

Location of the Targets of the *hpr-97*, *sacU32(Hy)*, and *sacQ36(Hy)* Mutations in Upstream Regions of the Subtilisin Promoter†

DENNIS J. HENNER,¹ EUGENIO FERRARI,^{2*} MARTA PEREGO,³ AND JAMES A. HOCH³

Genentech, Inc., South San Francisco, California 94080¹; Genencor, Inc., 180 Kimbal Way, South San Francisco, California 94080²; and Research Institute of Scripps Clinic, La Jolla, California 92037³

Received 31 August 1987/Accepted 21 October 1987

A number of mutations have been described with pleiotropic effects on the expression of genes for degradative enzymes in *Bacillus subtilis*. The *sacU32(Hy)* and *sacQ36(Hy)* mutations increase the expression of a wide variety of enzymes that degrade biological polymers. The phenotypes caused by mutations at the *hpr* locus are more restricted; they are known to increase expression of the alkaline and neutral proteases. The alkaline protease (*aprE*) promoter was analyzed to determine the target site for stimulation by these loci. Deletion of upstream regions of the *aprE* promoter could abolish or greatly reduce stimulation by mutations at these loci. A region upstream of -200 was necessary for full stimulation by an *hpr-97* mutation, whereas a region between -141 and -164 was necessary for full stimulation by the *sacU32(Hy)* and *sacQ36(Hy)* mutations. Northern analyses of mRNA preparations showed that the levels of *aprE* mRNA were increased in strains carrying the *sacU32(Hy)* or *hpr-97* mutation. Moreover, primer extension analysis of these mRNA preparations revealed that the transcription start point was identical to that in a wild-type strain. We hypothesize that upstream activation of the subtilisin promoter mediated by these genes is a mechanism for global responses to a variety of nutritional conditions.

In *Bacillus subtilis*, the expression of a variety of degradative enzymes can be increased by mutations at a number of loci. These loci are unlinked to the structural genes for the affected enzymes. Mutations at the *sacU* and *sacQ* loci can increase the expression of levansucrase, alkaline protease, neutral protease, xylanase, β -glucanase, α -amylase, and intracellular serine protease (1, 2, 10-13). Other mutations of the *sacU* locus decrease the expression of these enzymes. Another locus, designated *hpr*, has been defined which has a more restricted phenotype. Mutations at this locus lead to increased expression of alkaline and neutral proteases but do not affect the expression of levansucrase or α -amylase (8). Mutation at the *hpr* locus may have a small effect on the expression of intracellular serine protease (M. Ruppen, personal communication). Another gene, designated *prtR*, has recently been isolated whose overexpression on a plasmid can stimulate levansucrase and alkaline and neutral protease production (14, 21). No chromosomal mutation of this gene has been described.

The mechanism by which the *sacU32(Hy)* and *sacQ36(Hy)* mutations act to increase the expression of the levansucrase (*sacB*) gene has been the focus of a number of recent reports (2, 17). Mutations at both loci can increase the amount of *sacB* mRNA, probably by increasing the rate of transcription initiation. The mRNA start site was identical in strains that have *sacU32(Hy)*, *sacQ36(Hy)*, or wild-type alleles at these loci (17). How these mutations act to increase transcription of the *sacB* gene is unknown. Deletion analysis of the *sacB* promoter suggested that a region near -100 was necessary for complete stimulation of *sacB* transcription by the *sacU32(Hy)* and *sacQ36(Hy)* mutations (H. Shimotsu, M. Yang, and D. J. Henner, unpublished data).

The *prtR* and *sacQ* genes have been isolated and shown to encode 60- and 46-amino-acid polypeptides, respectively

(14, 20). They have no obvious similarity to other known polypeptides and only very limited similarity to each other. Overproduction of the *sacQ* or *prtR* polypeptides leads to stimulation of expression of these degradative enzymes (1, 14, 20, 21). It is not known whether these polypeptides act directly on their target genes or indirectly by stimulating the production of other regulatory factors. The nature of the *sacU* gene is unknown, and an initial report of its isolation has subsequently been retracted (G. Rapoport, personal communication). The *hpr* gene has been recently isolated, and the nature of stimulatory mutations at this site should soon be known (M. Perego and J. A. Hoch, unpublished data).

