Integration and Mapping of *Bacillus megaterium* Genes Which Code for Small, Acid-Soluble Spore Proteins and Their Protease

MICHAEL D. SUSSMAN,¹ PATRICIA S. VARY,^{2*} CHERYL HARTMAN,² AND PETER SETLOW¹

*Department of Biochemistry, University of Connecticut Health Center, Farmington, Connecticut 06032,¹ and
Department of Biological Sciences, Northern Illinois University, DeKalb, Illinois 60115²*

Received 2 November 1987/Accepted 24 June 1988

Four genes (*ssp* genes) coding for small, acid-soluble spore proteins of *Bacillus megaterium* and the gene for the protease that cleaves them during germination were cloned in the integratable plasmid pJH101. Each plasmid was integrated into the *B. megaterium* chromosome by a Campbell-type mechanism, allowing mapping of all five genes. The gene for the small, acid-soluble spore protein-specific protease (*gpr*) mapped near *rib*, and the *sspA* gene mapped between *argA* and *hisA*. The three other genes of the *ssp* gene family (*sspB*, *-D*, and *-F*) all mapped near *metC/D*, with the order: *sspF-sspD-metC/D-hemA-argO-sspB*. While neither *gpr* nor *sspF* has been mapped in *B. subtilis*, the positions of the *sspA*, *-B*, and *-D* loci are similar in *B. megaterium* and *B. subtilis*, suggesting that the members of this multigene family have not recently undergone significant movement on the chromosome. It appears that more gene rearrangement has occurred in the flanking genes than has occurred in the *ssp* family of genes producing the small, acid-soluble spore proteins.

Approximately 20% of the protein of *Bacillus megaterium* spores is made up of a group of small, acid-soluble spore proteins (SASP) which are synthesized only during sporulation (6). The SASP are degraded early in spore germination, providing amino acids for protein synthesis at this time, and their degradation is initiated by a protease that recognizes a specific amino acid sequence present in all SASP (6). There are two major types of SASP in *B. megaterium*, type A/C and type B, which differ significantly from each other in amino acid sequence (6, 11). While there is only a single type B SASP gene, there are at least seven genes which code for type A/C SASP of extremely similar amino acid sequence, with two of these genes coding for the two major SASP of this type, termed SASP-A and SASP-C in *B. megaterium* (6-8, 11).

The multigene nature of the SASP-A/C gene family as well as a single SASP-B gene have also been found in *B. subtilis*, and to date five SASP genes have been cloned and mapped in this organism (2, 3, 12). Strikingly, all five genes, including four SASP-A/C-type genes, are scattered about the chromosome (12). Major questions concerning these SASP genes, in particular, the multigene family of type A/C, include how and when they arose in evolution and how they became scattered on the chromosome. Data that might lead to answers to these questions include the chromosomal locations of SASP genes in an organism somewhat diverged from *B. subtilis*. Since genes coding for eight SASP (*ssp* genes), as well as the gene for the sequence-specific protease which acts on SASP during spore germination (*gpr* gene), have been cloned from *B. megaterium* (6-9, 11; M. D. Sussman and P. Setlow, unpublished results), it seemed worthwhile to devise methods to map these genes.

Considerable progress has been made in the mapping of the *B. megaterium* chromosome, using phage MP13; to date, approximately 40% of the chromosome has been joined by linkage groups, and a group of strains has been constructed to facilitate mapping (15a). Plasmids can also be introduced into *B. megaterium* by protoplast transformation (13, 16, 17), and plasmids used in *B. subtilis* replicate and express anti-

biotic resistance in *B. megaterium* (13, 15a). Given the successful mapping of cloned *B. subtilis* SASP genes with the integratable plasmid pJH101 (2, 5, 12), it seemed appropriate to determine whether this same plasmid carrying a *B. megaterium* gene would integrate into the *B. megaterium* chromosome at the site of the cloned gene, thus inserting a selectable marker adjacent to the gene of interest.

