JOURNAL OF BACTERIOLOGY, Nov. 1988, p. 5102–5109
0021-9193/88/115102-08$02.00/0

Vol. 170, No. 11

Localization of *Bacillus subtilis sacU*(Hy) Mutations to Two Linked Genes with Similarities to the Conserved Procaryotic Family of Two-Component Signalling Systems

DENNIS J. HENNER,[1]* MARIA YANG,[1] AND EUGENIO FERRARI[2]

Department of Cell Genetics, Genentech, Inc., 460 Point San Bruno Boulevard,[1] and Department of Research, Genencor, Inc.,[2] South San Francisco, California 94080

Received 16 May 1988/Accepted 9 August 1988

Mutations in the *sacU* region have a pleiotropic phenotype. Certain mutations designated *sacU*(Hy), for example, express degradative enzymes at high levels, are able to sporulate in the presence of glucose, have severely reduced transformation efficiencies, and are nonmotile. We isolated and sequenced the *sacU* gene region of *Bacillus subtilis*. Two open reading frames were found in the *sacU* region, and *sacU*(Hy) mutations were localized to both of these open reading frames. The two open reading frames have similarities to two widespread families of proteins that mediate responses to environmental stimuli.

In *Bacillus subtilis*, a variety of mutants have been isolated due to their ability to produce higher levels of particular degradative exoenzymes, such as amylase, proteases, or levansucrase (4, 22, 37). Upon further characterization, a number of these mutant strains were shown to produce higher levels of not only the original exoenzyme of interest but also most of the other exoenzymes examined (4, 21). The majority of the mutations were subsequently mapped to one genetic locus and are most commonly referred to as *sacU*(Hy) (40). These *sacU*(Hy) mutations have a pleiotropic phenotype. They overproduce a large number of enzymes that degrade polymeric substrates; these enzymes include alkaline and neutral protease, α-amylase, β-glucanase, levansucrase, and intracellular serine protease (1, 23, 33). Strains carrying these mutations can also sporulate efficiently in the presence of glucose, which wild-type strains of *B. subtilis* cannot do (23). These strains also lack flagella and are very poorly transformable (23). Other mutations, designated *sacU*, have also been mapped to the same location as the *sacU*(Hy) alleles or to a tightly linked location (22). These mutations almost completely abolish expression of the levansucrase and intracellular serine protease genes but have only a minimal effect on the expression of the other exoenzymes (22, 33).

The mechanism by which the *sacU*(Hy) mutations increase the level of expression of their target genes has been recently investigated. Studies of the levansucrase gene showed that strains carrying a *sacU*(Hy) mutation had an increased level of mRNA, and analysis of the mRNA by S1 analysis showed that the mRNA had the same transcriptional start site in *sacU*$^+$ and *sacU*(Hy) strains (3, 39). A study of the alkaline protease gene showed similar results (12), suggesting that the stimulation of expression of these two genes is by increased transcription of their promoters. Deletion analysis of the alkaline protease and levansucrase promoters suggested that the target site for this stimulation in *sacU*(Hy) mutant strains is approximately 110 nucleotides upstream of the transcriptional start site (12, 13).

Two other classes of regulatory mutants have been isolated which appear to have a phenotype similar to that of strains carrying the *sacU*(Hy) mutations. The *sacQ* and *prtR* genes encode 46- and 60-amino-acid proteins, respectively; overproduction of the *sacQ* or *prtR* polypeptides appears to be directly responsible for the phenotype of exoenzyme overproduction (1, 27, 43). Both of these genes appear to act in a manner similar to that of *sacU*(Hy) mutations, by stimulating transcription of their target genes (39, 42). The site necessary for stimulation of the alkaline protease and levansucrase in *sacQ* mutant strains appears to be the same as that defined for strains carrying *sacU*(Hy) mutations (12, 13); however, the target site for the *prtR* stimulation has not been defined. Which, if any, of these gene products interacts directly with this upstream target site has not yet been determined.

In this paper we detail the isolation and sequence determination of the *sacU* region of the *B. subtilis* chromosome. Two open reading frames (ORFs) that appear to be organized in an operon were identified. Sequence analysis of *sacU*(Hy) mutations showed that missense mutations in either ORF could lead to a SacU(Hy) phenotype. Comparisons of the deduced amino acid sequences of these two proteins showed similarities to two conserved families of bacterial proteins that transmit environmental stimuli. A model of how these genes might function to stimulate transcription of degradative enzymes in *B. subtilis* is presented.

MATERIALS AND METHODS

Strains. The *B. subtilis* strains used in this study are listed in Table 1. *Escherichia coli* SR101 [Δ(*lac-pro*) *supE thi* F' *traD36 proAB lacI*^{q}*Z* Δ*M15 tonA*], a derivative of strain JM101 (48), was used for plasmid constructions and as a host for M13 phages. *E. coli* transformation was performed by the calcium shock procedure (6), and transformants were selected on Luria-Bertini (LB) plates supplemented with carbenicillin at 50 μg/ml. *B. subtilis* was transformed by a previously published procedure (2), with selection on tryptose blood agar bases plates supplemented with chloramphenicol at 5 μg/ml. Screening for the SacU(Hy) phenotype was done by poking colonies onto LB plates supplemented with 1.5% Carnation powdered nonfat milk and 5 μg of chloramphenicol per ml. Screening for the SacU$^-$ phenotype was done in BG4088, which contains a single copy of a *sacB'-'lacZ* fusion (M. Yang and D. Henner, unpublished

* Corresponding author.