In these studies, the *aprE* promoter was analyzed to determine the site at which the *hpr-97*, *sacU32(Hy)*, and *sacQ36(Hy)* mutations act to stimulate its transcription. We determined that there are at least two regions involved in the stimulation of this promoter. The target site for *hpr-97* stimulation is separable from that of *sacU32(Hy)* and *sacQ36(Hy)* stimulation, and both target sites lie rather far upstream of the transcription start site. We also demonstrated that the *hpr-97* and *sacU32(Hy)* mutations act by increasing the amount of mRNA initiated at the same start point as in wild-type strains.

MATERIALS AND METHODS

Nucleic acids. The nucleotide sequence of DNA preparations was determined by the method of Sanger et al. (16) with M13 bacteriophages mp18 and mp19 for templates (22).

The generation and characterization of deletions are described in the accompanying paper (4). mRNA was isolated and primer extension and Northern (RNA) blot analyses were undertaken as described in the accompanying paper (4).

Strains. The following strains were used in this study: BG125 *trpC2 his-1 thr-5*, BG2335 *hpr-97 his-1 thr-5*, BG2334 *his-1 thr-5 scoC4*, BG29 *sacQ36(Hy) trpC2 ald*, and BG4049 *trpC2 thr-5 sacU32(Hy)*. Into these strains were integrated

* Corresponding author.

† Publication 4907-BCR from the Research Institute of Scripps Clinic.

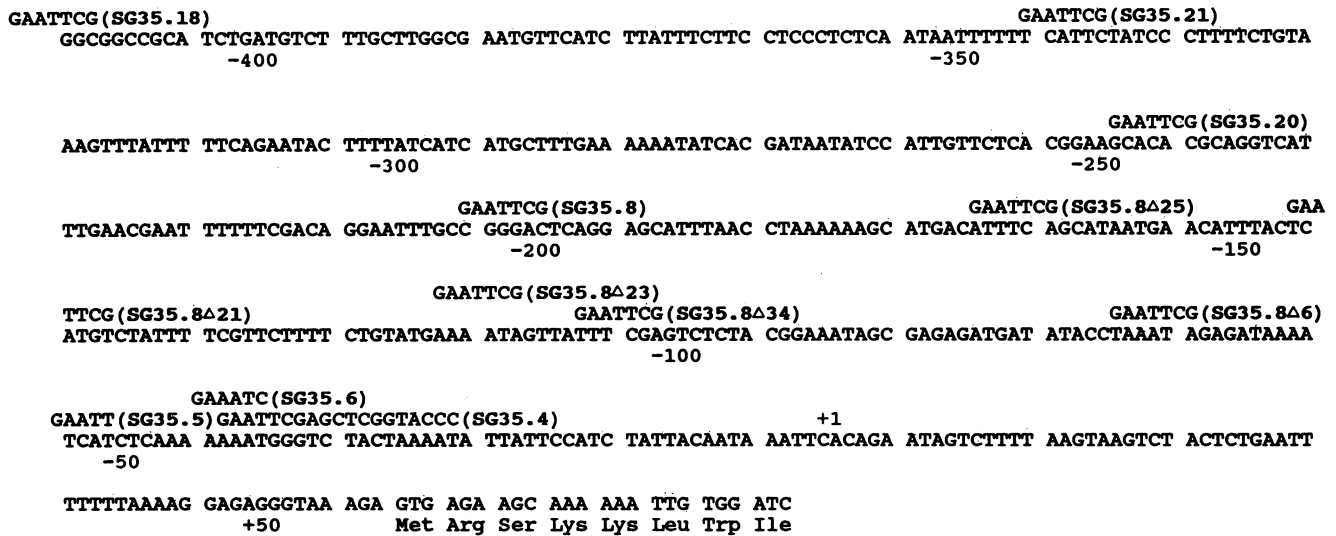


FIG. 1. Nucleotide sequence of the *aprE* promoter. The sequence was determined for both strands, and the sequence across all restriction sites was confirmed by sequencing across the site. The transcription initiation site is denoted by +1. Numbering is with respect to the transcription initiation site. The endpoints of each promoter derivative discussed in the text are shown, and the linker to the *EcoRI* or *BamHI* site used to construct the *lacZ* fusions is also shown for each derivative.

the single-copy vectors carrying various deletions upstream of the *aprE* promoter. The single-copy vectors, as well as the β -galactosidase assay methods, are described in the accompanying report (4). Construction of combinations of chromosomal markers was carried out by transformation or transduction (9).

