Mol Gen Genet (1992) 232:313–321

MGG

Isolation and characterization of *sfp*: a gene that functions in the production of the lipopeptide biosurfactant, surfactin, in *Bacillus subtilis*

Michiko M. Nakano, Nathan Corbell, Jerry Besson, and Peter Zuber

Department of Biochemistry and Molecular Biology, Louisiana State University Medical Center, Shreveport, LA 71130–3932, USA

Received July 10, 1991 / November 17, 1991

Summary. The *sfp* gene is required for cells of *Bacillus subtilis* to become producers of the lipopeptide antibiotic surfactin. *sfp* was isolated and its nucleotide sequence was determined. *sfp* was expressed in *Escherichia coli* and its putative product was purified for use in antibody production and in amino acid sequence analysis. The gene was plasmid-amplified in *B. subtilis*, where it conferred a Srf^+ phenotype on sfp^0 (surfactin non-producing) cells. Overproduction of Sfp in *B. subtilis* did not cause production of an increased amount of surfactin and resulted in the repression of a *lacZ* transcriptional fusion of the *srfA* operon, which encodes enzymes that catalyze surfactin synthesis. We propose that *sfp* represents an essential component of peptide synthesis systems and also plays a role, either directly or indirectly, in the regulation of surfactin biosynthesis gene expression.

Key words: Surfactin – Lipopeptide antibiotic – *Bacillus subtilis* – *sfp* – *srfA*

Introduction

Many microbial species, both prokaryotic and eukaryotic, are known for their ability to synthesize and release into the environment small peptides, some of which are endowed with bioactive properties (Herskowitz 1989; Kurjan and Whiteway 1990; Monaghan and Tkacz 1990; Vining 1990). These include antimicrobial agents, mineral scavengers, and substances that trigger developmental processes in responding microorganisms (Vining 1990). Microbial peptides can be gene-encoded or enzymatically produced by large multifunctional enzyme complexes. Several *Bacillus* species are known to produce peptide antibiotics through the multienzyme thiotemplate mechanism (Katz and Demain 1977; Kleinkauf and von Dohren 1987; Kurahashi 1974; Lipmann 1980). In this process, the constituent amino acids are first activated in the form of amino acyl adenylates and then covalently attached to the reactive centers of the multienzyme complex via a thioester linkage. Peptide bond formation and the movement of the growing peptide chain from one position of the thiotemplate to the next is catalyzed with the aid of a 4′ phosphopantetheine cofactor. Often the pathway concludes with a cyclization that results in the formation of a cyclic polypeptide. Although the biochemistry of peptide synthesis has been described in some detail, genetic characterization of the enzyme complexes and examination of the regulatory aspects of the process are only beginning (Martin and Liras 1989; Nakano and Zuber 1990a).

The spore-forming bacterium *Bacillus subtilis* is an important system for the molecular genetic study of peptide synthesis. In addition to being very amenable to genetic manipulation, it produces a number of peptide secondary metabolites (Katz and Demain 1977). One is the lipopeptide biosurfactant, surfactin (Arima et al. 1968), which is an eight-membered cyclic compound consisting of seven amino acids and a β-hydroxydecanoic acid moiety. Surfactin is synthesized, in part, by the multienzyme thiotemplate mechanism and is produced in stationary phase cultures of *B. subtilis* (Vater 1989). A genetic study has been undertaken to identify the genes required for the production of surfactin (Nakano et al. 1988; Nakano and Zuber 1990b). This has resulted in the characterization of three loci: (i) *srfA* is an operon encoding at least some of the enzymes that catalyze surfactin synthesis (Nakano et al. 1991a). (ii) *srfB* contains the *comP* and *comA* genes which comprise a signal transduction system involved in the competence development pathway and is required for the transcription of *srfA* (Weinrauch et al. 1989, 1990; Nakano and Zuber 1989; Dubnau 1989, 1991; Nakano et al. 1991b). (iii) *sfp* is a gene of unknown function that is required for surfactin production (Nakano and Zuber 1990b). *comP*, *comA*, and the 5′ half of *srfA* are required for competence development (Dubnau 1991; van Sinderen et al. 1990;

Offprint requests to: P. Zuber

Nakano et al. 1991a). *srfA* is also required for efficient sporulation (Nakano et al. 1991a). The involvement of *srfA* in genetic competence and sporulation suggest that there exists a peptide, which is synthesized by a mechanism involving the *srfA* gene products and may function in these developmental processes. Strains that do not produce surfactin possess *srfA* and *srfB*, but lack an intact *sfp* gene. *sfp* is the locus that was transferred to non-producer *B. subtilis cells*, thereby rendering them surfactin positive (Nakano et al. 1988). In this report, the isolation and the characterization of the *sfp* gene and gene product is described.

Materials and methods

Bacterial strains, plasmids, phage, and culture conditions. The *B. subtilis* strains used in this study are listed in Table 1. *Escherichia coli* strains AG1574 (obtained from A. Grossman; Nakano and Zuber 1989) and MV1190 {Δ*(lac-proAB), thi-1 supE,* Δ*(srl recA)306::*Tn*10* (Tet^r)[F′:*traD36, proAB, lacI^q*Δ*M15*]} were used for propagation of plasmid DNA. Plasmids pTV20 and pTV21 were obtained from P. Youngman. Plasmid pMMN13 (Nakano et al. 1991a) is a pGEM4 derivative that contains a *cat* gene (conferring chloramphenicol resistance) and is used as a vector for integrating fragments of DNA into the *B. subtilis* chromosome. The plasmids pMMN19, 20, 21, 24, 25, 26, and 35 are all derivatives of pMMN13 that contain the fragments of *B. subtilis* DNA indicated in Fig. 1.

Plasmid pZΔ328 is a derivative of the *lacZ* fusion vector pZΔ327 (Zuber 1985) which contains a multiple cloning sequence inserted into the unique *Hin*dIII site. It also contains a *cat* gene which is functional in *B. subtilis*.

