherein the mutation is effected by means of error-prone PCR.

6. Method as claimed in claim 4, wherein the mutation is effected by means of site-directed mutagenesis.

7. Method for the preparation of lipopeptide and peptide synthetases, which method comprises the steps of:

a) providing a range of DNA fragments encoding lipopeptide or peptide synthetase modules or domains

thereof, which DNA fragments contain at their 5'-end and 3'-end restriction enzyme recognition sites that generate non-palindromic sticky ends;

b) selecting and ligating at least two of these DNA fragments, in order to obtain a lipopeptide or peptide synthetase encoding DNA molecule;

c) introducing the lipopeptide synthetase encoding DNA molecule in a host cell;

d) expressing the DNA molecule in the host cell.

8. Method as claimed in claim 7, wherein the DNA fragments are generated by PCR using primers containing restriction enzyme recognition sites that generate non-palindromic sticky ends.

9. Method as claimed in claims 1-8, wherein the ligation is a random ligation.

10. Method as claimed in claims 1-9, wherein the ligation is effected in the presence of polyethylene glycol (PEG) and T4 DNA ligase such that linear DNA molecules are obtained.

11. Method as claimed in claims 1-10, wherein the ligation is effected on a solid surface.

12. Method for ligating DNA fragments encoding lipopeptide or peptide synthetase modules or domains thereof comprising the steps of:

a) providing a solid surface having coupled thereto a first DNA fragment encoding a start DNA sequence comprising at its 3'-end the sticky end of a non-palindromic restriction enzyme recognition site;

b) ligating thereto a DNA fragment encoding a lipopeptide or peptide synthetase module or domain thereof, which DNA fragment contains at its 5'-end a sticky end that is complementary to the sticky end at the 3'-end of the previously incorporated DNA fragment and at its 3'-end the other sticky end of the non-palindromic restriction enzyme recognition site;

c) removing the non-ligated DNA fragments by washing; and

d) optionally repeating steps b) and c) one or more times;

e) ligating one or more DNA stop fragments to the DNA fragment last incorporated, wherein the DNA stop fragments contain at their 5'-end a sticky end that is complementary to the sticky end at the 3'-end of the previously incorporated DNA fragment and at their 3'-end the other sticky end of the non-palindromic restriction enzyme recognition site to obtain a bound DNA molecule encoding a lipopeptide or peptide synthetase.

13. Method as claimed in claim 12, wherein the first DNA fragment encoding a start DNA sequence and the last of the end fragments correspond to chromosomal fragments of the host organism to allow integration of the DNA molecule into the host genome.

14. Method as claimed in claim 12 or 13, wherein one of the end fragments, but not the last end fragment is a selection marker, in particular an antibiotic selection marker.

15. Method for ligating DNA fragments encoding lipopeptide or peptide synthetase modules or domains thereof comprising the steps of:

a) providing a solid surface having coupled thereto a first oligonucleotide encoding one of the strands of a start DNA sequence, which strand comprises at least the nucleotides complementary to a non-palindromic restriction enzyme recognition site at its 3'-end;

b) annealing to the first oligonucleotide a complementary second oligonucleotide in a manner that generates the sticky end of a non-palindromic restriction enzyme recognition site at the 3'-end of the double strand thus obtained;

c) ligating a DNA fragment encoding a lipopeptide or peptide synthetase module or domain thereof, which DNA fragment contains at its 5'-end the sticky end of the same non-palindromic restriction enzyme

recognition site as the double strand obtained in step b) at its 3'-end ;

d) removing the non-ligated DNA fragments by washing; and

e) optionally repeating steps c) and d) one or more times;

f) ligating one or more DNA fragments encoding DNA stop sequences to the DNA fragment last incorporated, wherein the DNA fragments contain at their 5'-end a sticky end that is complementary to the sticky end at the 3'-end of the previously incorporated DNA fragment and, optionally except for the last incorporated DNA stop sequence, at their 3'-end the other sticky end of the non-palindromic restriction enzyme recognition site, to obtain a bound DNA molecule encoding a lipopeptide or peptide synthetase.

16. Method as claimed in claims 12-15, wherein the DNA molecule encoding a lipopeptide or peptide synthetase is used as a template in a PCR reaction to amplify the DNA molecule.

17. Method as claimed in claim 16, wherein the solid surface is a PCR tube.

18. Method as claimed in claim 16 or 17 wherein the DNA molecule remains bound to the solid surface during the PCR reaction and thereafter.

19. Method as claimed in claims 1-18, wherein the DNA molecule is introduced into naturally competent host cells.

20. Method as claimed in claims 1-19, wherein the host cell is a Bacillus species.

21. Method as claimed in claim 20, wherein the host cell is Bacillus subtilis.

22. Method as claimed in claims 1-21, further comprising the step of selecting a host cell harboring a DNA sequence encoding a (lipo)peptide synthetase which produced a peptide having a desired biological activity.

23. Method as claimed in claim 22, wherein the selection is effected by contacting the host cell or the

culture medium used for growing the host cell with an indicator micro-organism.

24. Method as claimed in claims 14 and 19-23, which comprises amplification, by means of the polymerase chain reaction and selected pairs of oligonucleotides, and isolation of DNA fragments that encode lipopeptide synthetase modules, including those encoding the fatty acid synthetase domains, the substrate specificity of which needs not to be known; ligation of mixtures of amplified module encoding DNA fragments in the presence of polyethylene glycol (PEG) and T4 DNA ligase, to induce the formation of linearly fused modules in a random order and number; transformation of natural competent Bacillus species cells that optionally have an inactivated surfactin synthetase and/or mycosubtilin synthetase operon; cultivation of said transformed cells under circumstances which lead to expression of the lipopeptide synthetase and production of the corresponding lipopeptide; screening the cells to specifically select for lipopeptides having novel biological activities; and application of the resulting lipopeptides.

25. Method as claimed in claim 24, wherein a mutation of the module encoding DNA fragments is additionally performed, for example effected by means of site-directed mutagenesis of particular amino acid positions within the modules or domains, which positions have been selected on the basis of molecular models of said modules or domains, that have been constructed, using molecular modelling techniques, on the basis of sequence similarity with the known three-dimensional structure of the phenylalanine activating domain of Gramicidin S.

26. Method as claimed in claim 24 or 25, wherein screening of the produced lipopeptides is done on a selective medium, which consists of a standard solid cultivation medium that is covered with a microbial organism (bacterium, fungus or yeast) against which a novel antibiotic is searched, wherein the cells to be

selected as antibiotic producing cells are those cells having the ability to lyse said microbe.

27. Method as claimed in claims 24-26, wherein the lipopeptide synthetase operon present on the wild type genome of the host bacterium has been exchanged with a hybrid synthetase, which is produced by fusing the wild type synthetase operon to the fatty acid synthetase encoding part of the mycA open reading frame of the mycosubtilin operon, resulting in the coupling of a ß-amino fatty acid to the peptides produced by this operon.

28. Method as claimed in claim 27, wherein in addition, the fatty acid synthetase encoding part may be modified by homologous recombination with, or the addition of, modules of polyketide synthetases, for example derived from natural or modified polyketide synthetases.

29. Method for the non-ribosomal preparation of novel peptides, comprising the preparation of a novel peptide synthetase by means of the method as claimed in claims 1-11 and 19-28 and using this peptide synthetase to produce the corresponding peptide.

30. Method as claimed in claim 29, wherein the peptide is produced in the host cell in which the peptide synthetase is expressed.

31. Method as claimed in claim 30, wherein the peptide synthetase is isolated from the host cell or the culture medium used for growing the host cell prior to using the peptide synthetase for the production of the corresponding peptide.

32. Lipopeptide having a novel structure and/or novel activity in comparison to known peptides, obtainable by the method as claimed in claims 24-28.

33. Lipopeptide having a novel structure and/or novel activity in comparison to known peptides, obtainable by the method as claimed in claims 29-31.

34. Lipopeptide having a novel structure and/or novel activity in comparison to known peptides, and being

the product of a peptide synthetase obtainable by the method as claimed in claims 1-11 and 19-23.