Consequently, a 0.7-kilobase (kb) *EcoRI-EcoRI* fragment carrying part of the coding sequence of the SASP-A gene (7), a 0.85-kb *Clal-HindIII* fragment containing two-thirds of the coding sequence of the *gpr* gene (Sussman and Setlow, unpublished results), a 0.7-kb *EcoRI-HindIII* fragment containing the entire coding sequence of the SASP-C gene (originally a *BglII-HaeIII* fragment cloned in pUC12 [8]), a 3.2-kb *EcoRI-EcoRI* fragment containing the entire coding sequence of the SASP-C-3 gene (9), or a 0.55-kb *EcoRI-PvuII* fragment containing the entire coding sequence of the SASP-C-4 gene (7) was ligated with appropriately digested pJH101, and the ligation mixtures were used to transform *Escherichia coli* RR101 to drug resistance. Colonies carrying desired inserts were grown and identified, and plasmid DNA was purified as described previously (3). Strain PV361 was used as the *B. megaterium* recipient. The strain is a derivative of strain QM B1551 cured of all seven plasmids by using sublethal doses of ethidium bromide (20 ng/ml), with subsequent freezing and subculturing. Purified plasmid was used to transform *B. megaterium* PV361 by a modification of the procedure of Von Tersch and Carlton (16) in which protoplasts were transformed with 10 µg of plasmid DNA in the presence of polyethylene glycol, plated on nonselective medium, and then incubated overnight. Plates were replica plated onto SNB plates (4) containing chloramphenicol (4 µg/ml) and incubated overnight at 30°C, and drug resistant transformants were single-colony purified twice. The transformation efficiencies of *B. megaterium* PV361 with pJH101 carrying either the SASP genes or the *gpr* gene were low (~1/µg) and were 4 to 5 orders of magnitude lower than parallel transformations with replicative plasmid pHV33 (data not shown). When chromosomal DNA of the wild-type strain was isolated, cut with *Clal*, run on agarose gel electrophoresis, transferred to nitrocellulose paper, and hy-

* Corresponding author.

FIG. 1. Southern blot of chromosomal DNA from *B. megaterium* strains. Chromosomal DNA from the parental *B. megaterium* strain (lane a) or a drug-resistant colony derived by transformation with pJH101 carrying a SASP-A gene fragment (lane b) was digested with *Cla*I, and 3 μ g was run on agarose gel electrophoresis, transferred to nitrocellulose paper, and hybridized as described in the text. The markers numbered 1 to 4 give the positions of DNA standards of 9.7, 4.5, 2.3, and 0.6 kb, respectively.

bridized with a probe made from the 0.7-kb *Eco*RI fragment carrying a part of the coding sequence of the SASP-A gene as described previously (7), only a single 2.9-kb hybridizing fragment was observed (Fig. 1, lane a). In contrast, *Cla*I-digested DNA from a drug-resistant colony transformed with pJH101 carrying the SASP-A gene did not give a 2.9-kb hybridizing fragment, but rather two fragments of 1 and 8 kb (Fig. 1, lane b). This is exactly the result predicted if this pJH101 derivative had integrated into the chromosome by a Campbell-type mechanism at the site of the SASP-A gene (7). Similar analyses of drug-resistant clones derived from transformation with pJH101 derivatives carrying the SASP-C, SASP-C-3, SASP-C-4, or *gpr* gene indicated that these plasmids had also integrated into the chromosome by a Campbell-type mechanism at the site of the cloned gene carried on the transforming plasmid (data not shown). This was also found previously, when pJH101 derivatives were integrated into the *B. subtilis* chromosome (2, 5, 12).

