Nucleic Acids Research, 1993, Vol. 21, No. 1 93–97

Nucleotide sequence of 5′ portion of *srfA* that contains the region required for competence establishment in *Bacillus subtilus*

Shoichi Fuma, Yoshiyuki Fujishima, Nathan Corbell[1], Cletus D'Souza[1], Michiko M.Nakano[1], Peter Zuber[1,*] and Kunio Yamane

Institute of Biological Sciences, University of Tsukuba, Tsukuba-shi, Ibaraki 305, Japan and [1]Department of Biochemistry and Molecular Biology, Louisiana State University Medical Center, 1501 Kings Highway, Shreveport, LA 71130-3932, USA

Received September 25, 1992; Revised and Accepted November 18, 1992 DDBJ accession no. D13262

ABSTRACT

The nucleotide sequence of the 20,535 base pairs of the 5′ end of the *srfA* operon, containing the region required for competence development, was determined. This included the *srfA* promoter region, the first open reading frame, *srfAA*, encoding surfactin synthetase I and part of the second open reading frame, *srfAB*, encoding surfactin synthetase II. Three amino acid-activating domains characteristic of those found in peptide synthetases could be discerned in both *srfAA* (activating Glu, Leu and D-Leu) and *srfAB* (activating Val, Asp, and D-Leu). The presence of a conserved spacer motif in the amino-terminal end of *srfAA* suggests that the *srfAA* product may not initiate surfactin synthesis. The portion of *srfA* that contains the region required for competence is composed of *srfAA* and the first amino acid-acitivating domain of *srfAB*.

INTRODUCTION

The synthesis of a number of peptide antibiotics in members of the genus *Bacillus* is catalyzed by large multienzyme complexes called peptide synthetases (1, 2). Through what has been called the multienzyme thiotemplate mechanism, the constituent amino acids are activated to form amino acyl-adenylates, covalently attached by thioesterification to the enzyme complex, and then linked together by peptide bond formation facilitated by a 4′-phosphopantetheine cofactor (1, 2). This model proposes that the enzymes are composed of subunits or domains that catalyze amino-acid activation, thioesterification and in some cases racemization of the constituent amino acid from the L to the D form. The isolation and characterization of the genes of several peptide synthetases has confirmed the multidomain nature of enzyme subunits. Each subunit is a large protein usually over 100,000 Daltons in molecular weight and contains 1 or more amino acid-activating domains. The domains are organized in an order that is colinear with the amino acid sequence of the peptide antibiotic (3). Each domain possesses an amino acid sequence motif that has recently been found to be the site of thioesterification (4). These sites contain a conserved serine that is believed to be the site of 4-phosphopantetheine attachment, suggesting that multiple pantetheine cofactors may be involved in peptide synthesis. Analysis of the *grs* operon (5) encoding gramicidin S synthetase suggests the involvement of a thioesterase in the multienzyme thiotemplate mechanism.

The synthesis of the lipopeptide surfactin by *Bacillus subtilis* is catalyzed by the multienzyme thiotemplate mechanism (6, 7). The genes that encode at least part of the multienzyme complex are contained within the *srfA* operon (8, 9). Sequence analysis (8) has identified four open reading frames; three encoding the peptide synthetase subunits making up the seven amino acid activating domains corresponding to the seven amino acids Glu-Leu-(D-Leu)-Val-Asp-(D-Leu)-Leu, and a gene encoding a product with significant sequence similarity to the thioesterase-like product of the *grsT* gene of the *grs* operon (5). Within the 5′ half of *srfA* is a sequence that encodes a product that is required for competence development (10), as insertion mutations within this region were found to render *B.subtilis* Com$^-$(9, 11) and block the transcription of the late competence genes (10). In this report, the nucleotide sequence of this region is presented and the corresponding primary structure of the *srfA* products is compared with other peptide synthetase-like proteins.