35. DNA molecule encoding a lipopeptide synthetase different from any known lipopeptide synthetase, comprising at least two DNA fragments encoding lipopeptide synthetase modules or domains thereof, at least one of which is specific for the lipid moiety of the lipopeptide synthetase.

36. DNA molecule as claimed in claim 35, comprising at least the mycA module encoding DNA fragment of the mycosubtilin operon of Bacillus subtilis.

37. Host cell harboring a DNA molecule as claimed in claims 35 or 36.

38. Host cell as obtained in the method as claimed in claims 1-11 and 19-28.

39. Library of DNA fragments encoding lipopeptide synthetase modules or domains thereof.

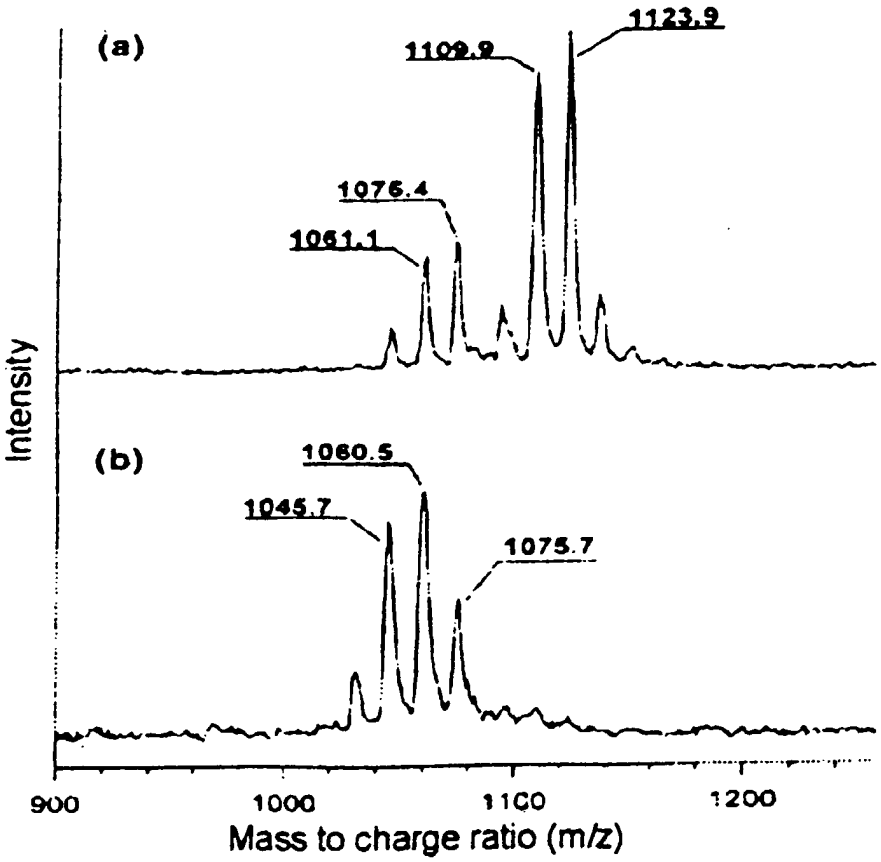
40. Library as claimed in claim 39, wherein one or more DNA fragments have been mutated as compared to the corresponding DNA fragment existing in nature.

41. Library as claimed in claims 39 or 40 for use in the production of a range of lipopeptide synthetases as described in the method according to claims 1-11 and 19-28.

42. PCR-tube having bound to its inner surface a DNA molecule encoding a lipopeptide or peptide synthetase, which tube is obtainable by performing the method of claims 15-17.

43. PCR-tube as claimed in claim 42, wherein the DNA molecule encoding a lipopeptide or peptide synthetase bound to the tube is for use a template in a PCR reaction.

Fig.1



	Condensation	Adenylation	Thiolation	Condensation/ Epimerization
MycA1	HHIISDGYSM -460-	SGSTGNPKGV -214- YKTGD -152-	FFDLGGNSI -177-	HHIISDGVSM
MycB1		SGTTGKPKGV -215- YKTGD -152-	FFDFGGDSI -179-	**HHLVIDGVSW**
MycB2	HHILMDGWCL -461-	SGSTGKPKGV -214- YKTGD -152-	FFDFGGDSI -178-	**HHLIVDGISW**
MycB3	HHIVMDGWCL -462-	SGTTGTPKGT -229- YRTGD -152-	FFDIGGHSL -178-	
MycB4	HHIISDGTSM -460-	SGTTGKPKGV -215- YKTGD -152-	FFDLGGHSL -178-	**HHIVTDGASM**
MycC1		SGSTGKPKGV -224- YRTGD -150-	FFELGGDSI -178-	HHLVIDGVSW
MycC2	HHILMDGWCL -461-	SGSTGKPKGV -214- YKTGD -151-	FFDVGGHSL	

Epimerization

MycB1	-97- AYNTEINDLLLTSLGV	-15-	EGHGREPVIP	-5-	TRTVGWFTSQYPVVLKM	-49- FNYLGQFD
MycB2	-97- AYNTEINDLLLSSLGL	-15-	EGHGREQVIP	-5-	SRTVGWFTSLFPVVLHI	-49- FNYLGQFD
MycC1	-98- AYTTEMNDLLLTGLGI	-15-	EGHGRESIIP	-5-	SRTVGWFTSQYPVFLPI	-49- FNFLGQFD

Fig.3

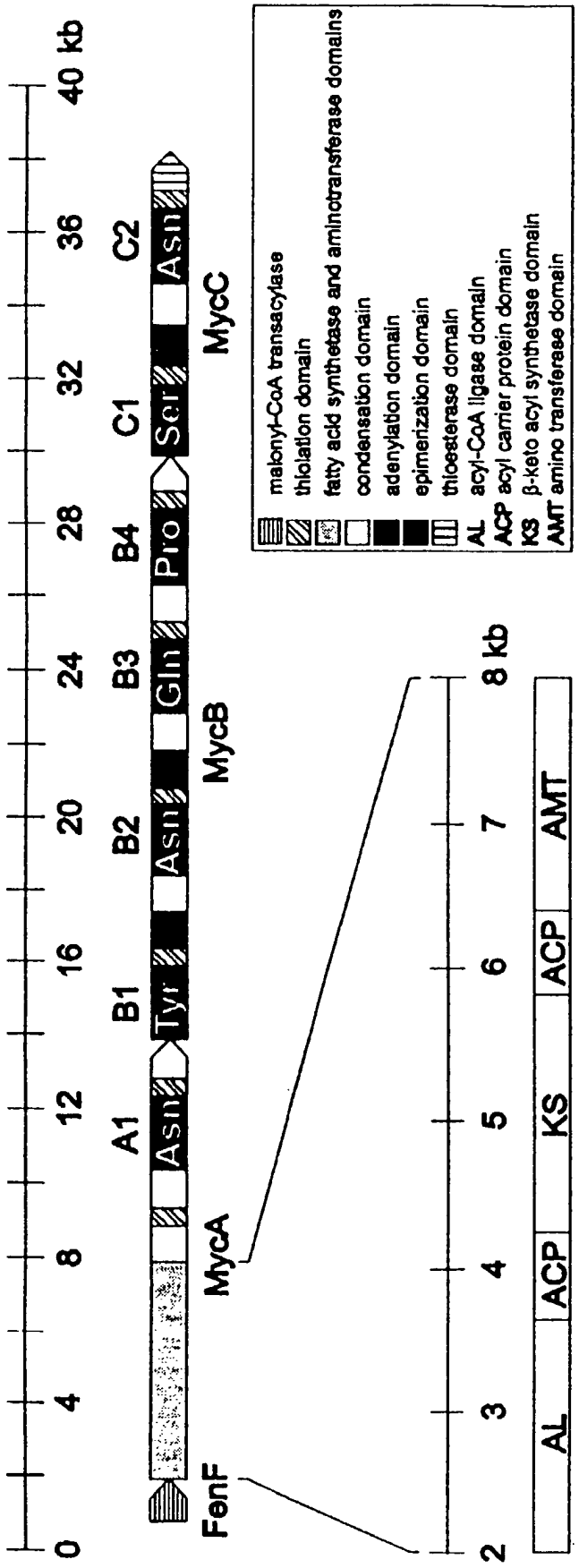

0
4
8
12
16
20
24
28
32
36
40 kb
A1
B1
B2
B3
B4
C1
C2
Asn
Tyr
Asn
Gln
Pro
Ser
Asn
FenF
MycA
MycB
MycC
malonyl-CoA transacylase
thiolation domain
fatty acid synthetase and aminotransferase domains
condensation domain
adenylation domain
epimerization domain
thioesterase domain
AL acyl-CoA ligase domain
ACP acyl carrier protein domain
KS β-keto acyl synthetase domain
AMT amino transferase domain
2
3
4
5
6
7
8 kb
AL
ACP
KS
ACP
AMT