RESULTS

Nucleotide sequence of the *aprE* promoter region. Although the nucleotide sequence of the *B. subtilis* promoter region has been previously reported (18), only a short distance upstream of the coding region has been published. To facilitate an examination of the effect of upstream regions on *aprE* expression, the sequence was determined for approximately 500 base pairs upstream of the initiation codon. The stimulation by all of the mutations studied here was found to occur between an upstream *NotI* site and the promoter. This sequence is shown in Fig. 1.

Deletion analysis of the subtilisin promoter. A series of deletions derivatives of the single-copy integration vector pSG35.1 was constructed as described in the accompanying paper (4). The endpoints of each deletion are shown in Fig. 1. Each derivative was transformed into strain BG125. Subsequently, the *sacU32*(Hy) and *sacQ36*(Hy) mutations were transduced into the BG125 parent by cotransduction with *hisA* and *thr*, respectively. A series of *hpr-79 aprE-lacZ* mutant strains was created by DNA transformation, with DNA from the BG125 series of strains as the donor strain and *hpr-97*-bearing strains as the recipient and chloramphenicol resistance as the selection. The β -galactosidase profile of each strain was then determined.

Figure 2 shows a comparison of the rates of β -galactosidase synthesis for several deletions in either the *hpr-97* or the *sacQ36*(Hy) background. Both the *sacQ36*(Hy) and *hpr-97*-mutations stimulated the apparent rate of β -galactosidase synthesis approximately 10-fold over that observed in wild-type strains. Their effects are due to stimulation of the initial rate, and this stimulation does not occur until the end of exponential growth. The *hpr-97* mutation maximally stim-

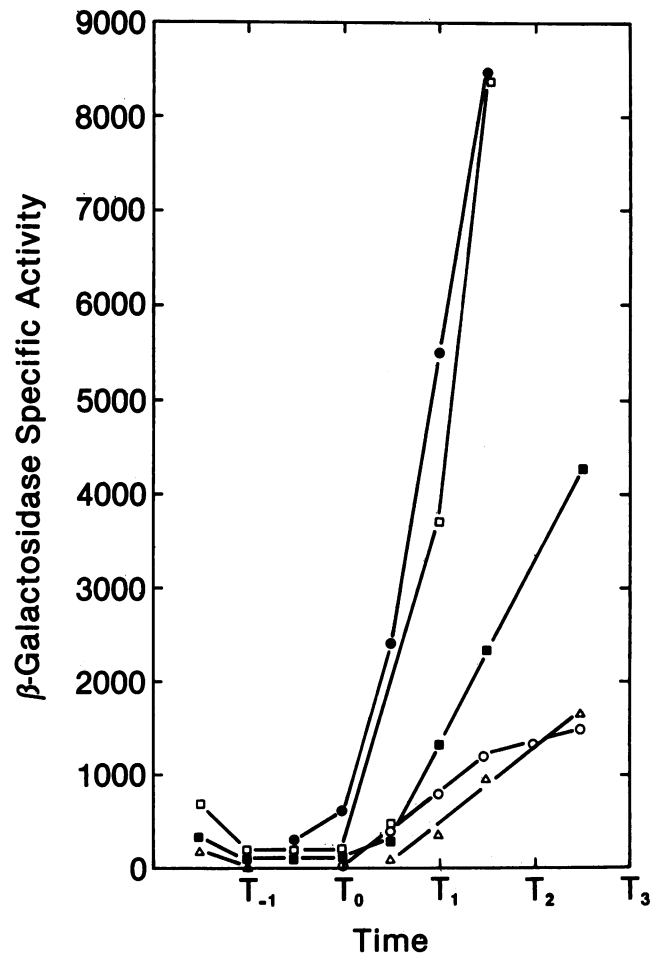


FIG. 2. β -Galactosidase profiles. T_0 indicates the time at which the culture left the exponential phase of growth. Genetic background and *aprE-lacZ* fusion derivative: ●, *hpr-97* and SG35.18; ○, wild type and SG35.18; □, *sacQ36*(Hy) and SG35.8Δ25; ■, *sacQ36*(Hy) and SG35.8Δ21; △, *sacQ36*(Hy) and SG35.8Δ6.