MATERIALS AND METHODS

Nucleotide sequence analysis

The sequence analysis at University of Tsukuba, Japan (nucleotide 3,524 to 20,535) was performed using DNA from *B.subtilis* 168*trpC2*. Two *Not* I linking clones, pNEXT13 and pNEXT61 (provided by M.Itaya, 12) located at 414 and 424 kb, respectively, from the *B.subtilis* replication origin were ^{32}P-labeled with the random primer DNA labeling kit (Takara Shuzo Co., Kyoto, Japan). Clones containing *srfA* DNA were isolated by plaque hybridization using ^{32}P-labeled pNEXT13 and

* To whom correspondence should be addressed

pNEXT61 as the probes from a 1 dashII library of *B.subtilis* chromosomal DNA (provided by N.Ogasawara). Four independent clones were isolated. DNA from clone 193 reacted with both pNEXT13 and 61 whereas clones 110 and 180 were detected by hybridization to pNEXT13. Clone 16 was detected by hybridization to pNEXT61. The relative location and rough restriction enzyme maps of these clones are shown in Figure 1. The nucleotide sequence was determined using an automated sequencer.

Nucleotide sequence analysis at Louisiana State University Medical Center-Shreveport (nucleotides 1 to 5032 and 7288 to 11,352) was performed using DNA purified from the 168 derivative JH642. Plasmids containing various *srfA* fragment (pMMN46, pMMN40, pMMN38 and pJC1, 9) were sequenced directly after treatment with T7 gene 6 exonuclease and using oligonucleotides (National Biosciences Inc.) as primers for dideoxynucleotide chain-termination sequencing reactions (13).

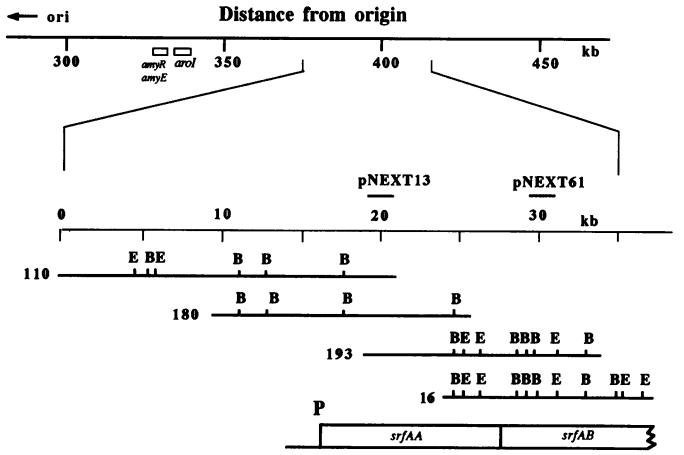

Figure 1. Restriction map of the *B.subtilis* chromosomal DNA from the region containing the part of the *srfA* operon. Shown at the top is a map of the chromosomal region between 300 and 450 kb with respect to the origin of replication of the *B.subtilis* genome. The region around 400 kb is expanded to show a map of the region containing the DNA corresponding to the two linking clones pNEXT13 and pNEXT61 (12). Below this map are the restriction maps of the λ dash clones that were identified by hybridization with pNEXT13 and pNEXT61. The restriction endonuclease sites *Bgl* II (B) and *Eco* RI (E) are indicated. At the bottom of the figure is indicated the approximate location of the *srfAA* and *srfAB* genes.

Alternatively, segments of the plasmid clones were inserted into M13 vectors for sequence analysis.

RESULTS AND DISCUSSION

Isolation of *srfA* DNA

Two molecular cloning projects resulted in the isolation of DNA from the *srfA* operon. One involved the isolation of Tn*917* transposon insertion mutations in *srfA* (14) followed by chromosome walking which involved the sequential 'outcloning' of overlapping *srfA* DNA fragments (9). This was accomplished by constructing plasmid integrants and then cleaving their chromosomal DNA, thereby releasing fragments containing the plasmid and flanking *srfA* DNA. It had been previously shown that the transposon insertion *srfA* ::Tn*917* OK120 was located downstream of the region of *srfA* that was required for the establishment of genetic competence (9). The 5′ region of the *srfA* operon upstream of the transposon was isolated and sequence analysis of this region was initiated.