Fig. 2

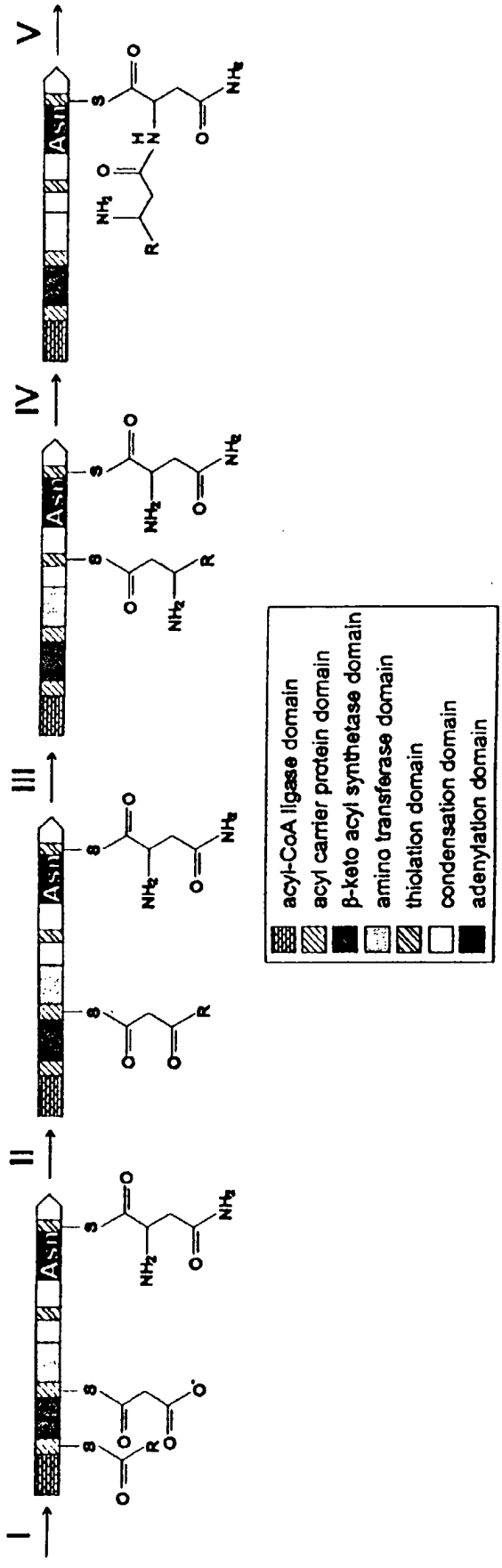

I
II
III
IV
V
Asn
acyl-CoA ligase domain
acyl carrier protein domain
β-keto acyl synthetase domain
amino transferase domain
thiolation domain
condensation domain
adenylation domain

Fig. 4

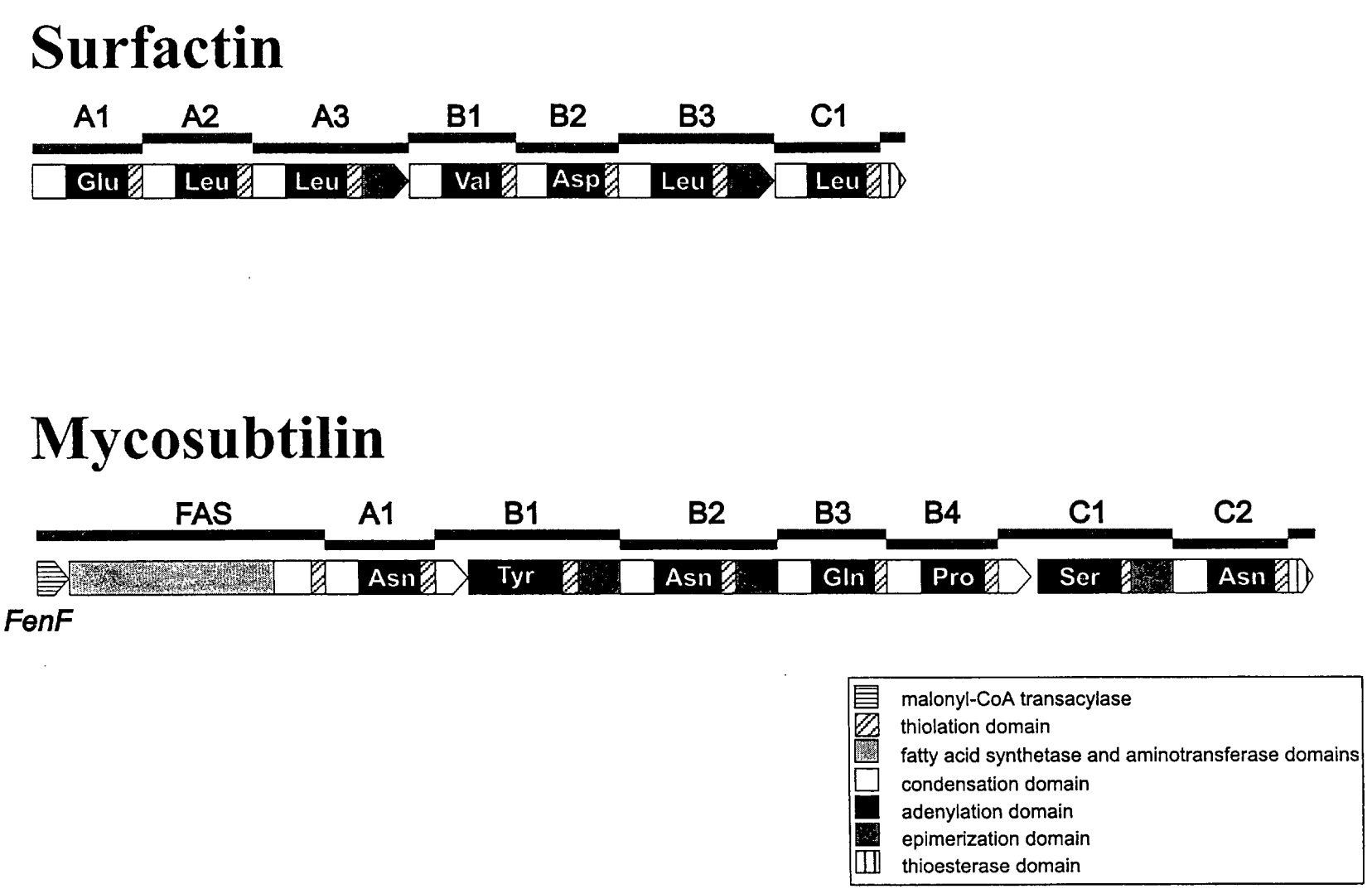


Fig. 5. Module composition of surfactin and mycosubtilin synthetases.