TABLE 1. *B. subtilis* + strains[a]

Strain	Genotype
I168	*trpC2*
1A10	*trpC2 thr-5 his-1*
1A72	*mtr-264*
1A95	*leuA8 sacU32*(Hy) *trpC2*
1A199	*leuA8 sacU200*(Hy) *trpC2*
1A201	*leuA8 sacU42*(−) *trpC2*
1A165	*sacU32*(Hy) *trpC2*
1A311	*amyE*$^+$*M amyR2 metB5 purB6 sacU9*(Hy) *str*
1A340	*amyE*$^+$*M amyR1 aroI116 metB5 sacU118*(Hy) *str trpB3*
1A200	*leuA8 sacU100*(Hy) *trpC2*
BG4088	*amyE*::(*sacB'-'lacZ erm*) *trpC2 thr-5 hisA1*
BG3019	*sacU32*(Hy) *trpC2 mot*$^+$

[a] All the strains were obtained from the Bacillus Genetic Stock Center, The Ohio State University, except BG3019 (39) and BG4088. Strain BG4088 is identical to strain BG4024 (39), except that the *erm* gene was used to integrate the single-copy *sacB'-'lacZ* fusion (D. Henner and M. Yang, unpublished data).

data). The SacU$^-$ phenotype could be distinguished from the wild type by the lack of blue color on medium C plates (30) containing 2% sucrose, 50 μg of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) per ml, and 5 μg of chloramphenicol per ml.

Plasmids and plasmid constructions. Plasmids pJH101 (9) and pJM102 are plasmids which cannot replicate in *B. subtilis* but which can integrate into the chromosome via homologous recombination. Plasmid pJM102 is a derivative of pUC18 (48) containing the pC194 *cat* gene subcloned into the unique *Nde*I site (M. Perego and J. A. Hoch, personal communication). Plasmid pEAK1 was constructed by Vicki Singer and is a subclone of plasmid pMG102 (11) in the integrative vector pJH101. The following plasmids are described by the restriction sites at the ends of the insert, and these are located by the nucleotide numbering in Fig. 1. A schematic diagram of the insert contained in each of the described plasmids is shown in Fig. 2 or 3. pDH49 contains the *Eco*RI(2383)-*Sal*I(3073) fragment cloned into the *Eco*RI site of pJM102. pDH51 contains the *Eco*R(1)-*Dra*I(256) fragment cloned into the *Eco*RI and *Hin*cII sites of pJM102. pDH54 contains the *Eco*RI(1)-*Stu*I(1484) fragment cloned into the *Eco*RI and *Hin*cII sites of pJM102. pDH55 contains the *Eco*RI(2383)-*Eco*RV(2137) fragment cloned into the *Eco*RI and *Hin*cII sites of pJM102. pDH59 contains the *Stu*I(1484)-*Eco*RV(2137) fragment cloned into the *Eco*RI and *Hin*cII sites of pJM102. pDH64 contains the *Stu*I(1484)-*Pst*I(1029) fragment cloned into the *Eco*RI and *Hin*cII sites of pJM102. In the above constructions, ends that were not complementary were repaired to flush ends with the Klenow fragment of DNA polymerase I and deoxynucleoside triphosphates before ligation.

To construct plasmids pDH69 and pDH74, chromosomal DNA from strains 1A95 and 1A311 was digested with *Eco*RV and *Hin*dIII, and DNA in the size range of 750 base pairs was eluted from an acrylamide gel and ligated into pJH101. Colony screening was done by hybridization with a synthetic 25-nucleotide oligonucleotide complementary to nucleotides 2390 through 2415. Colonies that hybridized to the probe were screened for the proper restriction map, and one plasmid of the expected structure with an *Eco*RV(2137)-*Hin*dIII(2896) insert was chosen from each DNA source. The entire sequence of the insert of each plasmid was confirmed. pDH69 had the sequence shown in Fig. 1, whereas pDH74 had a single nucleotide change corresponding to the *sacU*(Hy)9 mutation, as discussed in Results. Colony hybridization was done as described by Maniatis et al. (25). The oligonucleotide used as a probe was labeled by phosphorylation with polynucleotide kinase and 5'-[γ^{32}P] dATP (5,000 μCi/mmol; Amersham Corp.).

Plasmids pMY12 and pMY13 contain a 1,600-base-pair EcoRI(1)-*Bam*HI(1615) fragment cloned into the *Eco*RI and *Bam*HI sites of pJM102. The *Bam*HI site at nucleotide 1615 was introduced during a polymerase chain reaction as described below. Plasmids pMY12 and pMY13 were derived from DNA amplified from strains 1A199 and 1A200, respectively. The sequence of the inserted DNA was determined to show that each plasmid differed in only the single nucleotide corresponding to the particular *sacU*(Hy) mutation, as discussed in Results.

DNA manipulations. The procedures used for isolation, analysis, and construction of plasmid DNAs were carried out by standard methods (25). DNA sequencing was conducted by the dideoxy-chain termination method of Sanger et al. (36). Either single-stranded M13 phages (46) or plasmids pUC118 and pUC119 were used to generate templates for sequencing (45). The sequence was determined on both strands for the entire region depicted in Fig. 1, with the exception of the regions from nucleotide 1 through 90 and 2890 through 3078. Polymerase chain reactions (34) were done as suggested by New England BioLabs with a thermostable *Taq* polymerase. Synthetic oligonucleotides of 25 to 30 nucleotides in length were used as primers; the primers included mismatches from the known sequence which would create restriction sites. Approximately 1 μg of chromosomal DNA was used for the input, and 30 rounds of amplification were performed. Temperatures used were 96°C for 60 s, 46°C for 45 s, and 68°C for 315 s. After the amplification, the sample was extracted with an equal volume of phenol-$CHCl_3$ and precipitated with 2 volumes of ethanol. The sample was digested with the appropriate restriction enzymes, and the DNA fragments were purified by acrylamide gel electrophoresis. Each fragment was ligated into appropriately digested pUC118. Amplification of DNA by this method has been shown to lead to a high rate of nucleotide changes (34). To differentiate errors during amplification from a nucleotide change in the input DNA, two templates derived from separate amplification reactions were sequenced for each input DNA. Only a change consistent between the two templates could be derived from the input DNA; as discussed in Results, several of the mutations discovered by the sequence analysis were reintroduced into the chromosome, and their phenotype was confirmed.

Oligonucleotides were provided by the Genentech Organic Synthesis Group. Restriction enzymes, polymerases, and polynucleotides kinase were obtained from commercial sources.