The other cloning project involved the isolation of λ clones that hybridized to the *Not*I linking plasmids pNEXT13 and pNEXT61 (12). Nucleotide sequence analysis revealed that these clones contained *srfA* DNA. The genomic locations of the pNEXT plasmids and the λ DNA inserts are shown in Figure 1. Judging from the hybridization pattern of the pNEXT13 and 61 clones to the *srfA* DNA, it was concluded that the promoter of *srfA* lies proximal to the origin of replication and that transcription proceeds in the same direction as DNA replication.

Table 1. Discrepancies between 168 trp (Japan) and JH642 (USA) *srfA* 5′ DNA sequences

Nucleotides	Strain	Sequence	Amino Acid
8370 to 8388	168 trp	5′-GCGCGGTGCCTCGGCGTCTTG-3′	ARCLGVL
	JH642	5′-GCCGCGGTGCTCGGCGTCTTG-3′	AAVLGVL
9225 to 9235	168 trp	5′-GAAAGCCCG-3′	ESP
	JH642	5′-GAAACCGCG-3′	ETA
9804 to 9806	168 trp	5′-TGC-3′	C
	JH642	5′-TGG-3′	W
11058 to 11064	168 trp	5′-ACCAGA-3′	TR
	JH642	5′-ACAGAA-3′	TE

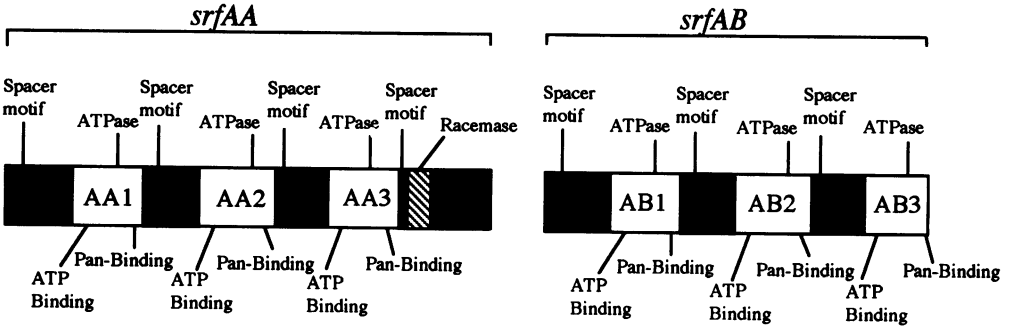

Figure 2. A diagram of the domain organization of the *srfAA* and *srfAB* products. White boxes indicate the amino acid acitivating domains, AA1, AA2, and AA3 in SrfAA; and AB1, AB2, and AB3 in SrfAB. The black boxes represent the non-homologous spacer seqences that separate each amino acid-acitvating domain. The approximate locations of the conserved motifs are indicated.

The competence region of *srfA* encodes products with significant sequence similarity to peptide synthetases

Nucleotide sequence determination of the 5' *srfA* region and *srfA* DNA downstream of Tn*917* OK120 covered an area of 20,535 bp. One large open-reading-frame (Orf) and part of another were found in this region (Fig. 2). The large complete Orf, *srfAA*, encodes a protein of 402, 034 Daltons (D) molecular weight and 3,587 amino acids. Based on the partial sequence of the second Orf, a product of 341, 929 D and 3072 amino acids was identified. Both strains used are derivatives of *B.subtilis* 168. Since the strain used to isolate *srfA* in Louisiana was JH642 and the strain used in Japan was 168trp, there were some discrepancies in the sequence that were not resolved (Table 1).