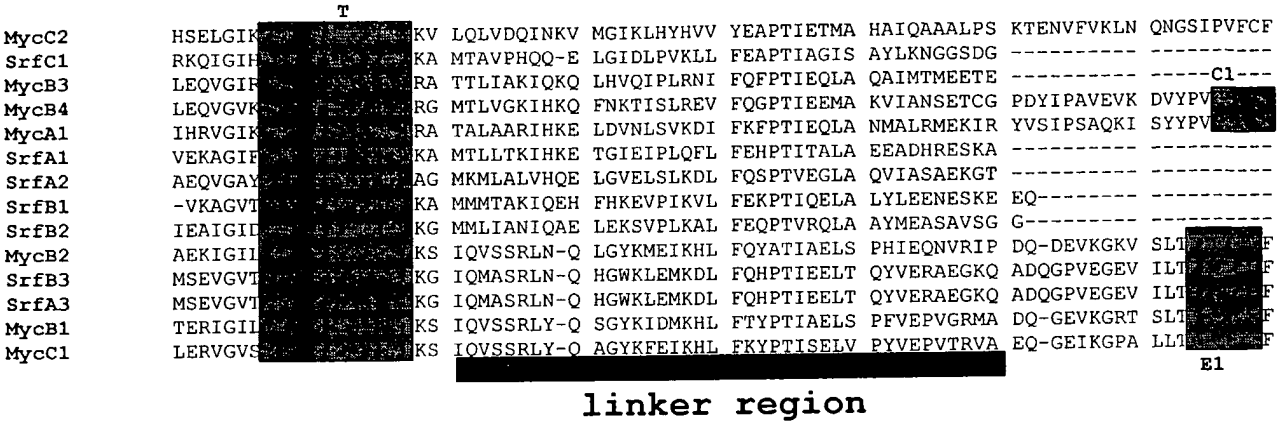


Fig. 6. Linker region between thiolation domain (T) and condensation domain (C) or epimerization domain (E).

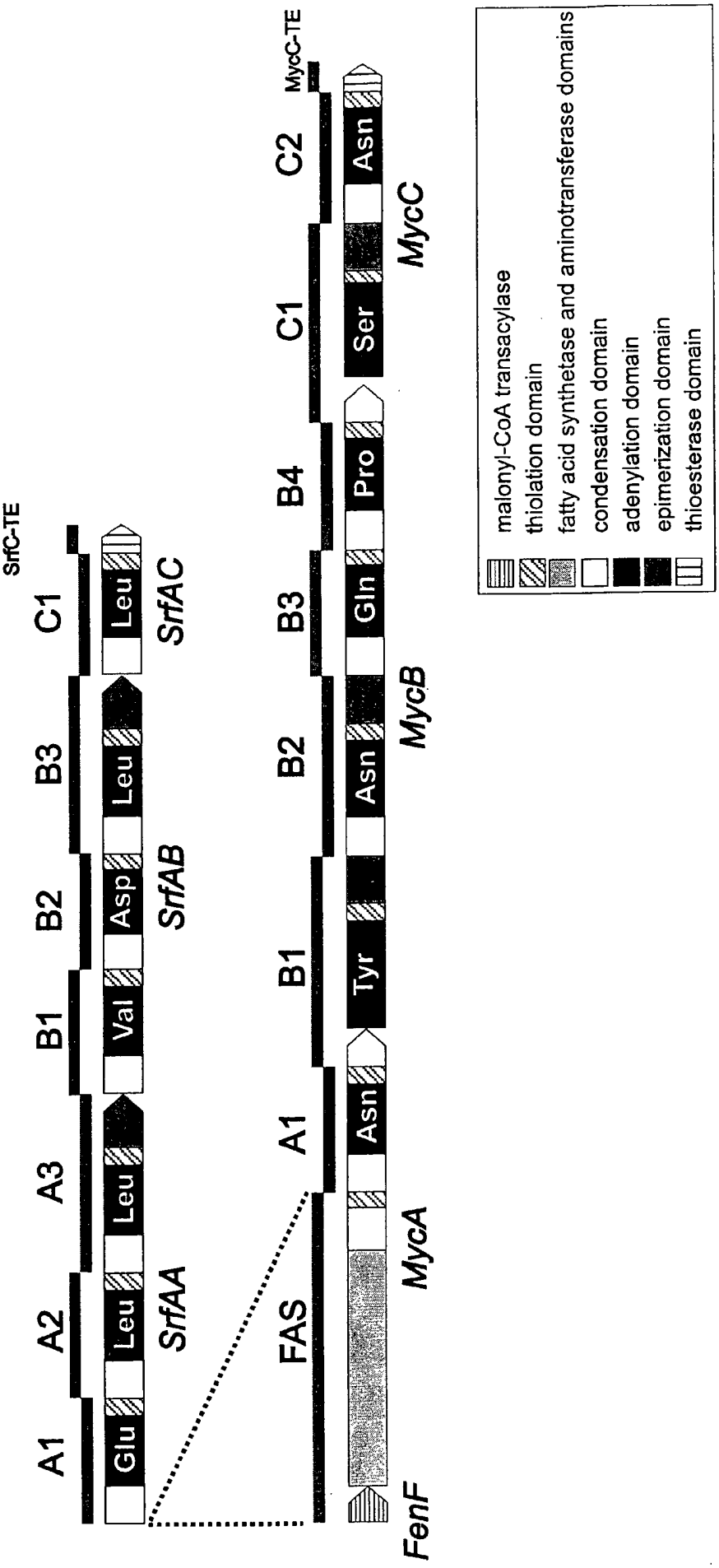


Fig. 7. Fusion of Myc fAS to first module of *srf* operon.

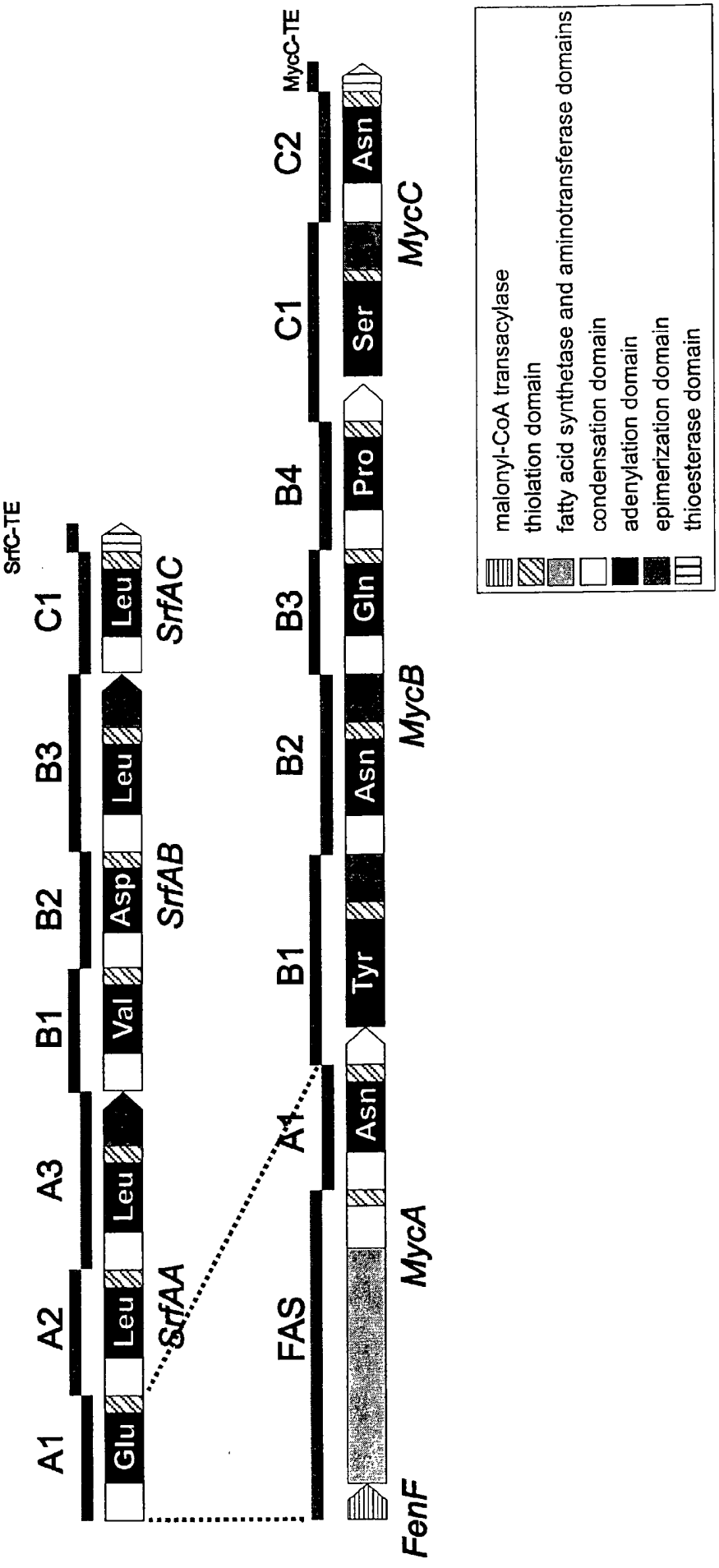


Fig. 8. Replacement of SrfA1 by Myc FAS + MycA1.