Table 1. *Bacillus subtilis* strains

Strain	Genotype	Reference
JH642	*trpC2 pheA1*	J. Hoch
OKB105	*pheA1 sfp*	Nakano et al. 1988
ZB307A	SPβ*c2del2*::Tn*917*::pSK10Δ6	Zuber and Losick 1987
OKB168	*pheA1* Tn*917 lac* Ω168	This report
OKB170	*pheA1 sfp* Tn*917 lac* Ω168	This report
LAB224	*pheA1 sfp* Tn*917lac* Ω224	This report
LAB225	*pheA1 sfp* Tn*917lac* Ω225	This report
LAB227	*pheA1 sfp* Tn*917lac* Ω227	This report
LAB358	*pheA1 sfp* SPβ*c2del2*::Tn*917*::pXL5 *(srfA-lacZ)*	This report
LAB364	*trpC2 pheA1* SPβ*c2del2*::Tn*917*::pXL5 *(srfA-lacZ)*	This report
LAB591	*trpC2 pheA1* pSV-sfp	This report
LAB592	*trpC2 pheA1* pNAC3	This report
LAB596	*trpC2 pheA1* pSV-sfp SPβ*c2del2*::Tn*917*::pXL5 *(srfA-lacZ)*	This report
LAB597	*trpC2 pheA1* pNAC3 SPβ*c2del2*::Tn*917*::pXL5 *(srfA-lacZ)*	This report

pXL5 is a pZΔ328 derivative that contains a 3 kb fragment carrying the promoter region and part of the first open reading-frame of the *srfA* operon (Nakano et al. 1991b). All the sequences required for *srfA* transcriptional regulation are contained in pXL5. Plasmid DNA was isolated according to published procedures (Ausubel et al. 1987). Plasmid pBD64 is a multicopy plasmid (copy number approximately 40), which replicates in *B. subtilis* (Gryczan et al. 1980) and contains neomycin and chloramphenicol resistance genes. The SPβ*rpsD-lacZ* fusion was constructed by F. Grundy and T. Henkin using the *Sau*3A-*Hin*dIII fragment of the *rpsD* gene (Grundy and Henkin 1990) containing 56 codons of the amino-terminal end of the coding region of *rpsD* and 134 bp upstream of the *rpsD* start point of transcription.

B. subtilis cells were propagated in 2×YT medium (Nakano et al. 1988) to provide cultures for SPβ lysogenization and to obtain material for chromosomal or plasmid DNA preparations. Neomycin at 5 μg/ml and

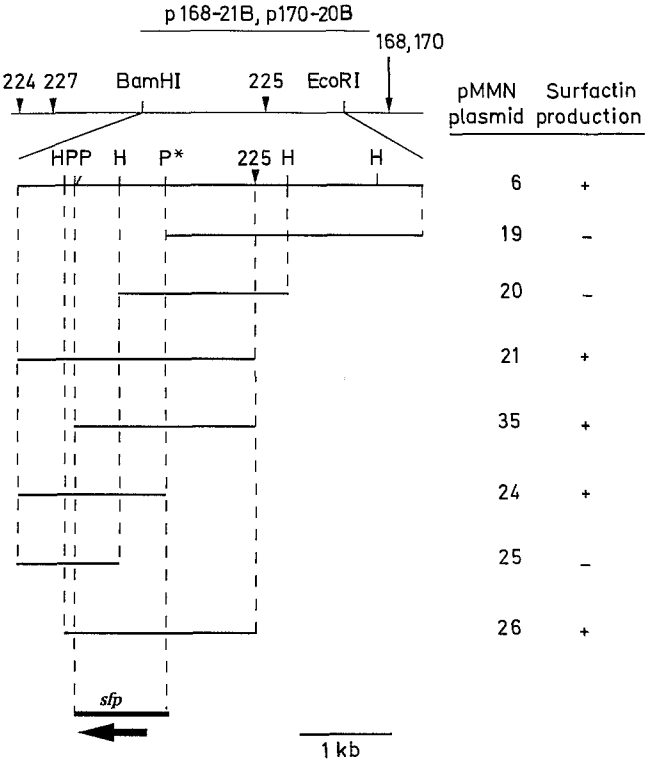

Fig. 1. Physical map of the *sfp* region and localization of the *sfp* gene. Shown at the *top* of the figure is the map of the *sfp* region including the location of the transposon insertions in strains OKB168, OKB170, LAB224, LAB227, and LAB225. The *Bacillus subtilis* genomic DNA contained within plasmids p168–21B and p170–20B is indicated. The results of complementation experiments designed to localize *sfp* are shown. The integrative plasmids indicated in the list beginning to the right of the physical map contain the *B. subtilis* DNA inserts indicated to the left of the list and shown *below* the physical map. Each plasmid was introduced into a derivative of SPβ which can recombine with pBR322 plasmid derivatives. The *arrow* indicates the orientation of the *sfp* open reading frame. The symbols for surfactin production next to each plasmid number are: (+) if the plasmid confers a Srf$^+$ phenotype and (−) if Srf$^-$. H, *Hin*dIII; P, *Pvu*II; the *asterisk* indicates *Pvu*II site present in *sfp* DNA but not in *sfp*^{0}

chloramphenicol at 5 μg/ml were added when selection was applied for plasmid- or phage-encoded antibiotic resistance. Competence medium (Dubnau and Davidoff-Abelson 1971) was used to cultivate cells for β-galactosidase enzyme assays (Nakano et al. 1988) and for obtaining protein samples for Western blot analysis. DS medium (Nakano et al. 1988) was used to cultivate cells for preparation of RNA. Plasmid-bearing *E. coli* cells were grown in 2 × YT medium containing 25 μg/ml ampicillin for plasmid preparation. Surfactin production was assayed on blood agar plates (Nakano et al. 1988).

Transformation and transduction. Preparation of competent *E. coli* cells and their transformation with plasmid DNA was carried out according to published protocols (Ausubel et al. 1987). Preparation of *B. subtilis* competent cells was performed according to the procedure of Niaudet and Ehrlich (1979). The induction and propagation of phage SPβ and its use in specialized transduction experiments have been described previously (Zuber and Losick 1987; Nakano and Zuber 1989). Genetic mapping using phage PBS–1 was performed as previously described (Hoch et al. 1967).

DNA sequencing. sfp and *sfp*^{0} DNAs were cloned into M13 mp18 and M13 mp19 (Messing 1983) and were subjected to nucleotide sequence analysis using the dideoxynucleotide chain-termination procedure (Sanger et al. 1977).

RNA isolation and primer extension. RNA was prepared from OKB105 cells which were grown in DS medium to a density corresponding to Klett 60. RNA was extracted as described by Igo and Losick (1986). For primer extension analysis, the primer 5′-GAGCACATCTCCCAGCAGGG-3′ was used. Primer extension was done as previously described (Nakano et al. 1991b).