A search for related proteins of the PIR data base revealed significant sequence similarity with several peptide synthetases, including tyrocidine synthetase 1 (15, 16), gramicidin synthetases 1 and 2 (3, 4, 5), α-aminoadipic acyl-cysteinyl-D-valine (ACV) synthetase (17, 18, 19), enterobactin synthetase subunit EntF (20), and the product of the anguibactin synthesis regulator AngR (21, 22). The characteristic domain structure of peptide synthetases is clearly evident upon examination of the putative *srfAA* and *AB* gene products. Previous work had shown that the first Orf

Table 2. The alignment of conserved motifs among peptide synthetase and peptide synthetase-like proteins.

Enzyme (amino acid)	Spacer motif		ATP-Binding		ATPase		Pan-Binding		Spacer motif COOH-end
SrfAA I (131-1006)	HHVISDGISM	- 460 -	SGTTGRPKGV	- 223 -	YRTGD	- 154 -	FFETGGHSL	- 178 -	
			Domain I						
SrfAA II (1186-2036)	HHLISDGVSI	- 462 -	SGTTGTPKGN	- 215 -	YKTGD	- 152 -	FFDIGGHSL	- 178 -	
			Domain II						
SrfAA III (2224-3004)	HHIISDGASV	- 466 -	SGTTGQPKGV	- 203 -	YRTGD	- 149 -	FFSLGGDSI	- 179 -	HHLVVDGVSW
			Domain III						
SrfAB I (143-1000)	HHIMMDGWSM	- 463 -	SGSTGKPKGV	- 213 -	YRTGD	- 151 -	FFMIGGHSL	- 181 -	
			Domain I						
SrfAB II (1182-2040)	HHIITDGSST	- 455 -	SGTTGKPKGV	- 221 -	YKTGD	- 151 -	FFDLGGHSL		
			Domain II						
SrfABIII (2220-3070)	HHIIADGVSR	- 468 -	SGTTGQPKGV	- 201 -	YRTGD	- 149 -	FFSLGGDSI		?
			Domain III						
GrsA (191-761)			SGTTGNPKGT	- 209 -	YKTGD	- 151 -	FYALGGDSI	- 177 -	HHLVVDGISW
TycA (179-751)			SGTTGKPKGT	- 209 -	YRTGD	- 153 -	FYSLGGDSI	- 177 -	HHLVVDGISW
GrsB I (145-1007)	HHILMDGWCF	- 459 -	SGTTGKPKGV	- 218 -	YRTGD	- 152 -	FFSLGGHSL	- 178 -	
			Domain I						
GrsB II (1186-2043)	HHIISDGVSM	- 461 -	SGSTGKPKGV	- 212 -	YRTGD	- 152 -	FFELGGHSL	- 178 -	
			Domain II						
GrsB III (2221-3089)	HHIISDGVSM	- 461 -	SGTTGKPKGV	- 224 -	YRTGD	- 151 -	FFTIGGHSL	- 178 -	
			Domain III						
GrsB IV (3268-4127)	HHIISDGISS	- 462 -	SGTTGKPKGV	- 214 -	YKTGD	- 152 -	FFELGGHSL		
			Domain IV						
ACVN.lacI (401-821)			SGTTGVPKGV	- 222 -	YRTGD	- 166 -	FFRLGGQSI	- 192 -	
			Domain I						
ACVN.lacII (1014-1897)	HHIILDGWSL	- 476 -	SGTTGKPKAV	- 221 -	YKTGD	- 154 -	FFALGGDSI	- 181 -	
			Domain II						
ACVN.lacIII (2079-3134)	HHSCFDGWSW	- 463 -	SGTTGKPKAV	- 216 -	YRTGD	- 156 -	FFRCGGDSI	- 179 -	HHLVVDTVSW
			Domain III						
EntF (137-1007)	HHKKVDGFSR	- 458 -	SGSTGRPKGV	- 223 -	YRTGD	- 158 -	FFALGGHSL		
AngR (602-1000)			SGSTGTPKGV	- 217 -	YRTGD	- 155 -	FFLSGGDAY		