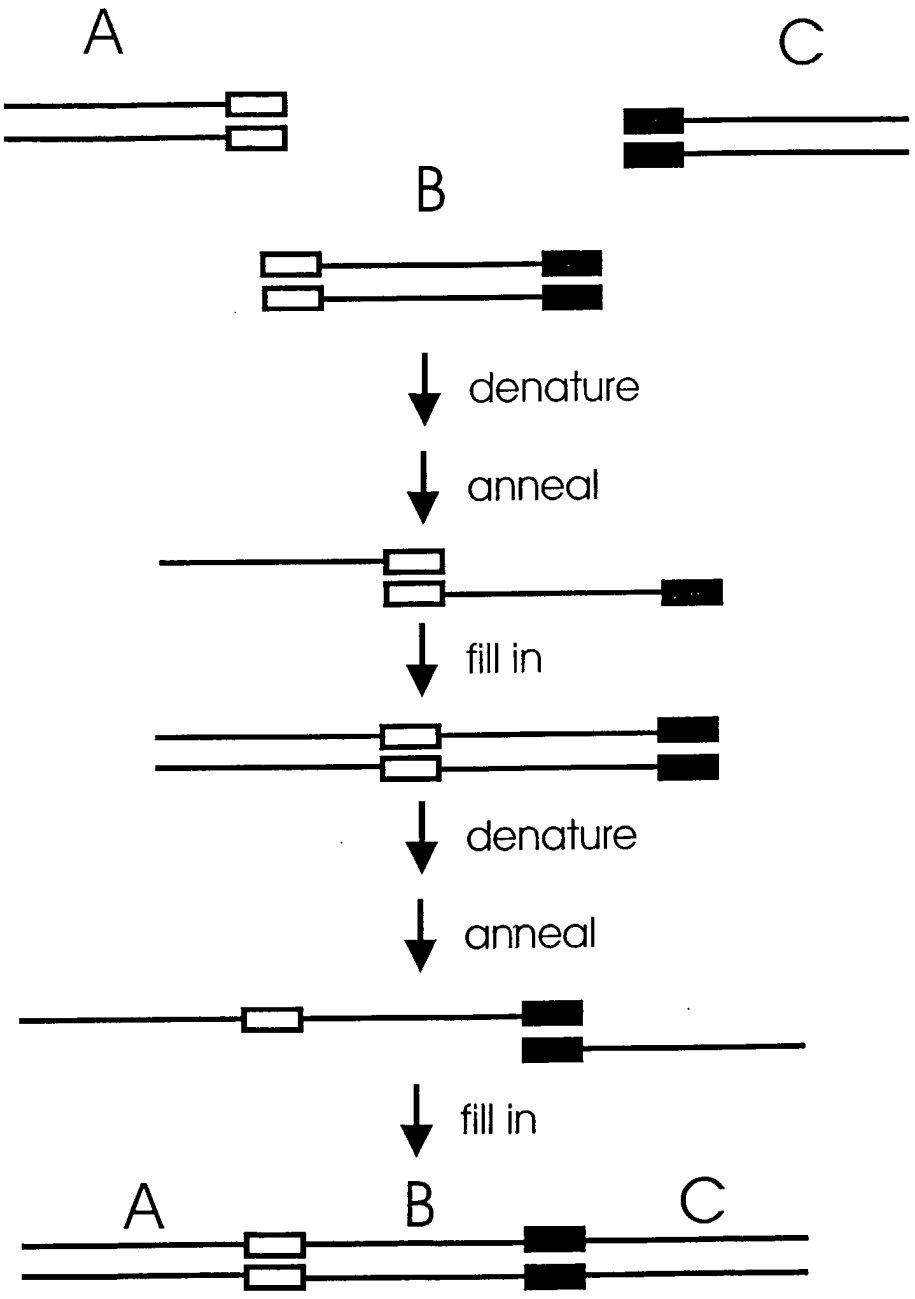


Fig. 9. Coupling of DNA modules by overlapping sequences.
▭ and ▬: overlapping sequences.

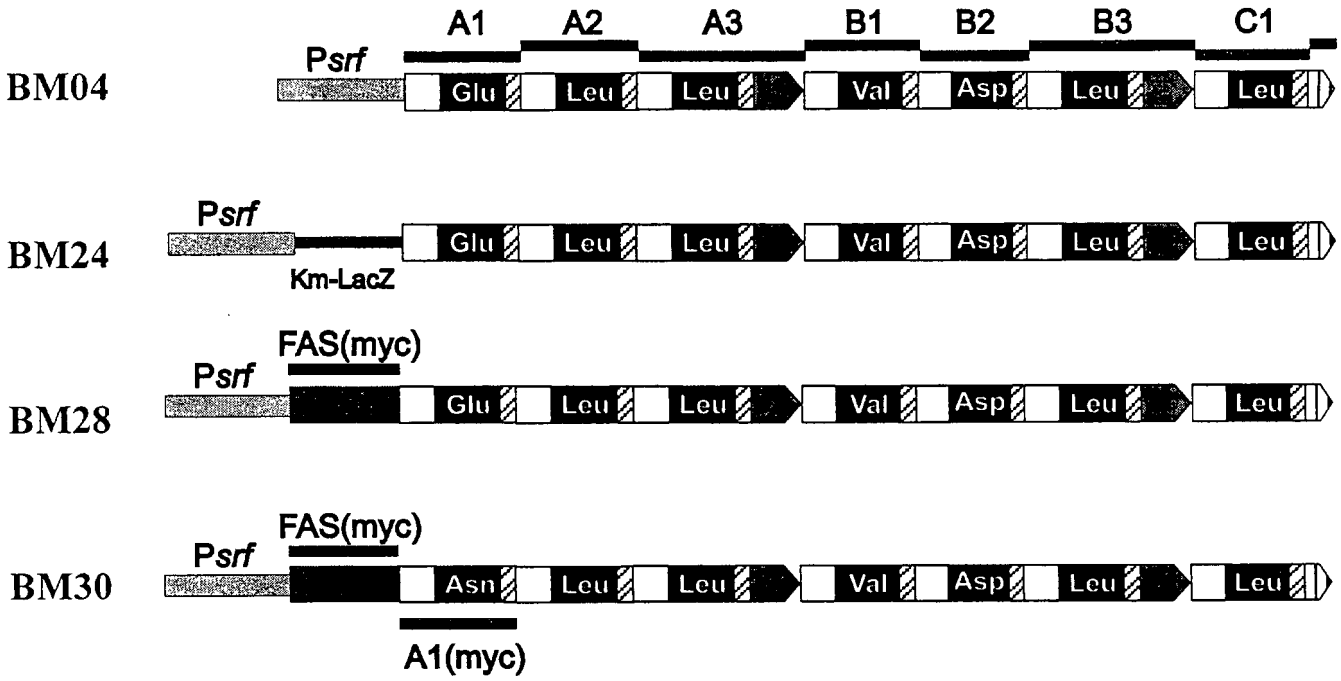


Fig. 10. Genetic organisation of BM04, 24, 28 and 30 in the *srf* locus and the phenotypic characteristcs of these strains.