TABLE 1. Effect of upstream deletion and *sacU*(Hy), *sacQ*(Hy), and *hpr* mutations on the rate of β -galactosidase accumulation from the *aprE* promoter in an *aprE-lacZ* fusion

| Promoter | Deletion endpoint | Initial rate of β -galactosidase accumulation (U/mg per h) | | | |
|--------------------|-------------------|--|------------------|------------------|------------|
| | | Wild type | <i>sacU</i> (Hy) | <i>sacQ</i> (Hy) | <i>hpr</i> |
| SG35.1 | -600 | 490 | | | 8,000 |
| SG35.18 | -412 | 750 | | | 6,800 |
| SG35.21 | -340 | 500 | | | 4,000 |
| SG35.20 | -244 | 1,700 | | | 4,000 |
| SG35.8 | -200 | 800 | 6,450 | 7,250 | 440 |
| SG35.8 Δ 25 | -164 | 575 | 8,400 | 9,950 | 500 |
| SG35.8 Δ 21 | -141 | 610 | 1,900 | 2,200 | 300 |
| SG35.8 Δ 23 | -113 | 560 | 750 | 2,200 | 340 |
| SG35.8 Δ 34 | -104 | 600 | 900 | 2,200 | 202 |
| SG35.8 Δ 6 | -64 | 375 | 800 | 900 | 260 |

ulated deletion SG35.18, whereas this deletion showed normal activity in a wild-type strain. Representative rates for the deletion series in a *sacQ36*(Hy) background are also shown in Fig. 2. Deletions SG35.8 and SG35.8 Δ 25 were maximally stimulated, whereas deletion SG35.8 Δ 6 was unresponsive to the *sacQ36*(Hy) mutation. Deletion SG35.8 Δ 21 was three- to fourfold less stimulated by the *sacQ36*(Hy) mutation but still showed about a threefold higher rate of synthesis than did the wild-type strain. Kinetics identical to those of SG35.8 Δ 21 were obtained for deletions SG35.8 Δ 23 and SG35.8 Δ 34.

Table 1 shows the rate of β -galactosidase accumulation in all of the strains tested. All of the deletions, with the exception of SG35.8 Δ 20, had similar rates of β -galactosidase accumulation in a wild-type background. Independent wild-type transformants carrying derivative SG35.8 Δ 20 showed consistently higher rates of β -galactosidase accumulation than did the other derivatives. Deletion of the sequence between -164 and -141 reduced stimulation by both *sacU32*(Hy) and *sacQ36*(Hy) from approximately 10-fold to about 3-fold. Further deletions to -64 still retained some stimulation by both *sacU32*(Hy) and *sacQ36*(Hy).

The region necessary for stimulation by the *hpr-97* mutation lies upstream of -200. Deletions from approximately -600 to -400 or -340 showed small decreases in the rate of β -galactosidase accumulation in an *hpr-97* background and in stimulation. A further deletion to -244 showed the same rate of β -galactosidase accumulation as the -340 deletion, but the stimulation was much reduced if the increased rate in the wild-type background was taken into consideration. A deletion to -200 showed a sharp cutoff in the rate of β -galactosidase accumulation, and this and further deletions all had lower rates of accumulation than the same deletions in the wild-type background.

Determination of *aprE* mRNA levels and initiation points in *sacU32*(Hy) and *hpr-97* mutations. The mechanism by which the *sacU32*(Hy) and *hpr-97* mutations stimulate expression was examined by determining the amount of *aprE* mRNA in a quantitative Northern analysis. A substantially larger amount of *aprE* mRNA could be detected in preparations extracted from strains carrying the *sacU32*(Hy), *scoC4*, and *hpr-97* mutations, as opposed to a strain that was wild type at both loci (Fig. 3). Northern analysis carried out with mRNA from a *sacU32*(Hy) *spo0A* strain gave no detectable band of hybridization, confirming that the increased signal seen in the *sacU32*(Hy) strain is due to subtilisin mRNA (data not shown). The *scoC4* mutation resides in the *hpr* locus (M. Perego and J. A. Hoch, unpublished data). At this level of resolution, there appeared to be only one size of

mRNA in all three strains, suggesting that these mutations do not function by activating different promoters. This point was also addressed by examining the mRNA start point in each strain. A single major start point was seen for each mRNA preparation (Fig. 4). The amount of mRNA detected by this method was also greater in the preparations from the *sacU32*(Hy), *scoC4*, and *hpr-97* strains, confirming the Northern blot analysis. A number of minor bands were also seen near +31 (Fig. 4). As discussed in the accompanying paper, we suspect that these bands are artifactual (4). No other bands were apparent on the entire gel, and no band was visible at the top of the gel, indicating that no transcript extended from upstream of the designated initiation point. Since both bands increase in the mutant strains and no signal was found at either location with mRNA from a *spo0A* strain (data not shown) we concluded that both bands are subtilisin