RESULTS

Isolation of the *sacU* gene. Plasmid pEAK1 overlapped a plasmid that had been previously shown to map near the *sacU* region of the *B. subtilis* chromosome (E. Ferrari, unpublished data). Plasmid pEAK1 was transformed into strain BG3019. DNA was extracted from a transformant and used to transform strain I168 to chloramphenicol resistance. Eighty percent of the chloramphenicol-resistant transformants had a SacU(Hy) phenotype, indicating a very close linkage of the insert of plasmid pEAK1 and the *sacU* gene. Lambda phage 1-26 (11) contains an approximately 8-kilo-

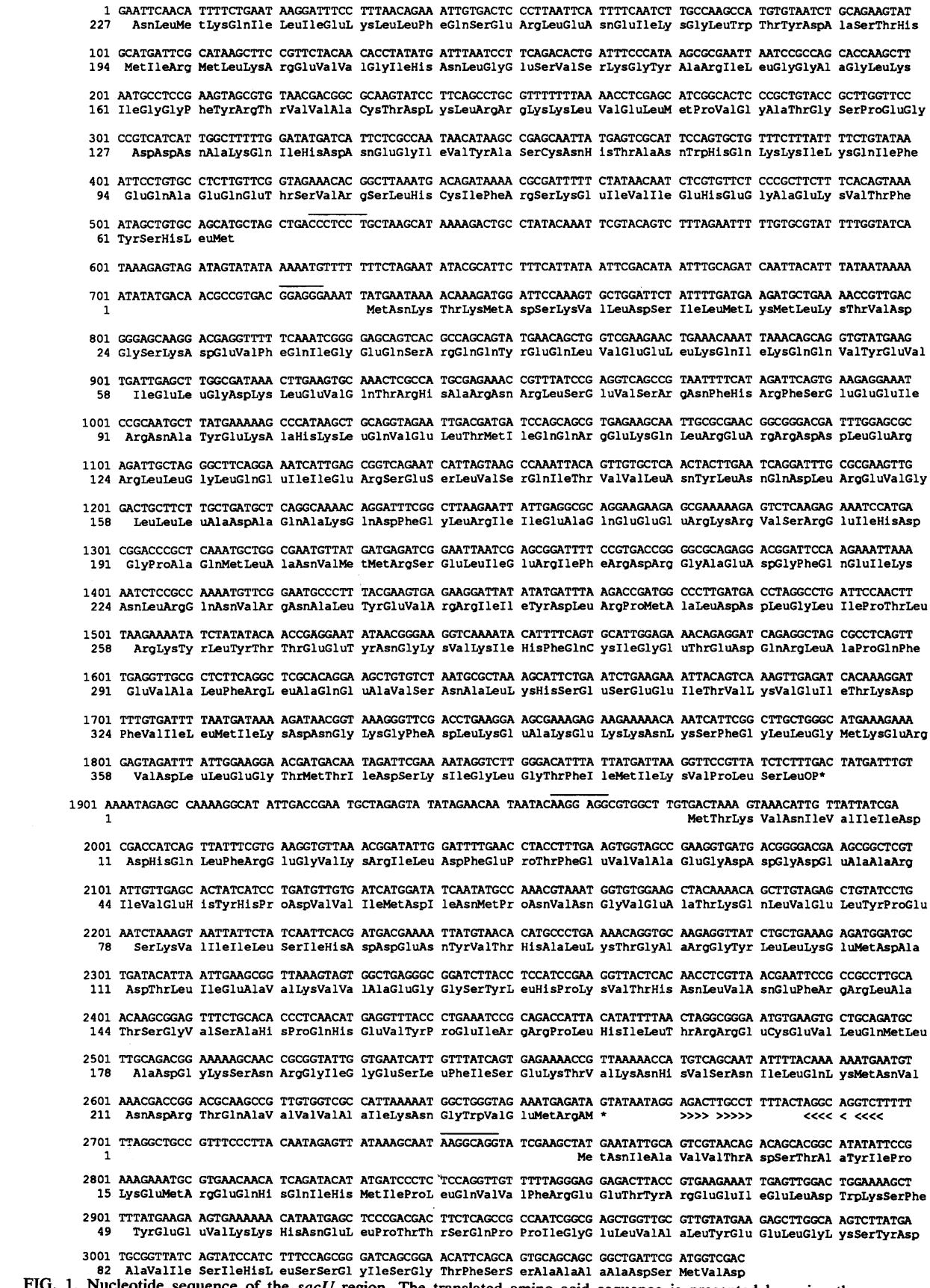

FIG. 1. Nucleotide sequence of the *sacU* region. The translated amino acid sequence is presented by using the presence of likely Shine-Dalgarno sequences and initiation codons to determine the start of open reading frames. ORF-L (nucleotides 1 through 515) reads from the reverse complement of the depicted sequence. Shine-Dalgarno sequences are overlined. A possible rho-independent terminator sequence after *degU* is designated (>><<) below the hairpin structure.

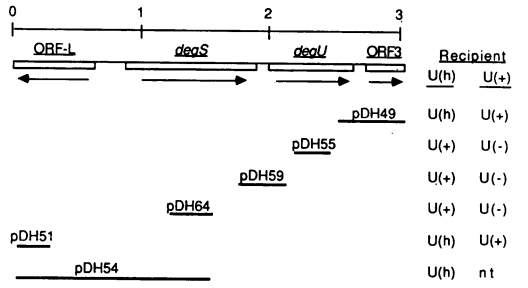

FIG. 2. Phenotype of integration mutations into the *sacU* region. A schematic diagram of the ORFs in the *sacU* region is shown, with a scale in kilobases. The direction of transcription of each ORF is shown by an arrow. Each of the plasmids contains the indicated region cloned into pJM102. The details of the plasmid constructions are described in Materials and Methods. The recipient strains for each plasmid were BG3019 [*sacU32*(Hy)] and BG4088 [*sacU amyE*::(*sacB'-'lacZ erm*)]. The phenotype of BG3019 transformants was determined on LB-skim milk plates. Protease halo size is not a good indicator to differentiate SacU$^+$ from SacU$^-$, and it is not clear whether the true phenotype of the integrations into BG3019 is SacU$^+$ or SacU$^-$. The phenotype of BG4088 transformants was determined on medium C plates containing sucrose and X-gal.

base (kb) insert, and entirely contains the 1.2-kb insert of plasmid pEAK1 in the middle (V. Singer and M. Chamberlin, personal communication). DNA extracted from lambda phage 1-26 was not able to transform strain BG3019 [*sacU32*(Hy)] from *sacU*(Hy) to *sacU*$^+$. The ends of the insert DNA of phage 1-26 were separately mapped with respect to the *sacU32*(Hy) locus (data not shown), and the end that was most closely linked was used to isolate further overlapping lambda phage. Several phage were isolated which could transform BG3019 from *sacU*(Hy) to *sacU*$^+$, indicating that those phage carried DNA extending over the *sacU32*(Hy) mutation. Restriction digestion analysis of those recombinant phage (data not shown) indicated that all of the phage that were able to transform the *sacU32*(Hy) mutation extended past the end of the insert of phage 1-26 and also had in common a 2.3-kb *Eco*RI fragment and a 2.5-kb *Hin*dIII fragment, indicating that the *sacU* gene must lie in this region. One phage containing this region, designated λ-U9, was chosen for further characterization of the *sacU* region.