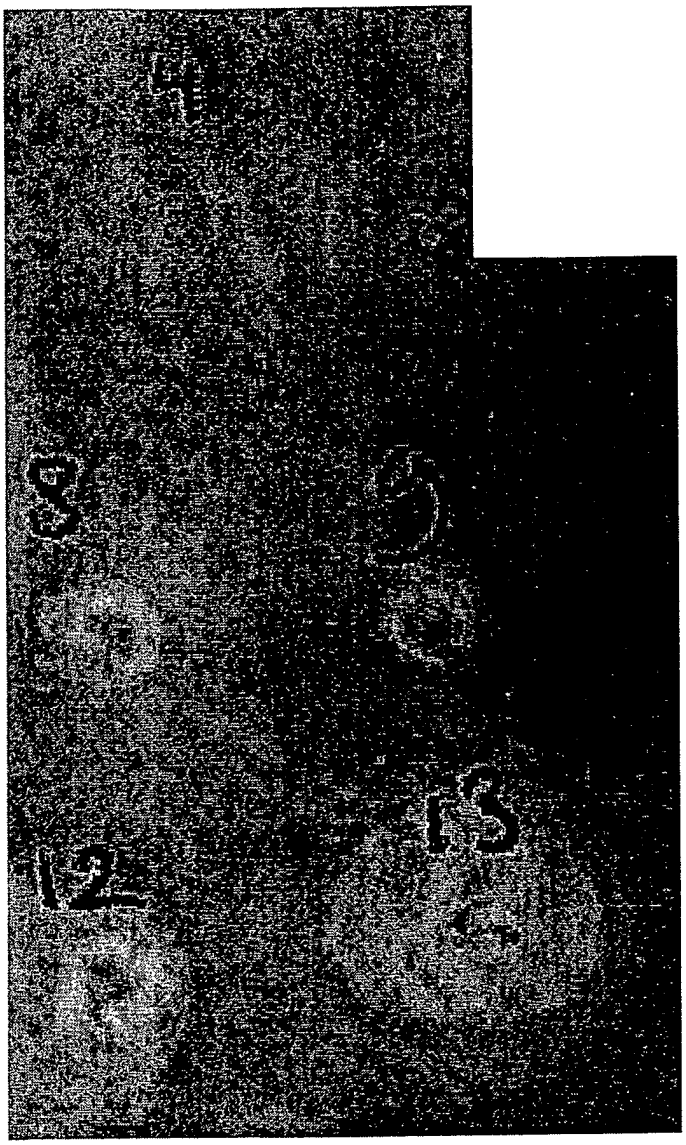
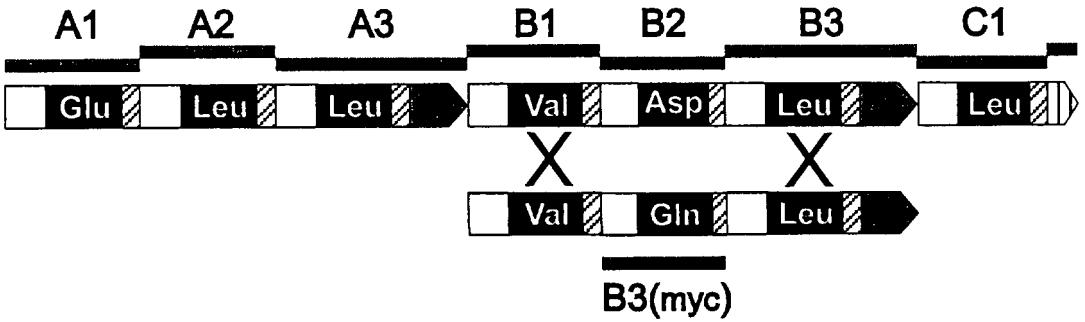


Fig. 11. Hemolytic activities of strains BM04 (12), BM24 (4), BM28 (13), BM30 (8) and BM2823 (9) as visualized ona blood agar plate.

Surfactin



Mycosubtilin

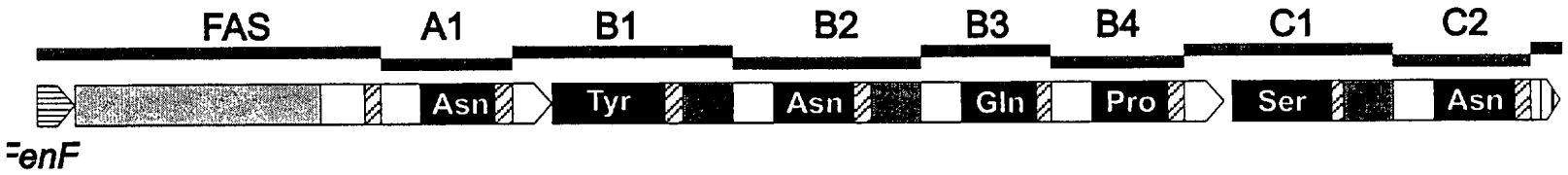
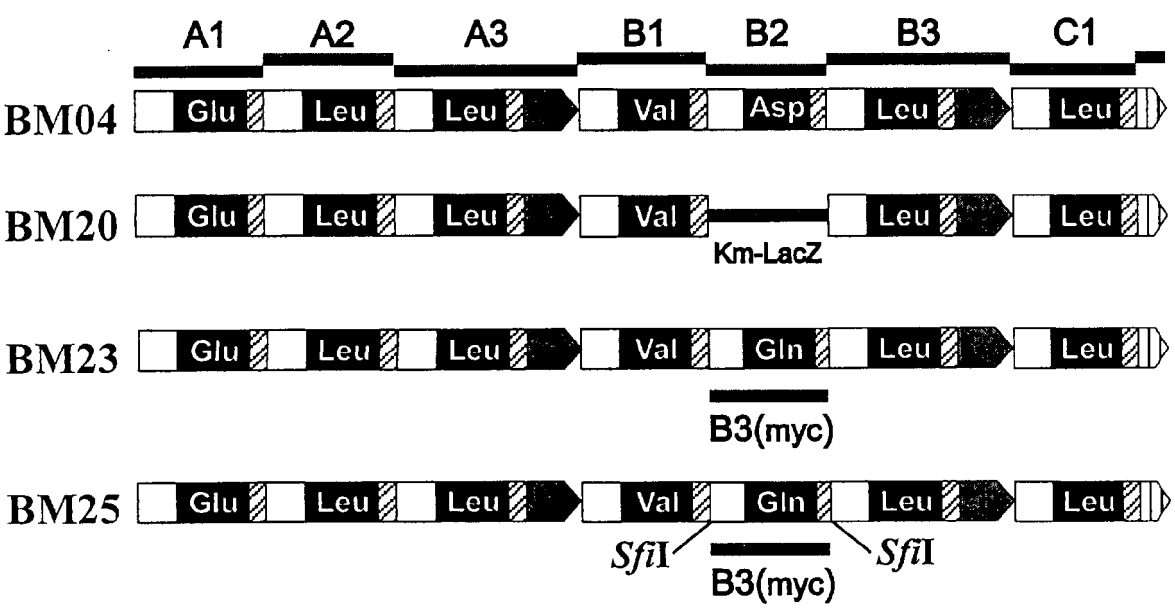


Fig. 12. Replacement of SrfB2 by MycB3 mediated by homologous recombination through SrfB1 and SrfB3 (indicated by the crosses).



	Km	Blue/white colonies	Hemolytic activity
BM04	s	white	+
BM20	r	blue	-
BM23	s	white	+
BM25	s	white	+

Fig. 13. Genetic organisation of BM04, 20, 23 and 25 in the *srf* locus and the phenotypic characteristics of these strains.