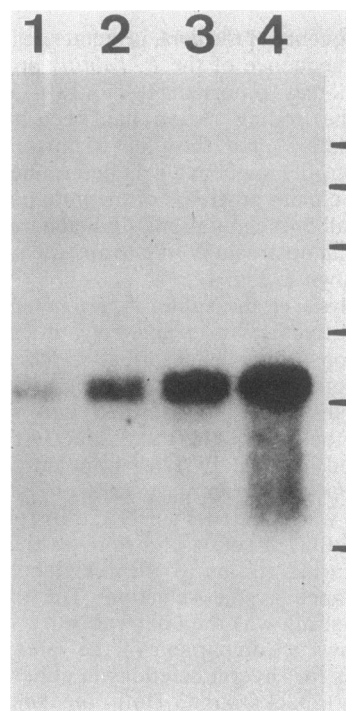


FIG. 3. Northern analysis of *aprE* mRNA. The marks at the right indicate the positions of RNA standards with molecular sizes from top to bottom of 9,500, 7,500, 4,400, 2,400, 1,400, and 300 base pairs. The relevant genotypes of the strains used for mRNA extraction are as follows (lanes): 1, wild type; 2, *scoC4*; 3, *hpr-97*; 4, *sacU32*(Hy).

derived and not due to nonspecific priming of another mRNA.

DISCUSSION

The Northern analyses and determination of the *aprE* mRNA start point suggest that both the *sacU32*(Hy) and *hpr-97* mutations act by increasing the steady-state level of mRNA. An identical finding has been reported for the actions of the *sacU32*(Hy) and *sacQ36*(Hy) mutations on the *sacB* promoter (2, 17). For both promoters, the same start site appears to be utilized in all strains, suggesting that a cryptic promoter is not activated by these mutations. A previous study has shown that *sacB* mRNA half-life is not changed by the presence of a *sacU32*(Hy) mutation (3), and we favor the hypothesis that these mutations act by increasing the rate of transcription initiation rather than by stabilizing the mRNA. A similar conclusion has been reached for the stimulation of the *aprE* promoter by the product of the *prtR* gene (19).

The analysis of the deletion derivatives of the *aprE* promoter showed that regions well upstream of the transcription start site are necessary for full stimulation by the *hpr-97*, *sacU32*(Hy), and *sacQ36*(Hy) mutations. These deletions roughly define the left edge of the regions necessary for this full stimulation. There are no obvious palindromic sequences, repeated elements, or sequences similar to

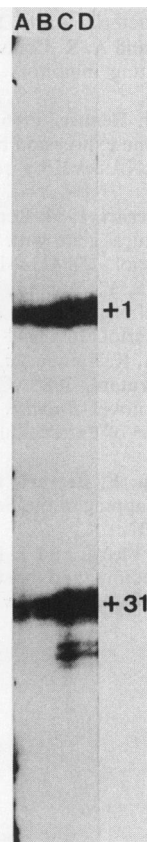


FIG. 4. Determination of the transcription start site for *aprE* in *hpr-97* and *sacU32*(Hy) strains. Primer extension analyses were carried out with mRNA from wild-type (lane A), *scoC4* (lane B), *hpr-97* (lane C), and *sacU32*(Hy) (lane D) strains. A control sequencing ladder was run beside the reactions to verify the +1 and +31 positions.