Isolation of sfp *gene product and production of anti-Sfp antisera. E. coli* cells bearing pUC8-sfp were cultivated overnight with shaking at 37° C in 3 l of 2 × YT medium containing 25 μg/ml ampicillin. The cells were harvested by centrifugation in Sorvall GS3 rotor at 5000 rpm for 10 min. The cells were suspended in 60 ml/of 20 mM HEPES, pH 8.0, 0.1 M KCl, 1 mM EDTA, 10% glycerol and disrupted by sonication. The lysate was cleared by centrifugation at 15 000 rpm in a Sorvall SA600 rotor and the supernatant was combined with ammonium sulfate to a concentration of 35%. After 2 h on ice, the precipitate was collected by centrifugation at 10 000 rpm for 10 min. The supernatant was subjected to further rounds of ammonium sulfate fractionation at concentrations of 50%, 65%, and 80%. The 80% ammonium sulfate pellet was dissolved in and dialyzed against 20 mM TRIS-HCl, pH 8.0. The solution was applied to a Sephadex G–100 column equilibrated with 20 mM TRIS-HCl, pH 8.0. The fractions enriched for the putative Sfp protein were pooled and applied to a DEAE-Sepharose column. The protein was eluted with a 0.1–1 M KCl gradient. The fractions containing the putative Sfp protein were again pooled and ammonium sulfate was added to 80% final concentration. The solution was stored overnight at 4° C and the precipitate was collected by centrifugation at 15 000 rpm for 15 min in a SA600 rotor. The pellet was dissolved in 20 mM TRIS-HCl, pH 8.0, 0.1 M KCl and stored at 4° C. For further purification, 50–100 mg of protein were applied to a preparative SDS-polyacrylamide gel and the protein was resolved by electrophoresis. A strip of the gel containing the putative Sfp protein was cut from the preparative gel, crushed and combined with incomplete Freund's adjuvant to create an emulsion which was injected intramuscularly into a rabbit. The injection was repeated every 2 weeks for a total of 6 weeks; blood was collected every 2 weeks. Anti-Sfp antibody was detected by Western blot analysis in serum collected 4 weeks after the first injection.

The anti-Sfp antibodies were pre-absorbed with *E. coli* cell extract (Harlow and Lane 1988) and purified by affinity adsorption. Sfp protein was purified as described above and resolved by preparative SDS-polyacrylamide gel electrophoresis. The protein was electrophoretically transferred to a nitrocellulose membrane. Anti-Sfp antiserum was diluted 1/50 in TBST (10 mM TRIS-HCl, pH 8.0, 150 mM NaCl, 0.05% Tween-20) and combined with the filter containing the immobilized Sfp protein. The filter was agitated in the diluted antiserum for 1 h at room temperature. The anti-Sfp antibodies were eluted by successive treatments with acidic glycine and alkaline triethylamine solutions according to published protocols (Harlow and Lane 1988). Western blot analysis was performed using a Genie Electrophoretic Blotter (Idea Scientific) and the Protoblot Western Blot AP system (Promega) according to the protocols provided by the manufacturer. Protein was prepared for N-terminal sequence analysis by electroelution from a preparative SDS-polyacrylamide gel. Sfp protein within a strip of polyacrylamide cut from the preparative gel was electroeluted at 120 V for 2 h into a salt bridge containing 7.5 M ammonium acetate. The solution within the salt bridge was collected and dialyzed against 20 mM TRIS-HCl, pH 8.0. The protein was precipitated by the addition of 4 volumes of acetone.

Results and discussion

Isolation of sfp

A surfactin-producing derivative (OKB105) of wild-type strain JH642 was constructed as reported previously (Nakano et al. 1988). This strain was given the genotypic designation *sfp* (surfactin producing) whereas the non-producing parent is referred to genotypically as *sfp*^{0}. Analysis of data from phage PBS–1-mediated generalized transduction showed that *sfp* was located in the *B. subtilis* genome at a site closely linked to *srfA* (Nakano et al. 1988). Tn*917* transposon insertions linked by transformation to *sfp* were identified by transforming competent cells of strain OKB105 *(sfp)* with chromosomal DNA from a Tn*917lacZ* library (Love et al. 1985) with selection for resistance to erythromycin-lincomycin

(MLS^r) conferred by the *erm* gene of Tn*917*. Transformants were screened for the Srf$^-$(surfactin negative) phenotype which would result if the transposon insertion lay within *sfp* or *srfA*, or was linked by transformation to *sfp*. As the cells used to create the Tn*917lacZ* library were *sfp*^{0}(Srf$^-$), in the latter case, the *sfp*^{0} DNA of the donor would replace the *sfp* locus of the surfactin-producing recipient. Five insertions were identified and all were linked by PBS–1 generalized transduction to the *aroI* and *mtlAB* region of the chromosome where *sfp* had been shown to be located (Nakano et al. 1988). One transposon, when reintrodued into *sfp* cells by transformation, conferred a Srf$^-$ phenotype in 100% of the transformants. This insertion was later found to be located in the *srfA* operon (Nakano et al. 1991a). Transformation of OKB105 cells with DNA from each of the other Tn*917* insertion mutants resulted in transformants, of which 45–80% were Srf$^-$. This indicates that the transposons were located near, but not in, *sfp*.

Transformation of OKB105 cells with DNA from cells of a strain which harbored one of the *sfp*-linked transposons, OKB168 (*sfp*^{0}, Tn*917*ΩOK168, Srf$^-$), resulted in both MLS^r Srf$^+$ (OKB170, *sfp*) and MLS^r Srf$^-$ (same phenotype as the OKB168 donor cells) transformants. In an attempt to isolate *sfp* and *sfp*^{0} DNAs, competent cells of OKB168 and OKB170 were transformed with plasmids pTV20 and pTV21 (Youngman et al. 1984) with selection for chloramphenicol resistance (Cm^r). Recombination between pTV20 and 21 with Tn*917* will result in the insertion of a pBR322 replicon and associated *cat* gene (conferring chloramphenicol resistance) into the middle of the transposon. Chromosomal DNA was purified from both pTV20 and pTV21 integrants and was digested with *Bam*HI, for which there is a single site in Tn*917*lac::pTV20/21, to release a fragment containing the transposon DNA, the pBR322 and *cat* DNA, and the DNA flanking the transposon. These fragments were self-ligated and used to transform *E. coli* with selection for ampicillin resistance (Amp^r). The plasmids p168–21B (containing DNA from the *sfp*^{0} strain) and p170–20B (containing DNA from the *sfp* strain, Fig. 1) thus formed were analyzed for the presence of the *sfp* DNA by using each to transform cells of the non-producing parent strain, JH642, with selection for Cm^r. Among the transformants obtained using p170–20B DNA were several that were Srf$^+$, indicating that this plasmid contained at least part of the *sfp* gene. Transformation of JH642 cells with p168–21B did not give rise to Srf$^+$ transformants.

Restriction enzyme analysis of each plasmid showed that the *sfp* and *sfp*^{0} regions were quite similar in their physical organization, which indicated that no major DNA rearrangements had occurred in the creation of the *sfp* or *sfp*^{0} loci. Some primary structural differences were detected as a *Pvu*II site was identified within the *sfp* region that was not present in the *sfp*^{0} DNA (Fig. 1). The structural authenticity of the *sfp* clones was confirmed by Southern blot hybridization analysis of p170–20B, p168–21B, and chromosomal DNA of both *sfp* and *sfp*^{0} cells (data not shown).