Sequence of the *sacU* gene region. Approximately 3 kb from the *sacU* region, which included the 2.3-kb *Eco*RI and 2.5-kb *Hin*dIII fragments mentioned above, was sequenced. The sequence showed four potential ORFs (Fig. 1 and 2). ORF-L extends from right to left for 515 nucleotides and is truncated at the end of the determined sequence. The other ORFs all extend from left to right. ORF1 encodes a protein with a translated molecular weight of 45,000. It is followed by 80 untranslated nucleotides and then ORF2, which encodes a protein with a translated molecular weight of 26,000. ORF2 is followed by an untranslated region of 100 nucleotides and then ORF3, which extends past the end of the sequenced region. The start of each ORF was assigned as the furthest upstream potential initiation codon that appeared to be preceded by a reasonable Shine-Dalgarno sequence. A potential rho-independent terminator follows ORF2, suggesting that ORF2 and ORF3 are separate transcription units. We have named the ORF1 and ORF2 genes *degS* and *degU*, respectively, and will refer to them by those names.

Integrative mutations in the *sacU* region. Subclones of DNA fragments from the sequenced region were constructed in the integrative plasmid pJM102, as described in Materials and Methods. These plasmids integrate by Campbell integration via homology of the plasmid insert with the chromosomal DNA. Each plasmid, containing the region shown schematically in Fig. 2, was transformed into strain BG3019 [*sacU32*(Hy)], and the phenotype was determined. Plasmids that do not interrupt either *degS* or *degU*, such as pDH49 and pDH51, had no effect on the SacU(Hy) phenotype (Fig. 2). Similarly, pDH54, which after integration would leave an intact copy of *degS*, has no effect on the SacU(Hy) phenotype. However, plasmids pDH64 and pDH55, which interrupt *degS* or *degU*, abolish the SacU(Hy) phenotype. Similarly plasmid pDH59, which spans *degS* and *degU*, also abolishes the SacU(Hy) phenotype. Our interpretation of these results is that *degU* is necessary for the SacU(Hy) phenotype. Since pDH59 does not interrupt either *degS* or *degU*, yet leads to a loss of the SacU(Hy) phenotype, it is probable that *degU* is transcribed with *degS*. The interruption of *degS* is not informative about whether *degS* is necessary for the SacU(Hy) phenotype, since disruption of the transcription of *degS* could be responsible for the loss of the SacU(Hy) phenotype.

To determine whether the interruption of *degS* or *degU* leads to a SacU$^-$ phenotype, the same plasmids were integrated into strain BG4088 (*sacU*$^+$ *sacB'-'lacZ*). Expression of the *sacB* gene is abolished in SacU$^-$ mutants, and the phenotype can be easily distinguished on plates containing sucrose and X-gal. Transformants with an integrated copy of either plasmid pDH55 or pDH64 no longer express the *sacB'-'lacZ* fusion and appear to have a SacU$^-$ phenotype. These results indicate that a loss of function of *degU*, and possibly *degS*, can lead to a SacU$^-$ phenotype.

Nature of *sacU*(Hy) mutations. To determine the nature of the *sacU*(Hy) mutations, this region of DNA was isolated from several strains carrying various *sacU* alleles by amplification with the polymerase chain reaction technique as described in Materials and Methods. After the amplification, the DNA fragments were subcloned into pUC118 for sequence analysis. Four mutations were located in five mutant backgrounds, two in *degS* and two in *degU*. Strain 1A199 [*sacU200*(Hy)] had a G-to-A change at position 1384, causing a Gly-to-Glu change at amino acid 218 in *degS*. Strain 1A200 [*sacU100*(Hy)] had a G-to-A change at position 1433, causing a Val-to-Met change at amino acid 236 in *degS*. Strains 1A95 and 1A165, which carry the *sacU32*(Hy) mutation, both had an A-to-T change at nucleotide 2006, which causes a His-to-Leu change at amino acid 12 of ORF2. Both strains 1A340 and 1A311, carrying the *sacU118*(Hy) and *sacU9*(Hy) mutations, respectively, had an identical G-to-A change at nucleotide 2290, which causes a Glu-to-Lys change at amino acid 107 in ORF2. No other nucleotide changes were found in the sequenced region (nucleotides 500 to 2670) in these strains.

To test whether these nucleotide changes were sufficient for a SacU(Hy) phenotype, several of the mutations were introduced into the chromosome via integrative plasmids. Plasmids pMY12, pMY13, pDH69, and pDH74 have the structures outlined in Fig. 3, and details of their construction are given in Materials and Methods. Transformants that arise via the integration events depicted in Fig. 3 by crossovers at B or C would have the single-nucleotide change from the plasmid introduced into an intact copy of the gene in the chromosome. Each plasmid was transformed into strain 1A10 (*sacU*$^+$), and the phenotype of the chloramphenicol-resistant transformants was checked on skim milk plates. None of the more than 100 pDH69 (*sacU*$^+$) transform-