MP13 transducing lysates were prepared (4) on strains carrying integrated pJH101, and the lysates were used to transduce a collection of strains with auxotrophic markers located at key positions around the chromosome (15a). Chloramphenicol-resistant colonies were then scored among prototrophic transductants. The *gpr* gene mapped near *rib*, but distal to *trpB* (Table 1; Fig. 2). The gene coding for SASP-A (now termed *sspA*; see below) was shown by two-factor crosses to lie between *argA* and *hisA* (Table 1; Fig. 2). The placement of this gene between these two loci

joined two previously unlinked linkage groups (1, 10, 15a). The complete sequence of markers in this region is now *azi-hisA-sspA-argA-leuB-leuC-leuA-ilv-4-hem-pheA*. The genes coding for SASP-C, SASP-C-3 and SASP-C-4 (now termed *sspB*, *sspD*, and *sspF*, respectively; see below) mapped near *argO*, a locus previously unmapped on the *B. megaterium* chromosome (Table 1; Fig. 3). No linkage was found between any of these three loci and *pyrB* or *purB*, which are presumed to flank *argO* (15, 15a) (Table 1).

There are two significant findings which can be drawn from the results we presented. The first is that pJH101 can be used in the mapping of cloned *B. megaterium* genes in the same way it has been used in *B. subtilis*. While the number of transformants obtained is low, this probably is due to either the low efficiency of the protoplast transformation procedure, as compared with transformation of competent cells in *B. subtilis*, or the small fragment size of the cloned DNA. However, transformants were obtained, and all were due to Campbell-type integration of pJH101 adjacent to the gene of interest. Given that a number of *B. megaterium* genes without selectable phenotypes have been cloned, pJH101 integration should prove a valuable addition to the genetic methods available in *B. megaterium*.

The second major conclusion concerns the chromosomal location of the genes which code for the various SASP in *B. megaterium*. While the *gpr* gene has not yet been mapped in *B. subtilis*, a number of *B. subtilis* *ssp* genes have been cloned, sequenced, and mapped (2, 3, 12). The *B. subtilis* *sspA* gene codes for one of the two major type A/C SASP of this organism (3). Comparison of the nucleotide sequences of the *B. subtilis* *sspA* and *B. megaterium* SASP-A genes reveals extensive homologies in both the 5' and 3' noncoding regions (Fig. 4) (3, 7), homologies not seen when the SASP-A gene is compared with other *B. subtilis* *ssp* genes. These findings strongly suggest that the *sspA* gene of *B. subtilis* is the homolog of the *B. megaterium* gene which codes for SASP-A. Consequently, we now term the latter gene *sspA*. Strikingly, the *B. megaterium* *sspA* gene maps in the same approximate position as the *sspA* gene of *B. subtilis* (2). This finding indicates that the *sspA* gene has not moved significantly on the chromosome after the divergence of *B. megaterium* and *B. subtilis*.

Comparison of the DNA sequences of other *B. subtilis* *ssp* genes with those of the *B. megaterium* genes coding for SASP-C, -C-3, and -C-4 also indicates that the *B. subtilis* *sspB* gene, which codes for the other major type A/C SASP of *B. subtilis*, is the homolog of the *B. megaterium* SASP-C

FIG. 2. Chromosomal mapping of *gpr* and *sspA* genes in *B. megaterium*. Values given are 100 - cotransduction frequency.