The *sfp* DNA was further localized by a complementation assay using SPβ-mediated specialized transduction (Nakano and Zuber 1989). The plasmid pGEM4 was cleaved with *Eco*RI and *Bam*HI and ligated with the *Eco*RI-*Bam*HI fragment of p170–20B. The resulting chimeric plasmid, pMMN4, was cleaved with *Eco*RI and ligated with an *Eco*RI fragment containing a selectable Cm^r gene. This plasmid, pMMN6, was used to transform competent cells of strain ZB307A, a lysogen of prophage SPβ*c2del2*::Tn*917*::pSK10Δ6 (Zuber and Losick 1987) with selection for Cm^r. The plasmid recombined with the SPβ prophage derivative by virtue of the pBR322 DNA homology present on both the phage and plasmid DNAs (Nakano and Zuber 1989; Zuber and Losick 1987). The Cm^r transformants were used to generate a specialized-transducing phage lysate. Cells of strain JH642 were combined with phage lysate and then applied to medium containing selecting levels of chloramphenicol. All Cm^r lysogens were found to be Srf$^+$ on erythrocyte agar medium, indicating that the *sfp* DNA lies within the *Eco*RI-*Bam*HI sequence of p170–20B and that *sfp* can act *in trans* to promote surfactin production. The possibility that the plasmid carrying *sfp* recombined with the *sfp*^{0} region of the chromosome was excluded by the following experiment. Chromosomal DNA prepared from a Cm^r Srf$^+$ transductant was used to transform JH642 and the transformants were selected for MLS^r which was conferred by the SPβ prophage derivative. If the plasmid recombined with the *sfp*^{0} region of the chromosome to give the Srf$^+$ strain, MLS^r transformants should be Srf$^-$. However, the MLS^r transformants were found to be Srf$^+$ indicating the linkage between MLS^r and *sfp*. Therefore it is clear that the phage carrying *sfp* complemented *in trans*. The complementation also shows that *sfp* is a gene(s) that is separate from the closely linked *srfA* operon.

Additional complementation experiments utilizing SPβ-mediated specialized transduction were performed by first inserting fragments from the *Eco*RI-*Bam*HI region (Fig. 1) into the integration vector pMMN13 (Nakano et al. 1991a). The resulting plasmids (listed in Fig. 1) were integrated into the SPβ-specialized transducing phage by transforming competent cells of ZB307A (Zuber and Losick 1987) with selection for Cm^r. Lysates of SPβ phage bearing each of the *sfp* DNA fragments were used to lysogenize cells of JH642 with selection for Cm^r. The transductants were then tested on erythrocyte agar medium to determine if they were Srf$^+$. In this way, the DNA region containing *sfp* was localized to a 1 kb *Pvu*II fragment (Fig. 1).

Nucleotide sequence of sfp *and* sfp^0

The nucleotide sequence of the *Pvu*II fragment (Fig. 2) showed that the *sfp* DNA contains an open reading frame of 224 amino acids, which is preceded by a sequence homologous to the ribosome-binding sites of *B. subtilis* genes. There are two sequences that resemble rho-independent transcription termination sites; one is located 72 to 106 bp upstream of the start codon of the putative *sfp* open reading frame and the other 117 to 132

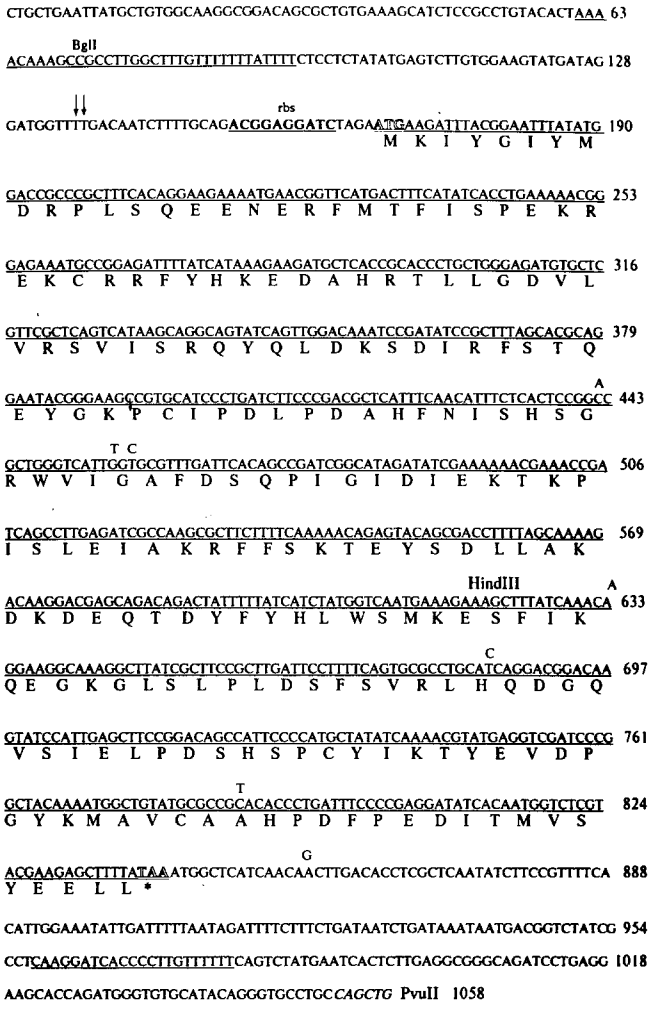

Fig. 2. Nucleotide sequence of *sfp*. The *Bgl*I, *Hin*dIII, and *Pvu*II sites are shown. The sequence of potential rho-independent transcription terminators are *underlined* at positions 61 to 95 and 958 to 979. The putative ribosome-binding site (rbs) is shown in *bold letters* that are also underlined. The letters below the nucleotide sequence indicate the amino acid sequence of Sfp. The transcription initiation sites determined by primer extension analysis are shown by *arrows* above the nucleotide sequence. The letters above the nucleotide sequence indicate the sites of base substitutions that exist in the sfp^0 sequence, except for the A at position 634 which is an insertion

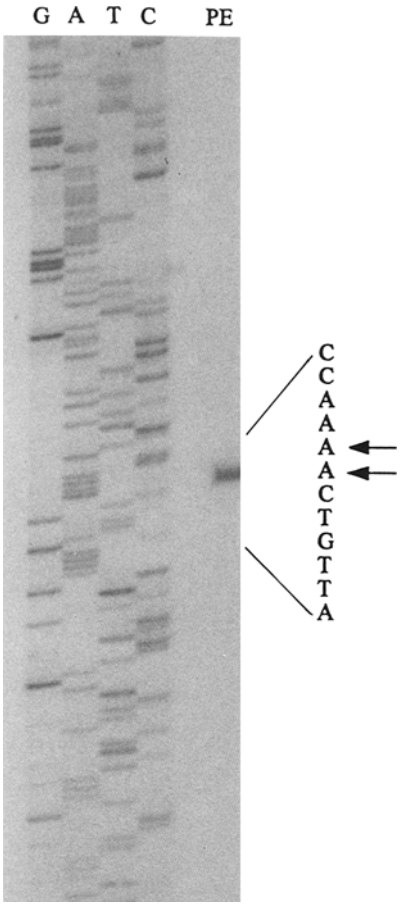

Fig. 3. Determination of the transcriptional start site for *sfp* by primer extension. The sequencing reactions (G, A, T, C) and primer extension reaction (PE) are shown. The nucleotide positions of the start site are marked by *arrows*