Indicated at the left are the enzyme amino acid sequences: SrfAA I (surfactin synthetase I, spacer and domain I); SrfAA II (surfactin synthetase I, spacer and domain II); SrfAA III (surfactin synthetase I, spacer and domain III); SrfAB I (surfactin synthetase II, spacer, domain I, and carboxy-terminal end spacer motif); SrfAB II (surfactin synthetase II, spacer and domain II); SrfAB III (surfactin synthetase II, spacer and domain III) the question mark at the end indicates that this sequence was not determined; GrsA (*Bacillus brevis*, gramicidin S synthetase I, 5); TycA (*Bacillus brevis*, tyrocidine synthetase I, 15, 16); GrsB I through IV (*Bacillus brevis*, gramicidin S synthetase II, domains I through IV and associated spacer motifs, 3); ACVN.lacI through III (*Nocardia lactamdurans* a-aminoadipyl-cysteinyl-D-valine synthetase, domains I through III, and associated spacer motifs, 17); EntF (Enterobactin synthetase subunit, 20); AngR (anguibactin regulatory protein, 21, 22). Next to the enzyme sequence names, in parahtheses, are the amino acid regions which contain the conserved motifs. Also shown are motifs which lie within the amino acid-activating domains (underlined); the putative ATP-binding site (23), ATPase motif (3), and the pan-binding sequence (4, 24). The spacer motif is a conserved region within the mostly non-homologous sequence between amino acid-activating domains.

Table 3. Conserved amino acid sequences at the end of the amino acid- activating domains that catalyze the racemization of constituent amino acids from the L to the D form.

SrfAA III (3254-3483)	Spacer motif - 100 -	AYHTEMNKILLTAFGL	- 16 -	EGHGREEIIE	- 5 -	SRTVGWFTSMYPMVLDM - 52 - FNYLGQFD
GrsA (752-981)	Spacer motif - 100 -	AYRTEINDILLTALGF	- 15 -	EGHGREEILE	- 5 -	ARTVGWFTSQYPVVLDM - 52 - FNYLGQFD
TycA (741-970)	Spacer motif - 100 -	AYQTEINDLLLAALGL	- 29 -			ARTVGWFTSQYPVLLDL - 52 - FNYLGQFD
ACVS.lacIII (3125-3329)	Spacer motif - 87 -	AYDTEVNDLLLTATGF	- 16 -	DGHFRELFEG	- 5 -	VRDTVGWFTTMHPFAVE - 41 - FNYLGKLA
ACVA.chrIII (3182-3387)	Spacer motif - 86 -	ALDSGMHEILLMAVGS	- 16 -	EGHGREDTID	- 5 -	SRTVGWFTSMYPFEIPK - 40 - FNYLGRLD

The sequences lie after the carboxy-terminal end spacer motif of SrfAA III, GrsA (5), TycA (15, 16), ACVN.lacIII (17), and ACVA.chrIII (*Acremonium chrysogenum* α-aminoadipyl-cysteinyl-D-valine synthetase, domain III and carboxy terminal region, 18) that could contain the racemase active site. Underlined are highly conserved amino acids.

encodes one of the enzymes that catalyzes surfactin synthesis (9). This enzyme appears to contain three amino acid-activating domains of approximately 800 amino acids each, which catalyze the activation and incorporation of glutamate, leucine, and D-leucine (6, D.Vollenbreich and J.Vater, unpublished). The domains are separated by relatively nonhomologous spacer sequences of 200 amino acids. Three domains can be discerned in the partial sequence of the second Orf *srfAB*. Table 2 shows the alignment of the conserved motifs possessed by *srfAA*, *srfAB*, and other peptide synthetases of known sequence. Shown are the putative ATP-binding (3, 23), ATPase (3), and 4′-phosphopantetheine-binding (pan-binding, 4, 24) motifs along with a conserved sequence (the spacer motif) found in the mostly non-homologous spacer region that lies between the domains (Fig. 2). From this alignment, it is clear that these motifs occupy nearly identical positions in the peptide synthetase sequences. That each domain contains a pan-binding sequence and that this is the site of amino acylation by thioester formation (4) suggests that trans-peptidation is carried out using multiple 4′-phosphopantetheine cofactors, each bound to the conserved serine. This is in contrast to the original model which included a single 4′-phosphopantetheine cofactor serving as a swinging arm that transfers the growing peptide chain from one amino acid position to the next (1, 2). A more recent model imagines a process of passing the growing peptide chain from one 4′-phosphopantetheine to the next during the transpeptidation process (4).