TABLE 1. Transductional mapping of *gpr* and *ssp*^a

Donor	Recipient ^b	Unselected/selected (no. of expt)
PV361::pJH101 <i>gpr</i>	JV109 <i>rib-2</i> PV195 <i>trpB3 pheA2</i> PV167 <i>trpB3 cysB24</i>	Cm ^r /Rib ⁺ 7/632 (4) Cm ^r /Phe ⁺ 0/185 (2) Cm ^r /Trp ⁺ 0/200 (1)
PV361::pJH101 <i>sspA</i>	PV238 <i>argA34 metA3</i> JV94 <i>hisA9 rpsL3</i> PV88 <i>leuC4 pyrB1 rpsL3</i> PV169 <i>leuC21 ilv-4 rpsL3</i>	Cm ^r /Arg ⁺ 165/1,439 (8) Cm ^r /His ⁺ 44/1,541 (8) Cm ^r /Leu ⁺ 0/249 (2) Cm ^r /Leu ⁺ 1/520 (3) Cm ^r /Ilv ⁺ 1/474 (3)
PV361::pJH101 <i>sspB</i>	PV33 <i>argO25</i> PV34 <i>argO26</i> PV148 <i>leuC4 rpsL3 hemA3</i> ^c PV175 <i>metC/D trpB3</i> ^d	Cm ^r /Arg ⁺ 3/200 (1) Cm ^r /Arg ⁺ 7/578 (3) Cm ^r /Hem ⁺ 0/502 (3) Cm ^r /Met ⁺ 0/200 (1)
PV361::pJH101 <i>sspD</i>	PV175 <i>metC/D trpB3</i> PV34 <i>argO26</i> PV90 <i>argO36 leuC4 rpsL3</i> PV88 <i>pyrB1 leuC4 rpsL3</i>	Cm ^r /Met ⁺ 12/600 (3) Cm ^r /Arg ⁺ 0/700 (3) Cm ^r /Arg ⁺ 0/450 (2) Cm ^r /Pyr ⁺ 0/250 (1)
PV361::pJH101 <i>sspF</i>	PV175 <i>metC/D trpB3</i> PV90 <i>argO36 leuC4 rpsL3</i> PV34 <i>argO26</i> PV88 <i>pyrB1 leuC4 rpsL3</i> PV148 <i>hemA3 leuC4 rpsL3</i>	Cm ^r /Met ⁺ 6/528 (3) Cm ^r /Arg ⁺ 0/500 (2) Cm ^r /Arg ⁺ 0/250 (1) Cm ^r /Pyr ⁺ 0/200 (1) Cm ^r /Hem ⁺ 0/250 (1)
PV175 <i>metC/D trpB3</i>	PV34 <i>argO26</i> PV90 <i>argO36 leuC4 rpsL3</i> JV70 <i>purB21</i> PV88 <i>pyrB23 leuC4 rpsL3</i> PV148 <i>hemA3 leuC4 rpsL3</i>	Met ⁻ /Arg ⁺ 0/450 (2) Met ⁻ /Arg ⁺ 9/500 (2) Met ⁻ /Pur ⁺ 0/250 (1) Met ⁻ /Ura ⁺ 0/250 (1) Met ⁻ /Hem ⁺ 12/600 (2)
PV90 <i>argO36</i>	PV148 <i>hemA3 leuC4 rpsL3</i>	Arg ⁻ /Hem ⁺ 75/424 (3)
PV34 <i>argO26</i>	PV148 <i>hemA3 leuC4 rpsL3</i>	Arg ⁻ /Hem ⁺ 16/896 (2)
JV70 <i>purB1</i>	PV34 <i>argO26</i>	Pur ⁻ Arg ⁺ 0/200 (1)

^a Preparation of transducing lysates on donor strains, the transduction procedure, and the selective media were all as described previously (4), except that the minimal MCT medium also contained 6 g of K₂HPO₄.

^b The sources of these strains have been described previously (13).

^c This strain could grow on δ -aminolevulinic acid and was thus designated *hemA*.

^d This strain requires methionine, but not homocysteine or cystathionine, and so is either *metC* or *metD*.

gene, while the *B. subtilis sspD* gene, which codes for a minor type A/C SASP, is the homolog of the *B. megaterium* SASP-C-3 gene (Fig. 4) (3, 8, 9). Consequently, we now term the *B. megaterium* SASP-C and -C-3 genes the *sspB* and *sspD* genes. The *B. megaterium* SASP-C-4 gene is not the homolog of the other cloned *B. subtilis* A/C-type *ssp* gene (termed *sspC* [3, 7]); consequently, we have termed the

SASP-C-4 gene the *sspF* gene. In *B. subtilis*, the *sspB* and *sspD* genes map at 70°C and 115°C on the chromosome, respectively (2), with the *argO* locus at 100°C (15). Strikingly, in *B. megaterium*, the *sspB* and *sspD* loci are also weakly linked to the *argO* locus and on opposite sides (Table 1; Fig. 3). Consequently, if *argO* maps in an equivalent region of the *B. megaterium* chromosome, and the majority

FIG. 3. Chromosomal mapping of *sspB*, *sspD*, and *sspF* genes in *B. megaterium*. Values given are 100 – cotransduction frequency.