bp downstream of the stop codon. Sequence analysis of the sfp^0 region revealed five base substitutions and one base insertion within the putative *sfp* coding region. The insertion creates a frameshift in the coding sequence that could potentially give rise to a truncated product of 165 amino acids. A search of the NBRF–PIR data base for proteins with primary structural similarity identified the *orfX* open reading frame located at the 5′ end of the *grs* operon of *Bacillus brevis* (Kratzschmar et al. 1989) as showing homology to the putative Sfp protein. The *grs* operon encodes the enzymes that catalyze the synthesis of the peptide antibiotic gramicidin S. *orfX* lies immediately upstream of the *grs* operon promoter region and only the 3′ end of *orfX* is so far cloned and sequenced. The determined C-terminal part of OrfX and Sfp have 39% identical residues and 22% conserved residues between amino acid residues 83 to 223 of Sfp (data not shown). The function that *orfX* might perform in the biosynthesis of gramicidin is not known. To begin to understand the functional role of Sfp and OrfX, isolation of the entire *orfX* and comparison of complete amino acid sequences of the two gene products are required.

Transcriptional mapping of sfp

The transcriptional start site for *sfp* was determined by primer extension analysis (see Fig. 3; the approximate start site is marked in Fig. 2). No sequence upstream of the transcriptional start site was identified that resembled a *B. subtilis* promoter. In fact the promoter activity, measured with SPβ-borne *sfp-lacZ*, was very low (data not shown). In contrast to the observed expression of *srfA-lacZ* (Nakano et al. 1988; Nakano and Zuber 1989), the *sfp-lacZ* fusion is not induced in stationary phase cells and is not regulated by the ComP-ComA signal transduction system.

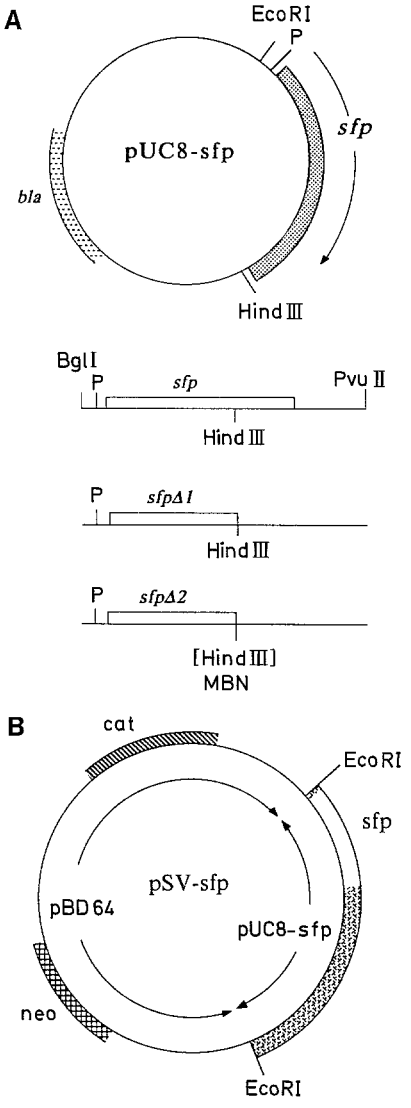

Fig. 4. **A** The structure of the plasmid pUC8-sfp which contains the *sfp* coding region. The maps *below* the plasmid illustration show the structures of the *sfp* DNA within pUC8-sfp and the deletion derivatives that contain the *sfp* alleles *sfpΔ1* and *sfpΔ2*. P indicates the site of the *sfp* promoter. MBN = mung bean nuclease. **B** The structure of plasmid pSV-sfp. Indicated are the sites of the pUC8-sfp and pBD64 moieties. The locations of the *neo* gene (conferring neomycin resistance) and the Cm^r gene *cat* are indicated

Expression of sfp *in* E. coli *and purification of the* sfp *gene product*

A 983 bp *Bgl*I-*Pvu*II fragment of plasmid pMMN24 containing the *sfp* open reading frame (Fig. 2) was inserted in the plasmid vector pUC8 to give plasmid pUC8-sfp (Fig. 4A). The orientation of the fragment was such that the *sfp* coding sequence was placed downstream of the *lac* operon promoter of the plasmid vector such that *sfp* would be under *lacI* control. Cleavage at the *Bgl*I site destroys the putative rho-independent transcription termination site which might impair *lac-sfp* expression. pUC8-sfp was introduced into cells of *E. coli* strain MV1190 by transformation. The resulting plasmid-bearing cells were propagated in 2 × YT medium and their

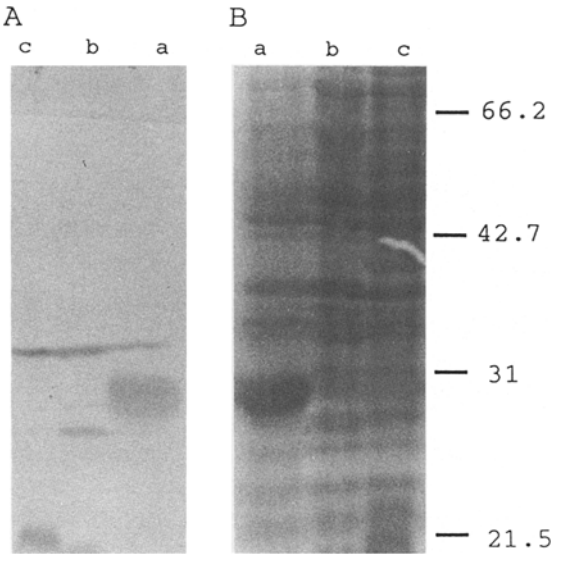

Fig. 5A, B. SDS-polyacrylamide gel electrophoresis (PAGE) profiles of protein from the cells of *Escherichia coli* strains containing pUC8-sfp and *sfpΔ1* and *sfpΔ2* derivatives. **A** Western blot of the gel using the anti-28 kDa Sfp antibody. **B** Coomassie blue-stained 12% SDS-polyacrylamide gel. The numbers on the right indicate the molecular weights of size standards in kilodaltons. Lane a, MV1190 cells carrying pUC8-sfp; lane b, MV1190 cells carrying pUC8-sfpΔ1; lane c, MV1190 cells carrying pUC8-sfpΔ2

proteins were examined by SDS-polyacrylamide gel electrophoresis. Figure 5B shows that cells containing pUC8-sfp produce abundant amounts of a protein of approximately 28 kDa.