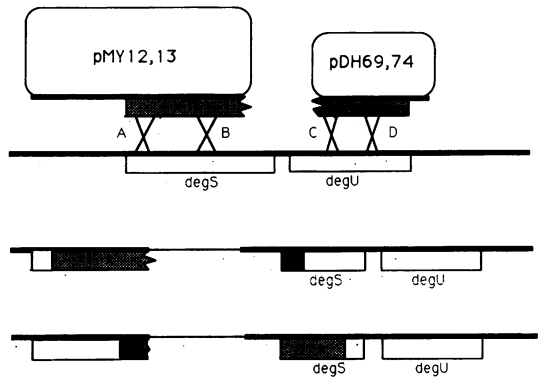

FIG. 3. Introduction of *sacU*(Hy) mutations via integration plasmids. The structures of plasmids pMY12, pMY13, pDH69, and pDH74 are indicated, not to scale. Lines A and B indicate the structures of the chromosomes after integration events A and B above. X indicates the location of a mutation. Symbols: (——) plasmid regions, (▬) chromosomal regions, (▭, ▩) *degS* and *degU* coding regions, (▩) portions derived from the plasmid-borne copy. Jagged lines at the end of the boxed regions indicate a truncated copy of the gene.

ants exhibited an altered phenotype, whereas approximately 50% of the pDH74 [*sacU9*(Hy)] transformants showed a SacU(Hy) phenotype on the skim milk plates. When the same plasmids were transformed into strain BG3019 [*sacU32*(Hy)], all of the pDH69 transformants still showed a SacU(Hy) phenotype. This was expected since the *sacU32* (Hy) mutation lies upstream of the cloned fragment. Unexpectedly, approximately 50% of the pDH74 transformants no longer had a SacU(Hy) phenotype. Presumably integration events that give a protein with both mutations no longer have a SacU(Hy) phenotype. Approximately 15% of the pMY12 [*sacU200*(Hy)] and 8% of the pMY13 [*sacU100*(Hy)] transformants had a SacU(Hy) phenotype. These data indicate that the single nucleotide changes identified for the *sacU9*, *sacU100*, and *sacU200* mutations are sufficient to cause a SacU(Hy) phenotype and that mutations in either *degS* or *degU* can cause such a phenotype.

Similarity of *degS* and *degU* with other proteins. A comparison of the deduced amino acid sequence of *degU* with other protein sequences showed that it is a member of a large family of bacterial proteins. Figure 4 shows a comparison of the N-terminal sequence of *degU* with the N termini of a subset of other members of this previously identified family (32). *degU* has about 25 to 30% identity in pairwise comparisons with the other members of this family, which is about the same level as pairwise comparisons within previously noted members of this family. The C terminus of *degU* showed similarity to the C termini of four other known proteins (Fig. 5). These proteins include the *B. subtilis gerE* protein and three *E. coli* proteins, including the product of the *malT* loci and two ORFs of unknown function in the *uvrC* operon (URF1 and URF2 in Fig. 5). No striking similarities of *degS* with other protein sequences were noted.

DISCUSSION

There are two families of bacterial proteins that make up pairs of proteins that transmit information regarding environmental stimuli to the appropriate transcriptional or enzymatic signals (for a review, see reference 32). The first family of proteins has been termed either the modulator or the sensor family. The modulator proteins have nonconserved N termini that either directly sense environmental information or receive sensory information via additional signal transduction proteins. A loosely conserved C-terminal domain of the modulators is presumably important in transmitting the information to the cognate member of the second family.

```
degU  - - M T K V N I V I I D D H Q L F R E G V K R I L D F E P T F E V V A E G D D G D E A A R I V E H Y
cheY  M A D K E L K F L V V D D F S T M R R I V R N L L K E L G F N N V E - E A E D G V D A L N K L Q A G
ompR  - M Q E N Y K N L V V D D D M R L R A L L E R Y L T E Q G F Q V R S V A N A E Q M D R L L T R E S F
phoB  - - - M A R R I L V V E D E A P I R E M V C F V L E Q N G F Q P V E A E D Y D S A V N Q L N E P W P
ntrC  - - M Q R G I A W I V D D D S S I R W V L E R A L T G A G L S C T - - T F E S G N E V L D A L T T K
spo0A - - M E K I K V C V A D D N R E L V S L L S E Y I E G Q E D M E V I G V A Y N G Q E C L S L F K E K
spo0F - - M M N E K I L I V D D Q Y G I R I L L N E V F N K E G Y Q T F - - Q A A N G L Q A L D I V T K E
sfrA  - - M Q T P H I L I V E D E L V T R N T L K S I F E A E G Y D V F - - E A T D G A E M H Q I L S E Y

degU  H P D V V I M D I N M P N V N G V E A T K Q L V - - E L Y P E S K V I I L S I H D D E N Y V T H A L
cheY  G Y G F V I S D W N M P N M D G L E L L K T I R A D G A M S A L P V L M V T A E A K K E N I I A A A
ompR  H L M V L - - D L M L P G E D G L S I C R R L R - - S Q S N P M P I I M V T A K G E E V D R I V G L
phoB  D L I L L - - D W M L P G G S G I Q F I K H L K R E S M T R D I P V V M L T A R G E E E D R V R G L
ntrC  T P D V L L S D I R M P G M D G L A L L K Q I K - - Q R H P M L P V I I M T A H S D L D A A V S A Y
spo0A D P D V L V L D I I M P H L D G L A V L E R L R E S D L K K Q P N V I M L T A F G Q E D V T K K A V
spo0F R P D L V L L D M K I P G M D G I E I L K R M K - - V I D E N I R V I I M T A Y G E L D M I Q E S K
sfrA  D I N L V I M D I N L P G K N G L L L A R E L R - - - E Q A N V A L M F L T G R D N E V D K I L G L

degU  K T G A R G Y L L K E M D A D T L I E A V K V V A E G G S Y
cheY  Q A G A S G Y V V K P F T A A T L E E K L N K I F E K L G M
ompR  E I G A D D Y I P K P F N P R E L L A R I R A V L R R Q A N
phoB  E T G A D D Y I T K P F S P K E L V A R I K A V M R R I S P
ntrC  Q Q G A F D Y L P K P F D I D E A V A L V D R A I S H Y Q E
spo0A D L G A S Y F I L K P F D M E N L V G H I R Q V S G N A S S
spo0F E L G A L T H F A K P F D I D E I R D A V K K Y L P L K S N
sfrA  E I G A D D Y I T K P F N P R E L T I R A R N L L S R T M N
```

FIG. 4. Similarity of *degU* with the effector family of proteins. The amino-terminal regions of the indicated proteins were aligned by hand. Residues that are identical in at least four of the eight aligned proteins are boxed. The references for the protein sequences used in this comparison are as follows: *cheY* (26), *ompR* (46), *phoB* (24), *ntrC* (5), *spo0A* (10), *spo0F* (44), and *sfrA* (8).