B. megaterium SASP-A gene: ATAAATTACCAAAAAATAATGATGATGCCATGTGTTTTGACACAATCTATATTCAACAGGAGGTGAAACACATG
 * * * * *
B. subtilis *sspA* gene: ATACAATTACCAAGTTTTCTGAATGAAGCCATGTGTTTTGACACATTCTATACTC-ACAAGGAGGTGAT-ACACATG
 * * * * *

B. megaterium SASP-C gene: GCAAA-CATTGAGTTACGCCGG(N₇₉)TGGTTCAACTCCAGCTAACTCAACCATTACAA(N₁₁)TAACGAGGAGATGAATACACATG
 * * * * *
B. subtilis *sspB* gene: GCAAAGCATTGGGTTACGCCAGG(N₃₀)TGGTGCAACTCCAGTTAACCACCACTACTAA-----TAAAGGAGATT-TACACATG
 * * * * *

B. megaterium SASP-C3 gene: AAGCTGGGGTTTT--ATT-TTCTCCACAACCATTACCAGTAT(N₅₅)AAAAAGGAGATGACTAATCATG
 * * * * *
B. subtilis *sspD* gene: AGGCCGGGCTTTTTATACTTCTTGCAACAATTGCCAGCAT(N₅₂)AAAAAGGAGATGAAAAAG-ATG
 * * * * *

FIG. 4. Comparison of upstream sequences from various *B. megaterium* and *B. subtilis* SASP genes. Data are taken from references 3, 7, 8, and 9. The underlined ATG residues are the translation initiation codons; asterisks between sequences denote mismatches.

of mapped loci have so far been shown to be roughly similar in location in both species (15a; unpublished data), then the *sspB* and *sspD* genes also map in a similar region in both species. The *sspF* gene of *B. megaterium* also maps close to the *sspD* gene; however, the homolog of the *sspF* gene has not yet been identified in *B. subtilis*. Interestingly, the order of markers around *argO* in *B. megaterium* was *metC/D-hemA-argO*, while in *B. subtilis* it is *metC-argO-metD-hemEFG* (14). The *hem* locus adjacent to *metC/D* and *argO* in *B. megaterium* has been designated *hemA* on the basis of the ability of the *hemA3* mutant to grow on minimal medium containing δ -aminolevulinic acid (14). We have reported previously that two other *hem* mutants of *B. megaterium* have loci that map near *ilv* and *leu* (10), as is the case for *hemA* mutants of *B. subtilis*, but these *B. megaterium* mutants cannot grow on δ -aminolevulinic acid and both produce porphobilogens (10). The gene order in the *argO* region as well as the apparent different location of *hemA* in the two organisms indicate that there has been rearrangement in these regions of the chromosome since the divergence of *B. megaterium* and *B. subtilis*. The mapping of the *sspF-sspB* region constitutes an entirely new linkage group in *B. megaterium* and now increases the regions spanned by linkages to approximately 60% of the chromosome distance.

The mapping of *B. megaterium* SASP genes we present provides further evidence that the type A/C SASP genes have not moved significantly on the chromosome since the divergence of *B. megaterium* and *B. subtilis*. Consequently, if these SASP genes evolved and became scattered on the chromosome by a transposition-like process, this process either was complete prior to the divergence of these two *Bacillus* species or takes place at an extremely slow rate. Clearly, it will be of interest to determine the chromosomal location of other type A/C SASP genes in *B. megaterium* and to compare these results with those in *B. subtilis*.

We are grateful to Wendy Steffenson, Mark Wisniewski, and Robert Matejka for help in mapping and to Alexandra Ho and Susan Goldrick for technical assistance.

This work was supported in part by grants from the National Science Foundation, DMB-8607259 (P.S.V.), and the Public Health Service National Institutes of Health, GM-19698 (P.S.).