A derivative of pUC8-sfp called pUC8-sfpΔ1 was constructed by cleaving the plasmid with *Hin*dIII, and then treating with T4 DNA ligase (Fig. 4A). This resulted in the removal of a fragment that contains 73 codons of the C-terminal end of the coding region of *sfp*. Another derivative (pUC8-sfpΔ2) was constructed, again by cleaving with *Hin*dIII and then treating with mung bean nuclease before ligation (Fig. 4A). This plasmid is the same as pUC8-sfpΔ1 except that there is a frameshift at the junction of the *sfp*-pUC8/*lac* sequence. Cells bearing these plasmids did not produce the 28 kDa protein, and in the case of the pUC8-sfpΔ1-bearing cells, a protein of approximately 26 kDa was produced (Fig. 5B). A truncated product was not observed in Coomassie blue-stained gels of protein from the pUC8-sfpΔ2 cells, but Western blot analysis using antibody raised against the 28 kDa protein revealed a product of 21 kDa (Fig. 5A). There was considerably less of the deletion products than wild-type product in *E. coli* (Fig. 5A). The *sfp* product, since it must perform its function in the proteolytic environment of the stationary phase *B. subtilis* cell (Freese and Heinze 1983), is probably quite resistant to intracellular proteases. This may be the reason for the observed abundance of Sfp produced in *E. coli* cells.

The 28 kDa protein was purified from 3 l of MV1190/pUC8-sfp cells grown in 2 × YT medium as described in Materials and methods. The amino-terminal sequence of the purified protein was determined (Dr. James Carlton, The Core Laboratories of Louisiana State University

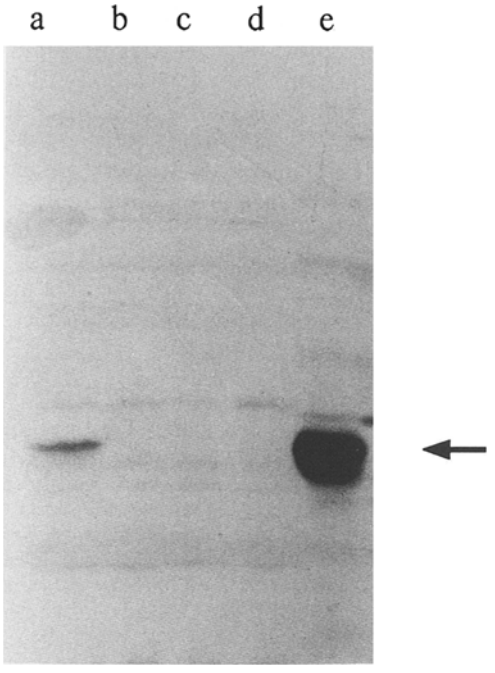

Fig. 6. Overproduction of Sfp in *B. subtilis* cells containing plasmid-amplified *sfp*. The photograph shows a Western blot of protein from the following *B. subtilis* strains: lane a, LAB591 (*sfp*^{0}, pSV-sfp); lane b, LAB592 (*sfp*^{0}, pNAC3); lane c, OKB105 (*sfp*); lane d, JH642 (*sfp*^{0}); and lane e, *E. coli* MV1190 cells containing pUC8-sfp. The *arrow* indicates the 28 kDa product. Protein was resolved on a 12% SDS-polyacrylamide gel

Medical Center, New Orleans) and found to be MKIYGIY × DRPL × Q ... which corresponds to the predicted Sfp N-terminal sequence of the *sfp* open reading frame (Fig. 2). It was thus shown that the 28 kDa protein produced by the pUC8-sfp-bearing cells is indeed the product of the *sfp* open reading frame.

In order to determine if a protein of the same molecular weight as that produced in *E. coli* cells carrying pUC8-sfp could be detected in *sfp* cells of *B. subtilis*, rabbit anti-Sfp antiserum was obtained. The antiserum was preabsorbed with *E. coli* cell extract followed by affinity purification (see Materials and methods). Samples of *sfp* and *sfp*^{0} cell cultures were sonicated and their proteins were resolved by SDS-polyacrylamide gel electrophoresis. Western blot analysis was performed, but no protein was detected that corresponded to the 28 kDa species produced in *E. coli* (Fig. 6, lanes c and d). This may be due to the poor expression of the *sfp* gene, as judged by the activity of the *sfp-lacZ* fusion. An attempt was made to overproduce the Sfp protein in *B. subtilis* by plasmid amplification of the *sfp* gene. This was done by inserting *Eco*RI-cleaved pUC8-sfp DNA into the *Eco*RI site of plasmid pBD64 (Gryczan et al. 1980), a multicopy plasmid (copy number approximately 40) that can replicate in *B. subtilis* cells. The resulting plasmid, pSV-sfp (Fig. 4B), was introduced into cells of strain JH642 with selection for neomycin resistance (Nm^r) conferred by the *neo* gene of pBD64. Another pBD64 derivative, pNAC3, containing the pUC8-sfpΔ2 construct, was also introduced by transformation into JH642 cells. As expected, the pSV-sfp bearing cells produced surfactin as judged by their phenotype on blood agar, but the pNAC3-bearing cells remained Srf$^-$. The plasmid-bearing JH642 cells were grown in competence medium, harvested by centrifugation and lysed by sonication. The proteins of the sonicate were resolved by SDS-polyacrylamide gel electrophoresis and subjected to Western blot analysis using α-Sfp antiserum. A protein of the same apparent molecular weight as the 28 kDa protein produced in *E. coli* cells harboring pUC8-sfp was observed in the JH642/pSV-sfp cells (Fig. 6, lane a). The protein was not detected in JH642 cells bearing pNAC3 (Fig. 6, lane b).

Overproduction of Sfp results in the repression of srfA-lacZ *expression*

Overexpression of Sfp in *B. subtilis* using a multicopy plasmid did not result in an increased amount of surfactin production; instead, thin-layer chromatographic analysis of partially purified surfacin showed that *B. subtilis* cells bearing pSV-sfp produced severalfold less surfactin than did OKB105 cells (data not shown). The effect of *sfp* amplification on surfactin production prompted an examination of *srfA* expression in *B. subtilis* cells carrying pSV-sfp. Cells of strain LAB591 containing the shuttle plasmid pSV-sfp and strain LAB592 containing pNAC3 were lysogenized with SPβ*srfA-lacZ* (SPβ*c2del2*::Tn*917*::pXL5; Nakano et al. 1991b) in order to construct strains that could be used to examine the effect of *sfp* overexpression on the transcription of *srfA*. Cells of the resulting strains, LAB596 (pSV-sfp, SPβ*srfA-lacZ*) and LAB597 (pNAC3, SPβ*srfA-lacZ*), were grown on DS solid medium containing X-gal. Colonies of LAB596 cells were white (Lac$^-$) whereas LAB597 colonies appeared blue (Lac$^+$). Cells of both strains were propagated in competence medium and *srfA*-directed β-galactosidase specific activity was assayed over time. Control cultures contained cells of the plasmid-less strains LAB364 (*sfp*^{0}, SPβ*srfA-lacZ*) and LAB358 (*sfp*, SPβ*srfA-lacZ*).