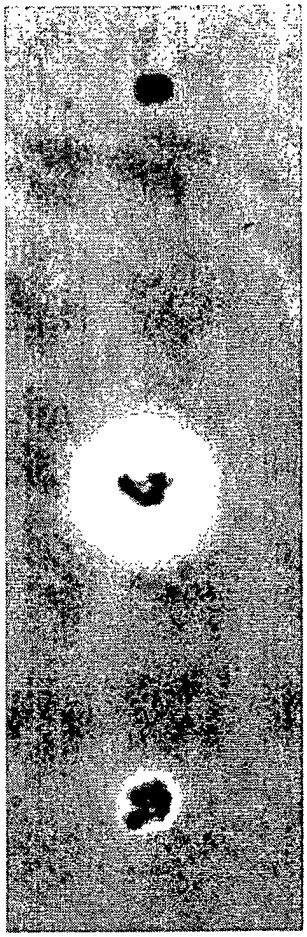
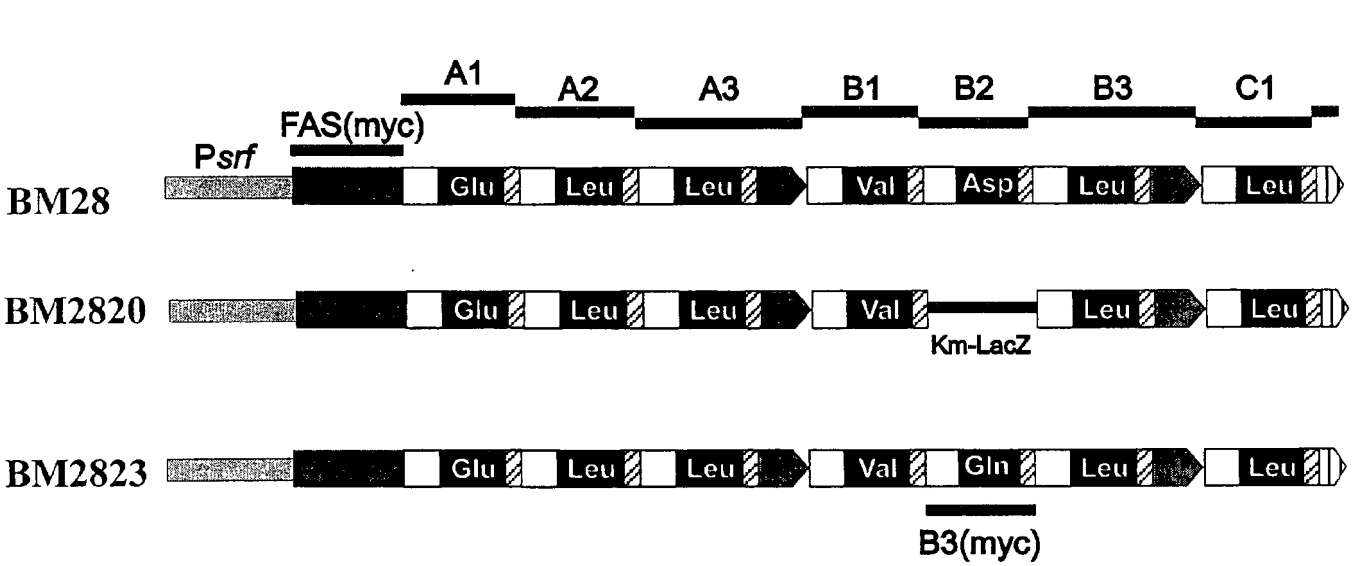


Fig. 14. Hemolytic activities of strains BM04, 20 and 23, determined on blood agar plates.



	Km	Blue/white colonies	Hemolytic activity
BM28	s	white	+
BM2820	r	blue	-
BM2823	s	white	+

Fig. 15. Genetic organisation of BM28, 2820 and 2823 in the *srf* locus and the phenotypic characteristics of these strains.

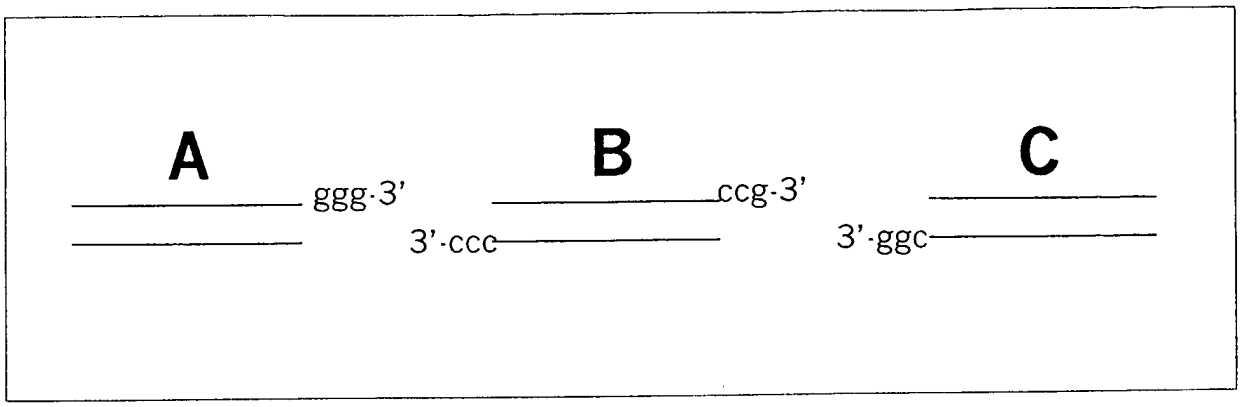


Fig. 16. Ligation of PCR fragments digested with *Sfi*I that generates non-palindromic sticky ends in order to ligate the DNA fragments in a predetermined order. B can only ligate in one orientation to A and C can only ligate in one orientation to B.

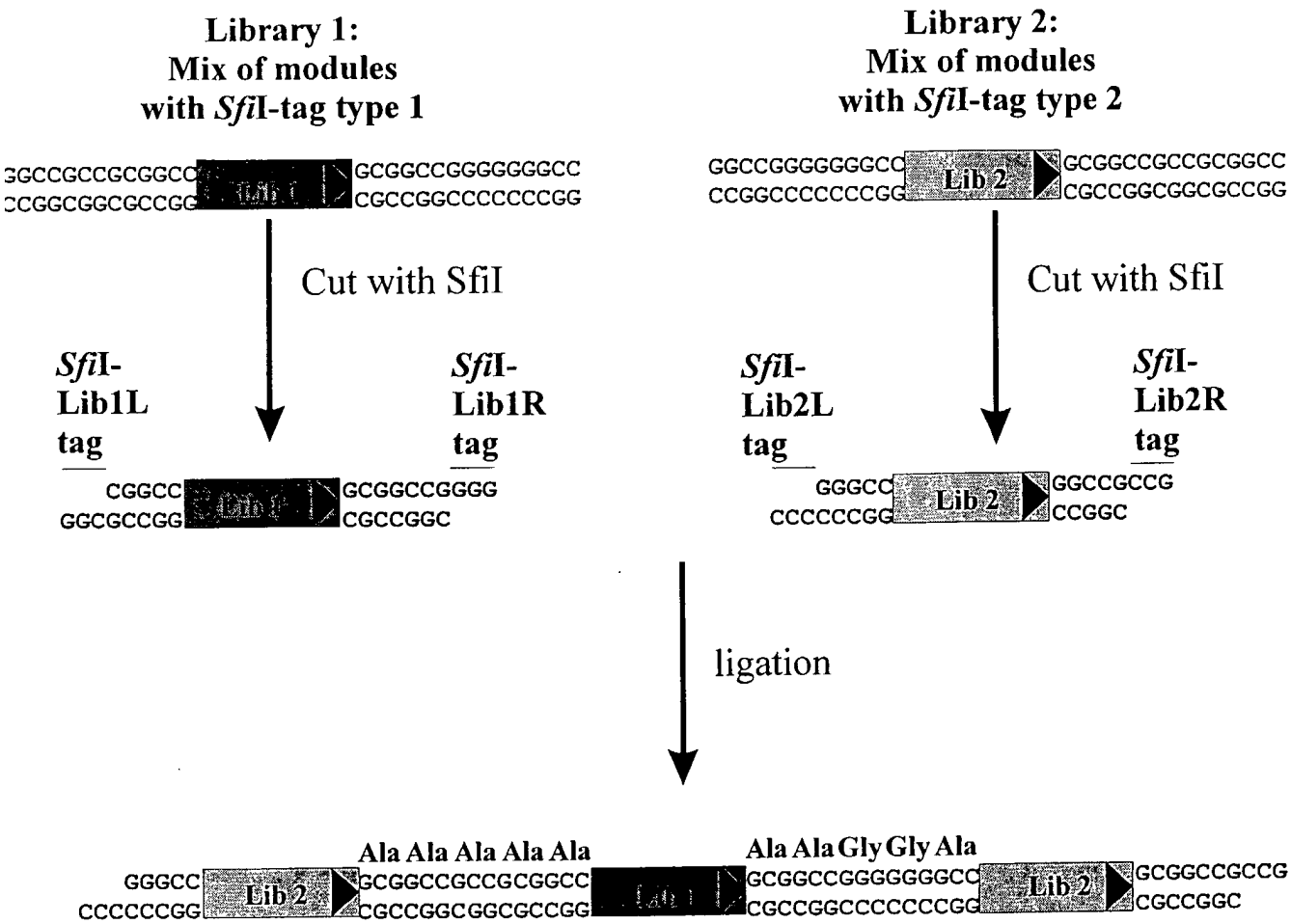


Fig. 17. Deatails of the *Sfi*I tags of library I and II modules. The amino acids that are encoded by these tags after ligation are indicated.