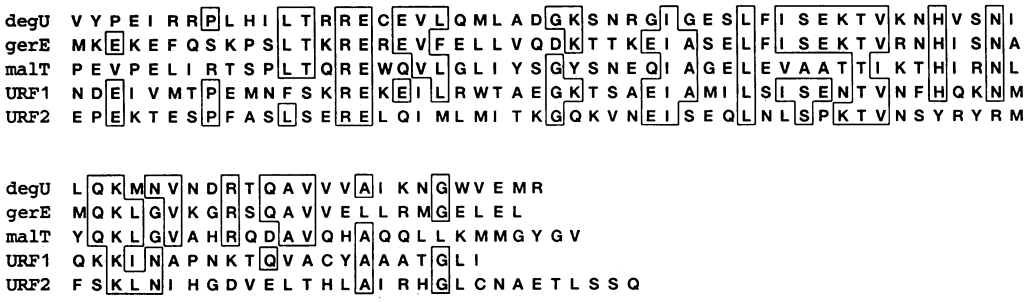

FIG. 5. Comparison of the C terminus of *degS* with other proteins. Residues that are identical in three of the five sequences are boxed. URF1 and URF2 refer to ORFs identified upstream of the *E. coli uvrC* gene. The references for the protein sequences used in this comparison are as follows: *gerE* (16), *malT* (7), URF1 and URF2 (38).

The second family of proteins, termed the effector or regulator family, have a conserved N-terminal domain that is modified by the cognate member of the modulator family (17). The best-understood pairs of these two component signal transduction systems are the *ntrB-ntrC* proteins involved in nitrogen regulation and the *cheA-cheB* and *cheA-cheY* proteins involved in chemotaxis. In both cases the modulator, *ntrB* or *cheA*, can transfer a phosphate molecule to the effector molecule, *ntrC*, *cheB*, or *cheY* (14, 28, 29, 47). The effectors can act in different ways. The *ntrC* protein can bind to a DNA sequence upstream of its target promoters and in its phosphorylated state activates transcription from these promoters (31). In the case of *cheB*, the phosphorylated protein has an increased methylesterase activity, increasing the rate of demethylation of the transmembrane receptors and initiating adaptation (A. Stock, personal communication). The phosphorylated *cheY* protein is somehow involved in an interaction with the flagellar motor to cause tumbly swimming behavior (Stock, personal communication). There have been two proteins described in *B. subtilis* which are members of the effector family, *spo0A* and *spo0F* (10, 44). Mutations in these genes lead to defects in sporulation (15). However, no cognate modulator proteins have been described for these sporulation proteins.

The membership of the *degU* gene product in this effector family of proteins suggests a model for how it might activate transcription of genes for degradative enzymes. We presume that the *degU* protein is phosphorylated by its cognate modulator protein. Whether the *degU* protein then directly interacts with the target sites located upstream of the *aprE* and *sacB* promoters (12, 13) has yet to be proven. However, this hypothesis is attractive for two reasons. The *ntrC* protein interacts with sites upstream of its target promoters to activate transcription (31). Also, the C terminus of the *degU* protein is similar to the C termini of a number of other positive regulatory proteins, which have been postulated to be DNA-binding proteins (7, 16, 38). The most striking similarity is to the *gerE* protein, a 72-amino-acid protein involved in the regulation of spore formation (16). It appears to be a transcriptional repressor of the *cotA* gene (35) and a transcriptional activator of the *cotC* gene (L. Zheng and S. Cutting, personal communication). However, whether this protein binds directly to DNA to cause these effects has not been demonstrated.

The most attractive hypothesis for the role of the *degS* gene product is that it is the cognate modulator for the *degU* gene product. There are some similarities between the amino acid sequence of the *degS* gene and other members of the modulator family, as pointed out in the accompanying paper (20). Also, an examination of the similarity of this protein with conserved domains of other modulator proteins by a computer analysis (19) shows a rather high alignment score (S. Parkinson, personal communication). However, purification of these proteins and demonstration of enzymatic activities will be needed to resolve this issue.

Several possibilities for how the *sacU*(Hy) mutations localized in the *degU* gene cause the SacU(Hy) phenotype come to mind. The first is that these changes alleviate the need for phosphorylation and cause this protein to become a constitutive transcriptional activator. The second is that the change makes the protein a better substrate for a related modulator protein and that "crosstalk" (28a) is activating the *degU* protein. The third is that the mutations increase the rate of phosphorylation or decrease the rate of autodephosphorylation of this protein (17). Mutations of *ntrC* have been isolated which seem analogous to the *degU* [*sacU*(Hy)] mutations. These mutations allow activation of target promoters in strains that are deleted for *ntrB* (18). Analysis of these mutant proteins suggests that one class of them is indeed active in the absence of phosphorylation, whereas others behave as though they are more effectively phosphorylated by "crosstalking" modulators (S. Kustu, personal communication). If the *degS* gene product were indeed a kinase, the phenotype of the *sacU*(Hy) mutations in that gene could be explained by either increased kinase activity or by activity in the absence of the natural signal.

We have previously suggested that the *sacU*(Hy) stimulation is part of a global regulatory system involved in the regulation of degradatory enzymes, perhaps in response to the need for alternative carbon or nitrogen sources (13). The nature of the operon described in this paper is consistent with this hypothesis, since the other members of this two-component family of sensory proteins are involved in such global regulatory systems as osmolarity changes, phosphate limitation, nitrogen limitation, and chemotaxis (32). However, we are still left with the puzzle of what signal this complex regulatory system senses. When the upstream *sacU*(Hy) stimulation site of either the *aprE* or *sacB* promoter is removed, there is little effect on the expression of these promoters in wild-type strains (12, 13), suggesting that this regulatory system has a very small role under the physiological conditions used for the assays. A major puzzle remains to determine the physiological trigger for this global regulatory system.