Figure 7A and B shows the pattern of *srfA-lacZ* expression through exponential and stationary phases of growth. Both *sfp* and *sfp*^{0} cells showed post-exponential induction of *srfA-lacZ* expression whereas *srfA-lacZ* expression in the LAB596 cells was low and poorly induced, suggesting that the expression of plasmid-amplified *sfp* leads to the repression of *srfA* transcription. Expression of *srfA-lacZ* in LAB597 was indistinguishable from that in LAB364, indicating that the plasmid-amplified *sfp* deletion mutant had no effect on the expression of *srfA-lacZ*. A control fusion was used to confirm that the effect of *sfp* overexpression was specific for *srfA-lacZ*. This fusion construct was composed of the 5′ end of the *rpsD* gene, encoding *B. subtilis* ribosomal protein S4 (Grundy and Henkin 1990), fused to the same promoterless *spoVG-lacZ* moiety (Zuber 1985) used to construct the *srfA-lacZ* transcriptional fusions. An SPβ*rpsD-lacZ* lysate was used to lysogenize cells of LAB591 (*sfp*^{0}, pSV-sfp) and LAB592 (*sfp*^{0}, pNAC3). The resulting lysogens were grown in competence

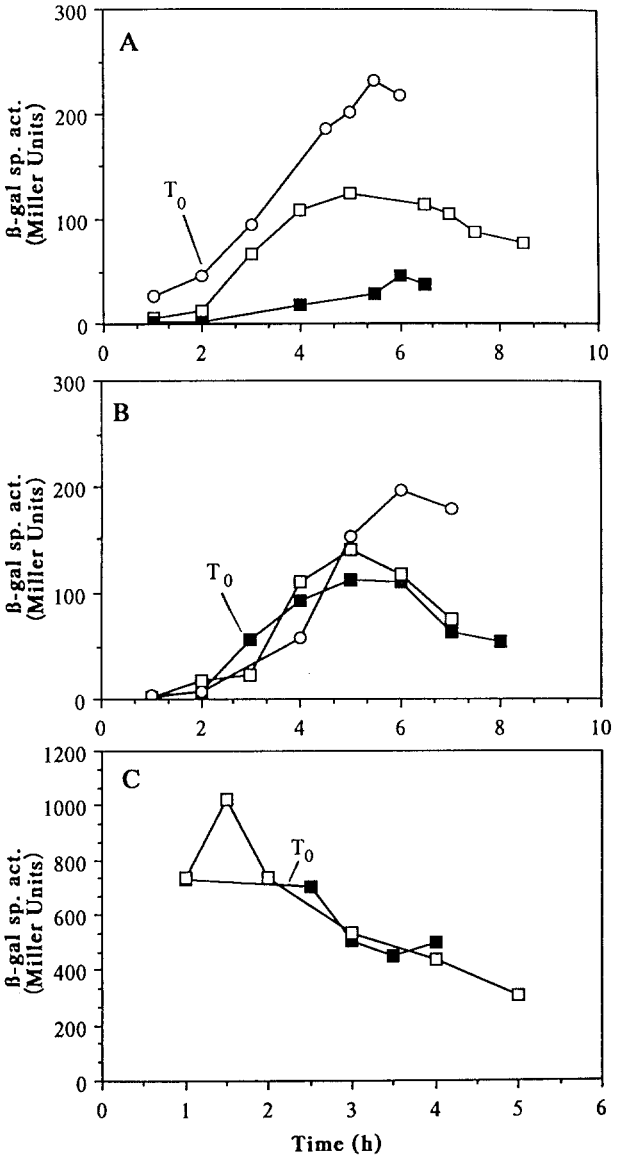

Fig. 7A–C. Expression of *srfA-lacZ* in cells containing plasmid-amplified *sfp*. *srfA*-directed β-galactosidase specific activity was assayed in samples collected at 1 h intervals from cultures of **A, B** SPβ*srfA-lacZ* or **C** SPβ*rpsD-lacZ* lysogens. Cells were grown in competence medium. **A** O, LAB358 *sfp* SPβ*srfA-lacZ*; □, LAB364 *sfp*^{0} SPβ*srfA-lacZ*; ■, LAB596 *sfp*^{0} SPβ*srfA-lacZ*, pSV-sfp. **B** O, LAB358 *sfp* SPβ*srfA-lacZ*; □, LAB364 *sfp*^{0} SPβ*srfA-lacZ*; ■, LAB597 *sfp*^{0} SPβ *srfA-lacZ*, pNAC3 *(sfpΔ2)*. **C** □, *sfp*^{0}/*pSV-sfp*/ SPβ*rpsD-lacZ*; ■, *sfp*^{0}/*pNAC* 3/SPβ*rpsD-lacZ*. T_0 represents the end of exponential growth phase

medium and *rpsD*-directed β-galactosidase specific activity was assayed over time. The level of *rpsD-lacZ* expression in the pSV-sfp cells was the same as that of the cells bearing pNAC3 (Fig. 7C). These data show that the repressive effect of *sfp* overexpression is specific for *srfA-lacZ*.

Among the three identified loci involved in surfactin production, *srfA, srfB,* and *sfp,* only *sfp* has not yet been assigned a functional role. This paper shows that Sfp is homologous to the putative product of *orfX* which is associated with the *grs* operon. By analogy with *sfp, orfX* may encode a product required for the synthesis of gramicidin. *sfp* in single copy, in concert with *srfA,* which encodes the surfactin synthetase enzymes, brings about the production of surfactin. If the *sfp* gene is amplified in order to cause overproduction of the Sfp protein, expression of the *srfA-lacZ* transcriptional fusion is reduced and the preliminary result showed that less surfactin is produced in the *sfp*-overproducing strain. These findings may suggest that *sfp* serves both a regulatory role as well as a more direct role in surfactin synthesis. Further studies of both *orfX* and *sfp* may throw light on the role of this new component in the complex mechanism of peptide synthesis.

Acknowledgements. The *rpsD-lacZ* fusion was provided by Drs. T. Henkin and F. Grundy. We are grateful to Dr. H. Su and Dr. R. Courtney for their assistance in obtaining the α-Sfp antiserum and to Dr. James Carlton of the LSUMC-New Orleans core laboratory facility for sequence analysis of the Sfp protein. Support was provided through PHS grant GM39476 (to P.Z.) from the National Institutes of Health and from a BRSG grant from LSU–MC at Shreveport, Louisiana, USA. The sequence data are available from GenBank/EMBL/DDBJ under accession number X63158.