A potentially interesting observation is the position of the spacer motif. Domains that initiate peptide synthesis such as tyrocidine synthetase A, gramicidin S synthetase A , and the first domain of ACV synthetase lack this motif at their amino terminal end. However, the first open-reading frame, *srfAA*, which encodes the enzyme that activates the first amino acid (glutamate) of the peptide portion of surfactin, contains the spacer motif suggesting that it may not initiate surfactin synthesis. One possibility is that an enzyme that activates the β-hydroxy fatty acid moiety initiates the synthesis of surfactin. The component that is responsible for the incorporation of the fatty acid moiety has not been purified and the gene encoding this enzyme subunit has not been identified.

A region is identified which is conserved among the enzymes that catalyze the L to D racemization of a constituent amino acid

A region of extended sequence similarity was found at the carboxy terminal region of the SrfAA (third domain) and the corresponding regions of the enzymes TycA, GrsA, and ACV synthetase third domain. Each of the these enzymes catalyze the racemization of their cognate constituent amino acid from L to the D form; L to D-leu in the case of SrfAA domain 3, L to D-phe by TycA and GrsA, and L to D-val by ACV synthetase domain 3. These putative racemase active sites, which show sequence similarity among these four enzymes are aligned in Table 3. Biochemical analysis of the racemization reaction catalyzed by GrsA suggested that a sulfhydryl was required for the reaction (25), however there is no conserved cysteine within this region. Interestingly, this putative racemization site is found only at the carboxyl end of the peptide synthetase subunits.

The portion of *srfA* which includes the region required for competence development encodes four peptide synthetase domains

The transposon insertion Tn*917*OK120 lies within a *Bgl* II- *Eco* RV fragment extending from nucleotide position 14,982 to 15,237 based on restriction enzyme analysis. This fragment corresponds to the spacer region between the SrfAB domain 1 and SrfAB domain 2. Insertion mutations located upstream from this point confer a competence negative phenotype as well as a loss of surfactin production (9). This would suggest that a peptide, synthesized through the activity of the first four domains of the surfactin synthetase, is required for competence development. A project is underway to determine if the peptide synthesizing activity of *srfA* is required for competence by examining the effect of Ser to Ala mutations within the pan-binding motifs of each domain (26, C. D., M. M., and P. Z. manuscript in preparation). Mutations in the pan-binding sequence of each of the first four domains have little effect on competence development but eliminate surfactin production. This would suggest that the peptide synthesizing capacity of the first four domains of *srfA*, although necessary for surfactin production, is not required for competence development. There exists the possibility that the amino acid adenylating activity is somehow required to regulate competence. A number of adenylating enzymes, such as EntE of *E.coli* which functions in the synthesis of the depsipeptide enterobactin (27), resemble truncated peptide synthetases because they lack the pan-binding sequence but contain the ATP and ATPase motifs (3). Perhaps one or more of the amino acid-activating domains functions in competence as adenylating enzymes. Further mutational analysis is required to determine the role of *srfA* in competence establishment.

ACKNOWLEDGEMENTS

The authors are grateful to M.Itaya for *Not* I linking clones, N.Ogasawara for λ dash library, J.Vater, D.Vollenbreich, and M.A.Marahiel for information and valuable discussion. This research was supported by a grant-in-aid from Ministry of Education, Sciences, and Culture of Japan (to K. Y.); and grants

GM39476 (to P.Z.) and GM45898 (to P.Z.) from the National Institutes of Health, USA. The GenBank/EMBL/DDBJ accession number for the seqeunce presented is D13262.