Nomenclature. A recent report by Tanaka and Kawata (41) described the isolation of a gene, *iep*, which is identical to the gene designated *degU* in this report. We are jointly proposing with Tanaka and Rapoport that these genes be designated *deg*, to reflect their role in the regulation of

degradative enzymes. The *sacQ* gene will be renamed *degQ*, and the *prtR* gene will be renamed *degR*.

ACKNOWLEDGMENTS

We thank Georges Rapoport and Frank Kunst for sharing their manuscript (20) before publication. We also thank those cited in personal communications for their discussions. We acknowledge James Hoch for pointing out the similarity between the *degU* and *gerE* genes. We acknowledge Stewart Causey for technical assistance.

LITERATURE CITED

1. **Amory, A. F., F. Kunst, E. Aubert, A. Klier, and G. Rapoport.** 1987. Characterization of the *sacQ* genes from *Bacillus licheniformis* and *Bacillus subtilis*. J. Bacteriol. **169**:324–333.
2. **Anagnostopoulos, C., and J. Spizizen.** 1961. Requirements for transformation in *Bacillus subtilis*. J. Bacteriol. **81**:741–746.
3. **Aymerich, S., G. Gonzy-Tréboul, and M. Steinmetz.** 1986. 5′-Noncoding region *sacR* is the target of all identified regulation affecting the levansucrase gene in *Bacillus subtilis*. J. Bacteriol. **166**:993–998.
4. **Ayusaw, D., Y. Yoneda, K. Yamane, and B. Mauro.** 1975. Pleiotropic phenomena in autolytic enzyme(s) content, flagellation and simultaneous hyperproduction of extracellular α-amylase and protease in a *Bacillus subtilis* mutant. J. Bacteriol. **124**:459–469.
5. **Buikema, W. J., W. W. Szeto, P. V. Lemley, W. H. Orme-Johnson, and F. M. Ausubel.** 1985. Nitrogen fixation specific regulatory genes of *Klebsiella pneumoniae* and *Rhizobium meliloti* share homology with the general nitrogen regulatory gene *ntrC* of *K. pneumoniae*. Nucleic Acids Res. **13**:4539–4555.
6. **Cohen, S. N., A. C. Y. Chang, and L. Hsu.** 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor RNA. Proc. Natl. Acad. Sci. USA **69**:2110–2114.
7. **Cole, S. T., and O. Raibaud.** 1986. The nucleotide sequence of the *malT* gene encoding the positive regulator of the *Escherichia coli* maltose regulon. Gene **42**:201–208.
8. **Drury, L. S., and R. S. Buxton.** 1985. DNA sequence analysis of the dye gene of *Escherichia coli* reveals amino acid homology between the Dye and OmpR proteins. J. Biol. Chem. **260**:4236–4242.
9. **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.
10. **Ferrari, F. A., K. Trach, D. Le Coq, J. Spence, E. Ferrari, and J. A. Hoch.** 1985. Characterization of the spo0A locus and its deduced product. Proc. Natl. Acad. Sci. USA **82**:2647–2651.
11. **Gilmann, M. Z., J. S. Glenn, V. S. Singer, and M. J. Chamberlin.** 1984. Isolation of sigma-28-specific promoters from *Bacillus subtilis* DNA. Gene **32**:11–20.
12. **Henner, D. J., E. Ferrari, M. Perego, and J. A. Hoch.** 1988. Location of the targets of the *hpr-97*, *sacU32*(Hy), and *sacQ36*(Hy) mutations in upstream regions of the subtilisin promoter. J. Bacteriol. **170**:296–300.
13. **Henner, D. J., M. Yang, L. Band, H. Shimotsu, M. Ruppen, and E. Ferrari.** 1986. Genes of *Bacillus subtilis* that regulate the expression of degradative enzymes, p. 81–90. *In* M. Alacevic, D. Hranueli, and Z. Toman (ed.), Genetics of industrial microorganisms. Proceedings of the fifth international symposium on the genetics of industrial microorganisms.
14. **Hess, T. F., N. Oosawa, N. Kaplan, and M. I. Simon.** 1988. Phosphorylation of three proteins in the signalling pathway of bacterial chemotaxis. Cell **53**:79–87.
15. **Hoch, J. A.** 1976. Genetics of bacterial sporulation. Adv. Genet. **18**:69–98.
16. **Holland, S. K., S. Cutting, and J. Mandelstam.** 1987. The possible DNA-binding nature of the regulatory proteins encoded by *spo*IID and *ger*F, involved in the sporulation of *Bacillus subtilis*. J. Gen. Microbiol. **133**:2381–2391.
17. **Keener, J., and S. Kustu.** 1988. Protein kinase and phosphoprotein phosphatase activities of nitrogen regulatory proteins NTRB and NTRC of enteric bacteria: roles of the conserved amino-terminal domain of NTRC. Proc. Natl. Acad. Sci. USA **85**:4976–4980.
18. **Keener, J., P. Wong, D. Popham, J. Wallis, and S. Kustu.** 1987. A sigma factor and auxiliary proteins required for nitrogen-related transcription in enteric bacteria, p. 159–175. *In* W. S. Reznikoff, R. R. Burgess, J. E. Dahlberg, C. A. Gross, M. T. Record, Jr., and M. P. Wichers (ed.), RNA polymerase and the regulation of transcription. Elsevier Scientific Publishing Co., Inc., New York.
19. **Kofoid, E. C., and J. S. Parkinson.** 1988. Transmitter and receiver modules in bacterial signalling proteins. Proc. Natl. Acad. Sci. USA **85**:4981–4985.
20. **Kunst, F., Debarbouille, M., T. Msadek, M. Young, C. Mauel, D. Karamata, A. Klier, G. Rapoport, and R. Dedonder.** 1988. Deduced polypeptides encoded by the *Bacillus subtilis sacU* locus share homology with two-component sensor-regulator systems. J. Bacteriol. **170**:5093–5101.
21. **Kunst, F., M. Pascal, J. Lepesant-Kejzlarova, J. A. Lepesant, A. Billault, and R. Dedonder.** 1974. Pleiotropic mutations affecting sporulation conditions and the synthesis of extracellular enzymes in *Bacillus subtilis*. Biochemie **56**:1481–1489.
22. **Lepesant, J. A., F. Kunst, J. Lepesant-Kejzlarova, and R. Dedonder.** 1972. Chromosomal location of mutations affecting sucrose metabolism in *Bacillus subtilis* Marburg. Mol. Gen. Genet. **118**:135–160.
23. **Lepesant, J. A., F. Kunst, M. Pascal, J. Lepesant Kejzlarová-Lepesant, M. Steinmetz, and R. Dedonder.** 1976. Specific and pleiotropic regulatory mechanisms in the sucrose system of *Bacillus subtilis*, p. 58–69. *In* D. Schlessinger (ed.), Microbiology—1976. American Society for Microbiology, Washington, D.C.
24. **Makino, K., H. Shinagawa, M. Amemura, and A. Nakata.** 1986. Nucleotide sequence of the *phoB* gene, the positive regulatory gene for the phosphate regulation of *Escherichia coli* K-12. Mol. Biol. **190**:37–44.
25. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. **Mutoh, N., and M. I. Simon.** 1986. Nucleotide sequence corresponding to five chemotaxis genes in *Escherichia coli*. J. Bacteriol. **165**:161–166.
27. **Nagami, Y., and T. Tanaka.** 1986. Molecular cloning and nucleotide sequence of a DNA fragment from *Bacillus natto* that enhances production of extracellular proteases and levansucrase in *Bacillus subtilis*. J. Bacteriol. **166**:20–28.
28. **Ninfa, A. J., and B. Magasanik.** 1986. Covalent modification of the *gln*G product, NRI, by the *gln*L product, NRII, regulates the transcription of the *gln*ALG operon *Escherichia coli*. Proc. Natl. Acad. Sci. USA **83**:5909–5913.