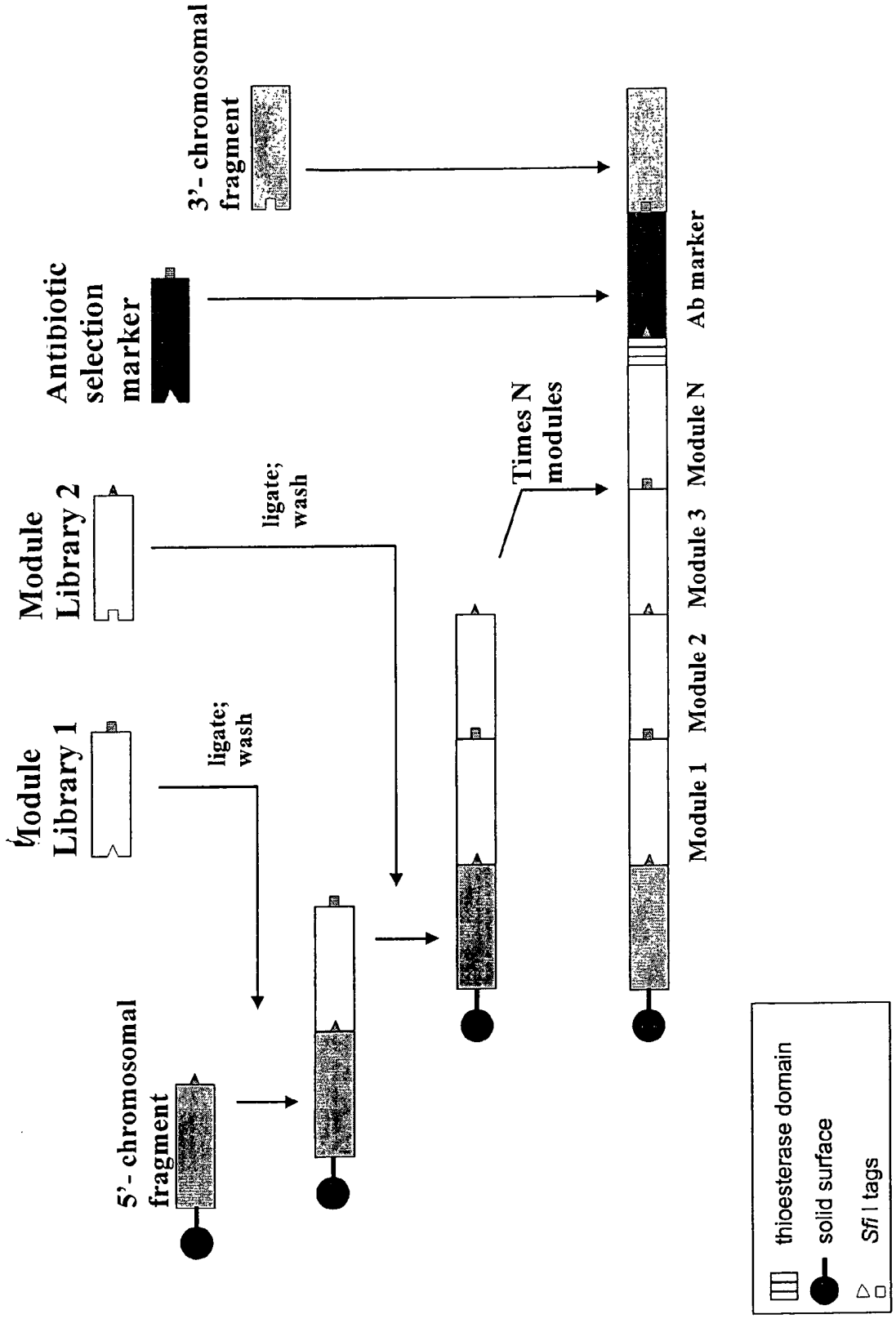


Fig. 18. Solid surface assembly of modules resulting in novel hybrid lipopeptide synthetases.

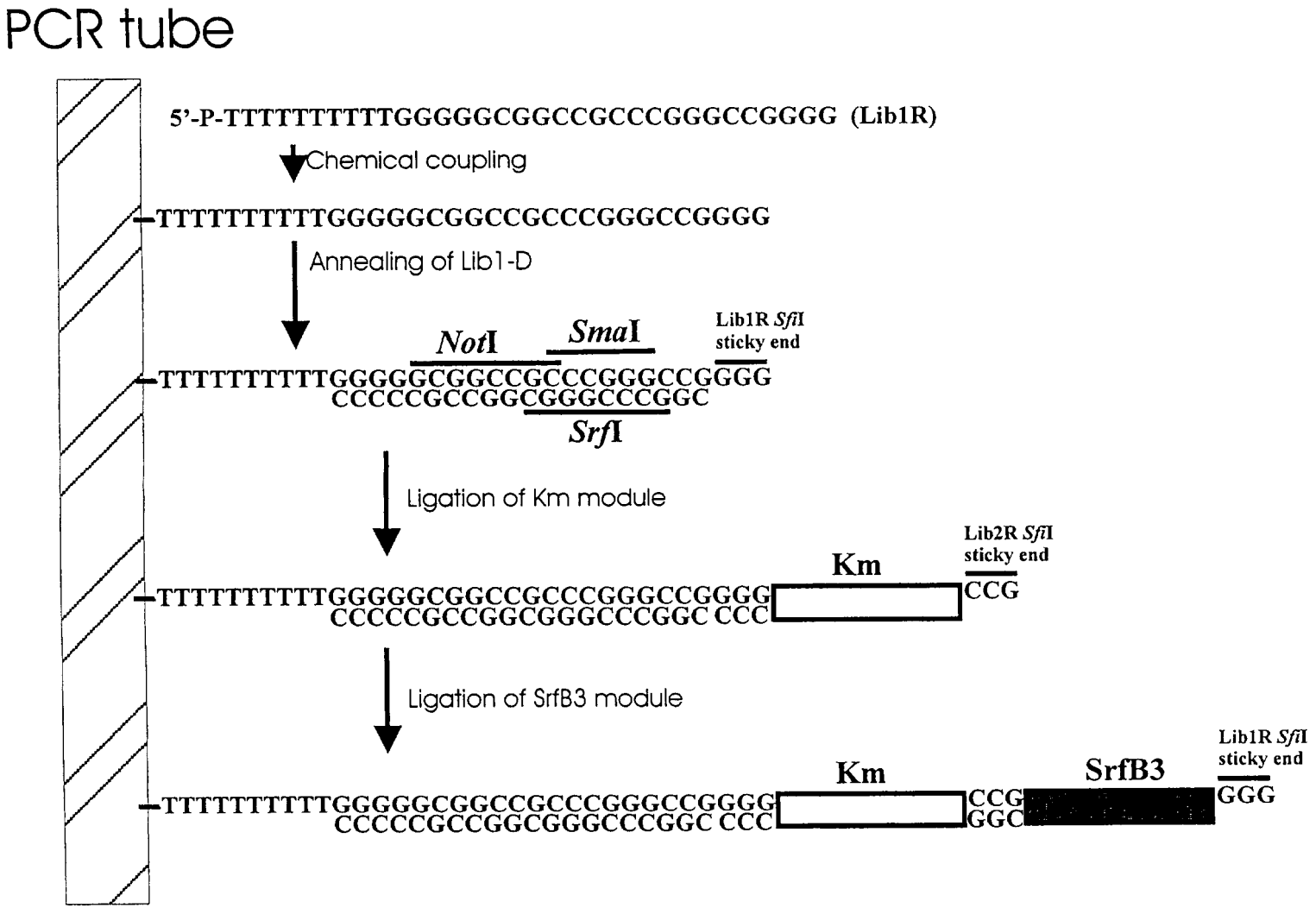


Fig. 19. Chemical coupling of oligo's to NucleoLink PCR tubes and the subsequent ligation of *Sfi*I tagged modules.