REFERENCES

1. Kleinkauf, H. and Dohren, H. v. (1990) *Eur. J. Biochem.*, **192**, 1–15.
2. Lipmann, F. 1980. *Adv. Microbial. Phys.*, **21**, 227–260.
3. Turgay, K., Krause, M., and Marahiel, M. A. (1992) *Mol. Micro.* **6**: 529–548.
4. Schlumbohm, W., Stein, T., Ullrich, C., Vater, J., Krause, M., Marahiel, M. A., Kruft, V., and Wittman-Liebold, B. (1991) *J. Biol. Chem.* **266**, 23135–23141.
5. Kratzschmar J., Krause, M., and Marahiel, M. A. (1989) *J. Bacteriol.*, **171**, 5422–5429.
6. Ullich C., Kluge, B., Zbigniew, P., and Vater, J. (1991) *Biochem.* **30**, 6503–6508.
7. Vater, J. (1989) In Schlunegger, U. P. ed. Biologically Active Molecules. Springer-Verlag Berlin Heidelberg, pp. 27–38.
8. Grandi, G., Cosmina, P., Rodriguez, F., De Rerra, F., Perego, M., v. Sinderen, D., and Venema, G. (1992) Eleventh International Spores Conference, Woods Hole, MA.
9. Nakano, M. M., Magnuson, R., Myers, A., Curry, J., Grossman, A. D., and Zuber, P. (1991) *J. Bacteriol.*, **173**: 1770–1778.
10. Dubnau, D. 1991. *Microbiol. Rev.*, **55**, 395–424.
11. Sinderen, D. v., Withoff, S., Boels, H., and Venema, G. (1990) *Mol. Gen. Genet.* **224**, 396–404.
12. Itaya, M., and Tanaka, T. (1991) *J. Mol. Biol.*, **220**, 631–648.
13. Sanger, F., S. Nicklen,and A. R. Coulson. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* , **74**, 5463–5467.
14. Nakano, M. M., Marahiel, M. A., and Zuber, P. (1988) *J. Bacteriol.*, **170**: 5662–5668.
15. Mittenhuber, G., Weckermann, R., and Marahiel, M. A. (1989) *J. Bacteriol.*, **171**, 4881–4887.
16. Weckermann, R., Furbaβ, R., and Marahiel, M. A. (1988) *Nucl. Acids Res.* **16**: 11841.
17. Coque J. J. R., Martin J. F., Calzada J. G., and Liras, P. (1991) *Mol. Micro.*, 5, 1125–1133.
18. Gutierrez S., Diez, B., Montenegro, E., and Martin, J. F. (1991) *J. Bacteriol.*, **173**, 2345–2365.
19. Smith , D. J., Earl, A. F., and Turner, G. (1990) *EMBO J.* **9**: 2743–2750.
20. Rusnak F., Sakaitani, M., Drueckhammer, D., Reichert, J., and Walsh, C. T. (1991) *Biochem.*, **30**, 2916–2927.
21. Farrell, D. H., Mikesell, P., Actis, L. A., and Crosa,J. H. (1990) *Gene* , **86**, 45–51.
22. Koster, W. L., Actis, L. A., Waldbeser, L. S., Tolmasky, M. E., and Crosa, J. H. (1991) *J. Biol. Chem,* . **35**, 23892–23833.
23. Saraste, M., Sibbald, P. R., and Wittinghofer, A. (1990) *Trends Biochem. Sci.* **15**: 430–434.
24. Zuber, P. (1991) *Curr. Op. Cell Biol.* **3**, 1046–1050.
25. Kanda, M., Hori, K., Kurotsu, T., Miura, S., Yamada, Y., and Saito, Y. (1981) *J. Biochem.*, **90**, 765–771.
26. D'Souza, C., Nakano, M. M., and Zuber, P. (1992) Eleventh International Spores Conference, Woods Hole, MA.
27. Staab, J. F., Elkins, M., and Earhart, C. F. (1989) *FEMS Lett* , **59**:15–20.
28. Nakano, M. M., Xia, L., and Zuber, P. (1991) *J. Bacteriol.*, **173**, 5487–5493.