28a. **Ninfa, A. J., E. G. Ninfa, A. N. Lupas, A. Stock, B. Magasanik, and J. Stock.** 1988. Crosstalk between bacterial chemotaxis signal transduction proteins and regulators of transcription of the Ntr regulon: evidence that nitrogen assimilation and chemotaxis are controlled by a common phosphotransfer mechanism. Proc. Natl. Acad. Sci. USA **85**:5492–5496.

29. **Oosawa, K., J. F. Hess, and M. I. Simon.** 1988. Mutants defective in bacterial chemotaxis shows modified protein phosphorylation. Cell **53**:89–96.
30. **Pascal, M., F. Kurst, J. A. Lepesant, and R. Dedonder.** 1971. Characterization of two sucrose activities in *Bacillus subtilis* Marburg. Biochemie **53**:1054–1066.
31. **Reitzer, L. J., and B. Magasanick.** 1986. Transcription of *glnA* in *E. coli* is stimulated by activator bound to sites far from the promoter. Cell **45**:785–792.
32. **Ronson, C. W., B. T. Nixon, and F. Ausubel.** 1987. Conserved domains in bacterial regulatory proteins that respond to environmental stimuli. Cell **49**:579–581.
33. **Ruppen, M. E., G. L. Van Alstine, and L. Band.** 1988. Control of intracellular serine protease expression in *Bacillus subtilis*. J. Bacteriol. **170**:136–140.
34. **Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich.** 1988. Primer-

directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**:487–491.
35. **Sandman, K., L. Kroos, S. Cutting, P. Youngman, and R. Losick.** 1988. Identification of the promoter for a spore coat protein gene in *Bacillus subtilis* and studies on the regulation of its induction at a late stage of sporulation. J. Mol. Biol. **200**:461–473.
36. **Sanger, F., S. Niclen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
37. **Sekiguchi, J., N. Takada, and H. Okada.** 1975. Genes affecting the productivity of α-amylase in *Bacillus subtilis* Marburg. Biochemie **53**:1059–1066.
38. **Sharma, S., T. F. Stark, W. G. Beattie, and R. E. Moses.** 1986. Multiple control elements for the *uvrC* gene unit of *Escherichia coli*. Nucleic Acids Res. **14**:2301–2318.
39. **Shimotsu, H., and D. J. Henner.** 1986. Modulation of *Bacillus subtilis* levansucrase gene expression by sucrose and regulation of the steady-state mRNA level by *sacU* and *sacQ* genes. J. Bacteriol. **168**:330–388.
40. **Steinmetz, M., F. Kurst, and R. Dedonder.** 1976. Mapping of mutations affecting synthesis of exocellular enzymes in *Bacillus subtilis*. Mol. Gen. Genet. **148**:281–285.
41. **Tanaka, T., and M. Kawata.** 1988. Cloning and characterization of *Bacillus subtilis iep*, which has positive and negative effects on production of extracellular proteases. J. Bacteriol. **170**:3593–3600.
42. **Tanaka, T., M. Kawata, Y. Nagami, and H. Uchiyama.** 1987. *prtR* enhances the mRNA level of the *Bacillus subtilis* extracellular proteases. J. Bacteriol. **169**:3044–3050.
43. **Tomioka, N., M. Honjo, K. Funahashi, K. Manabe, A. Akaoka, I. Mita, and Y. J. Furutani.** 1985. Cloning, sequencing, and some properties of a novel *Bacillus amyloliquefaciens* gene involved in the increase of extracellular protease activities. J. Biotechnol. **3**:85–96.
44. **Trach, K. A., J. W. Chapman, P. J. Piggot, and J. A. Hoch.** 1985. Deduced product of the stage 0 sporulation gene *spo0F* shares homology with the *spo0A*, *ompR* and *sfrA* proteins. Proc. Natl. Acad. Sci. USA **82**:7260–7264.
45. **Viera, T., and J. Messing.** 1987. Production of single-stranded plasmid DNA. Methods Enzymol. **153**:3–11.
46. **Wurtzel, E. T., M. Y. Chou, and M. Inouye.** 1982. Osmoregulation of gene expression: DNA sequence of the *ompR* gene of the *ompB* operon of *Escherichia coli* and characterization of its gene product. J. Biol. Chem. **257**:13685–13691.
47. **Wylie, D., A. Stock, C.-Y. Wong, and J. Stock.** 1988. Sensory transduction in bacterial chemotaxis involves phosphotransfer between che proteins. Biochem. Biophys. Res. Commun. **151**: 891–896.
48. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.