LITERATURE CITED

- Callahan, J. P., I. P. Crawford, G. F. Hess, and P. S. Vary. 1983. Cotransductional mapping of the *trp-his* region of *Bacillus megaterium*. J. Bacteriol. 154:1455-1458.
- Connors, M. J., S. Howard, J. A. Hoch, and P. Setlow. 1986. Determination of the chromosomal locations of four *Bacillus subtilis* genes which code for a family of small, acid-soluble spore proteins. J. Bacteriol. 166:412-416.
- Connors, M. J., J. M. Mason, and P. Setlow. 1986. Cloning and nucleotide sequencing of genes for three small, acid-soluble proteins from *Bacillus subtilis* spores. J. Bacteriol. 166:417-425.
- English, J. E., and P. S. Vary. 1986. Isolation of recombination-defective and UV-sensitive mutants of *Bacillus megaterium*. J. Bacteriol. 165:155-160.
- Ferrari, F. A., A. Nguyen, D. Lang, and J. A. Hoch. 1983. Construction and properties of an integrable plasmid for *Bacillus subtilis*. J. Bacteriol. 154:1513-1515.
- Fliss, E. R., M. J. Connors, C. A. Loshon, E. Curiel-Quesada, B. Setlow, and P. Setlow. 1985. Small, acid-soluble spore proteins of *Bacillus*: products of a sporulation-specific, multigene family, p. 60-66. In J. A. Hoch and P. Setlow (ed.), Molecular biology of microbial differentiation. American Society for Microbiology, Washington, D.C.
- Fliss, E. R., C. A. Loshon, and P. Setlow. 1986. Genes for *Bacillus megaterium* small, acid-soluble spore proteins: cloning and nucleotide sequence of three additional genes from this multigene family. J. Bacteriol. 165:467-473.
- Fliss, E. R., and P. Setlow. 1984. The complete nucleotide sequence and the transcriptional and translational start sites of the protein C gene of *Bacillus megaterium*. J. Bacteriol. 158:809-813.
- Fliss, E. R., and P. Setlow. 1984. *Bacillus megaterium* spore protein C-3: nucleotide sequence of its gene and the amino acid sequence at its spore protease cleavage site. Gene 30:167-172.
- Garbe, J. C., G. F. Hess, M. A. Franzen, and P. S. Vary. 1984. Genetics of leucine biosynthesis in *Bacillus megaterium*. J. Bacteriol. 157:454-459.
- Hackett, R. H., B. Setlow, and P. Setlow. 1986. Cloning and nucleotide sequence of the *Bacillus megaterium* gene coding for small, acid-soluble spore protein B. J. Bacteriol. 168:1023-1025.
- Hackett, R. H., and P. Setlow. 1987. Cloning, nucleotide sequencing, and genetic mapping of the gene for small, acid-soluble spore protein γ of *Bacillus subtilis*. J. Bacteriol. 169:1985-1992.
- Kieselburg, M. K., M. Weickert, and P. S. Vary. 1984. Analysis of resident and transformed plasmids in *Bacillus megaterium*. Biotechnology 2:254-259.
- Kiss, I., I. Berek, and G. Ivanovics. 1971. Mapping the delta-aminolevulinic acid synthetase locus in *Bacillus subtilis*. J. Gen. Microbiol. 66:153-159.
- Piggot, P. J., and J. A. Hoch. 1985. Revised linkage map of *Bacillus subtilis*. Microbiol. Rev. 49:158-179.
- Vary, P. S., and Y.-P. Tao. 1988. Development of genetic methods in *Bacillus megaterium*, p. 403-407. In A. T. Ganesan and J. A. Hoch (ed.), Genetics and biotechnology of bacilli, vol. 2. Academic Press, Inc., San Diego.
- Von Tersch, M. A., and B. C. Carlton. 1983. Megacinogenic plasmids of *Bacillus megaterium*. J. Bacteriol. 155:872-877.
- Vorobjeva, I. P., I. A. Khmel, and L. Alföldi. 1980. Transformation of *Bacillus megaterium* protoplasts by plasmid DNA. FEMS Microbiol. Lett. 7:261-263.