References

Arima K, Kakinuma A, Tamura G (1968) Surfactin, a crystalline peptidelipid surfactant produced by *Bacillus subtilis:* Isolation, characterization and its inhibition of fibrin clot formation. Biochem Biophys Res Commun 31:488–494

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (1987) Current protocols in molecular biology. Greene Publishing Associates and Wiley-Interscience, NY

Dubnau D (1989) The competence regulon of *Bacillus subtilis.* In: Smith I, Slepecky RA, Setlow P (eds) Regulation of procaryotic development: A structural and functional analysis of bacterial sporulation and germination. American Society for Microbiology, Washington, DC, pp 147–166

Dubnau D (1991) The regulation of genetic competence in *Bacillus subtilis.* Mol Microbiol 5:11–18

Dubnau D, Davidoff-Abelson R (1971) Fate of transforming DNA following uptake by competent *Bacillus subtilis.* I. Formation and properties of the donor-recipient complex. J Mol Biol 56:209–221

Freese E, Heinze J (1983) Metabolic and genetic control of bacterial sporulation. In: Hurst A, Gould GW (eds) The bacterial spore, vol 2. Academic Press, London, pp 101–172

Grundy FJ, Henkin TM (1990) Cloning and analysis of the *Bacillus subtilis rpsD* gene, encoding ribosomal protein S4. J Bacteriol 172:6372–6379

Gryczan T, Shivakumar AG, Dubnau D (1980) Characterization of chimeric plasmid cloning vehicles in *Bacillus subtilis.* J Bacteriol 141:246–253

Harlow E, Lane D (1988) Antibodies. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, pp 313–315

Herskowitz I (1989) A regulatory hierarchy for cell specialization in yeast. Nature 342:749–757

Hoch JA, Barat M, Anagnostopoulos C (1967) Transformation and transduction in recombination-defective mutants of *Bacillus subtilis.* J Bacteriol 93:1925–1937

Igo MM, Losick R (1986) Regulation of a promoter that is utilized by minor forms of RNA polymerase holoenzyme in *Bacillus subtilis.* J Mol Biol 191:615–624

Katz E, Demain AL (1977) The peptide antibiotics of *Bacillus:* chemistry, biogenesis, and possible functions. Bacteriol Rev 41:449–474

Kleinkauf H, von Dohren H (1987) Biosynthesis of peptide antibiotics. Annu Rev Microbiol 41:259–289

Kratzschmar JM, Krause M, Marahiel MA (1989) Gramicidin S biosynthesis operon containing the structural genes *grsA* and *grsB* has an open reading frame encoding a protein homologous to fatty acid thioesterases. J Bacteriol 171:5422–5429

Kurahashi K (1974) Biosynthesis of small peptides. Annu Rev Biochem 43:445–459

Kurjan J, Whiteway M (1990) The *Saccharomyces cerevisiae* pheromone response pathway. Semin Dev Biol 1:151–158

Lipmann F (1980) Bacterial production of antibiotic polypeptides by thiol-linked synthesis on protein template. Adv Microbiol Physiol 21:227–260

Love PE, Lyle MJ, Yasbin RE (1985) DNA-damage-inducible *(din)* loci are transcriptionally activated in competent *Bacillus subtilis*. Proc Natl Acad Sci USA 82:6201–6205

Martin JF, Liras P (1989) Organization and expression of genes involved in the biosynthesis of antibiotics and other secondary metabolites. Annu Rev Microbiol 43:173–206

Messing J (1983) New M13 vectors for cloning. Methods Enzymol 101:20–78

Monaghan RL, Tkacz JS (1990) Bioactive microbial products: focus upon mechanism of action. Annu Rev Microbiol 44:271–301

Nakano MM, Zuber P (1989) Cloning and characterization of *srfB*, a regulatory gene involved in surfactin production and competence in *Bacillus subtilis*. J Bacteriol 171:5347–5353

Nakano MM, Zuber P (1990a) Molecular biology of antibiotic production in *Bacillus*. CRC Crit Rev Biotech 10:223–240

Nakano MM, Zuber P (1990b) Identification of genes required for the biosynthesis of the lipopeptide antibiotic surfactin in *Bacillus subtilis*. In: Hoch J, Ganesan AT (eds) Genetics and biotechnology of bacilli, vol 3. Academic Press, New York, pp 397–405

Nakano MM, Marahiel MA, Zuber P (1988) Identification of a genetic locus required for biosynthesis of the lipopeptide antibiotic surfactin in *Bacillus subtilis*. J Bacteriol 170:5662–5668

Nakano MM, Magnuson R, Myers A, Curry J, Grossman AD, Zuber P (1991a) *srfA* is an operon required for surfactin production, competence development, and efficient sporulation in *Bacillus subtilis*. J Bacteriol 173:1770–1778

Nakano MM, Xia L, Zuber P (1991b) Transcription initiation region of the *srfA* operon which is controlled by the ComP-ComA signal transduction system in *Bacillus subtilis*. J Bacteriol 173:5487–5493

Niaudet B, Ehrlich SD (1979) In vitro genetic labeling of *Bacillus subtilis* cryptic plasmid pHV400. Plasmid 2:48–58

Sanger R, Nicklen W, Coulson AR (1977) DNA sequencing with chain termination inhibitors. Proc Natl Acad Sci USA 74:5463–5467

van Sinderen D, Withoff S, Boels H, Venema G (1990) Isolation and characterization of *comL*, a transcription unit involved in competence development of *Bacillus subtilis*. Mol Gen Genet 224:396–404

Vater J (1989) Lipopeptides, an interesting class of microbial secondary metabolites. In: Schlunegger UP (ed) Biologically active molecules. Springer-Verlag, Berlin Heidelberg, pp 27–37

Vining LC (1990) Functions of secondary metabolites. Annu Rev Microbiol 44:395–427

Weinrauch Y, Guillen N, Dubnau DA (1989). Sequence and transcription mapping of *Bacillus subtilis* competence genes *comB* and *comA*, one of which is related to a family of bacterial regulatory determinants. J Bacteriol 171:5362–5375

Weinrauch Y, Penchev R, Dubnau E, Smith I, Dubnau D (1990) A *Bacillus subtilis* regulatory gene product for genetic competence and sporulation resembles sensor protein members of the bacterial two-component signal-transduction systems. Genes Dev 4:860–872

Youngman P, Perkins JB, Losick R (1984) A novel method for the rapid cloning in *Escherichia coli* of *Bacillus subtilis* chromosomal DNA adjacent to Tn*917* insertions. Mol Gen Genet 195:424–433

Zuber P (1985) Localizing the site of *spo0*-dependent regulation in the *spoVG* promoter of *Bacillus subtilis*. In: Hoch JA, Setlow P (eds) Spores IX. American Society for Microbiology, Washington, DC, pp 149–156

Zuber P, Losick R (1987) Role of AbrB in Spo0A- and Spo0B-dependent utilization of a sporulation promoter in *Bacillus subtilis*. J Bacteriol 169:2223–2230

Communicated by T. Lengeler