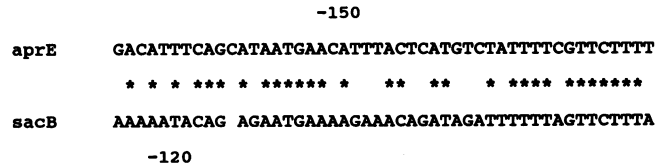


FIG. 5. Comparison of *aprE* and *sacB* upstream regions containing the target for *sacU32*(Hy) and *sacQ36*(Hy) stimulation. The *sacB* sequence is from Shimotsu and Henner (17). The asterisks indicate identical nucleotides. The positions with respect to the transcription start site are indicated.

known protein-binding sites defined by these deletions (6). The analysis of the deletion mutations is not as clear-cut as we would have desired. There is definitely a clear difference between the deletion at -200 and all those with larger upstream regions for *hpr-97* stimulation. But whether there is a single site between -244 and -200 or whether there are a number of upstream sites whose effects are cumulative between -200 and -400 is difficult to determine from these data. Similarly, the -164 and -141 deletions clearly differ for *sacU32*(Hy) and *sacQ36*(Hy) stimulation. However, further deletions show a consistent stimulation of up to three-fold. The *sacU32*(Hy) and *sacQ36*(Hy) mutations are known to have very pleiotropic effects on cells, and one could postulate that this further stimulation is a secondary effect. It is also possible that a second target site exists within or downstream of the promoter region that mediates this additional stimulation. Deletion mutations of the upstream regions of the *nif* promoters show a similar complex pattern of stimulation (7).

A similar study using the *sacB* promoter suggested that a region near -100 was necessary for full stimulation of the *sacB* promoter by the *sacU32*(Hy) and *sacQ36*(Hy) mutations (H. Shimotsu and D. J. Henner, unpublished data). Interpretation of these data was complicated by the finding that a second region downstream of the promoter appeared to mediate both sucrose induction and stimulation by *sacU32*(Hy) and *sacQ36*(Hy) (Shimotsu and Henner, unpublished data). For the *aprE* promoter, the region between -164 and -141 appears to be necessary for full stimulation by the *sacU32*(Hy) and *sacQ36*(Hy) mutations. Comparison of the DNA sequences in this region showed some similarities (Fig. 5). However, the similarities are primarily in a T-rich region and it is difficult to judge their significance. A comparison of this region with other known target genes of the *sacU32*(Hy) and *sacQ36*(Hy) mutations showed similarities in the regions upstream of the genes. The transcription start points of some of these genes have not been characterized. However, the comparisons picked up primarily stretches of T's and there was no convincing consensus sequence.

Although positive stimulation of promoters upstream of the RNA polymerase recognition site has been defined for a number of systems (for a review, see reference 15), it is not so common for sites upstream of -100 to stimulate transcription. The most analogous case appears to be that of the *nif* genes. Both the *nifA* and *ntrC* gene products appear to be transcription stimulatory factors whose target sites are often found 100 to 150 base pairs upstream of the transcription start site and which can function when placed more than 1,000 base pairs upstream (7). Studies are in progress to determine whether the spacing of the target sites for the *hpr-97*, *sacU32*(Hy) and *sacQ36*(Hy) mutations is critical for their activity.

Two central questions are raised by this research. What are the factors that interact with the target sites to cause stimulation, and what are the physiological roles of these stimulatory systems? Although the *sacU32*(Hy), *sacQ36*(Hy), and *hpr-97* mutations stimulate expression of *aprE*, there is no evidence that they directly interact with the target sites. At least two other genes have been characterized which also influence expression of the *aprE* gene. Overexpression of the *prtR* gene, a 60-amino-acid polypeptide, on a high-copy plasmid stimulates expression of levansucrase, neutral protease, and alkaline protease (14). This phenotype appears to be very similar to that caused by the *sacU32*(Hy) and *sacQ36*(Hy) mutations and might have the same target site(s). Another region of DNA, encoding two small polypeptides, has been isolated which inhibits both sporulation and alkaline protease production (5). Analysis of some of the *aprE-lacZ* deletion derivatives described here suggests that the inhibition is caused at an upstream site on the *aprE* promoter (I. Smith, personal communication). Why should the cell have such a complex set of controls for a gene that is nonessential for growth, at least in a laboratory environment? We speculate that, since the enzymes stimulated by these mutations are all involved in the degradation of complex polymers, these controls are part of a global regulatory system involved in the search for alternative nutrient sources.

ACKNOWLEDGMENTS

Excellent technical assistance in these studies was rendered by Louise Band, Maria Yang, Walter Escobar, and Sandra Howard.

The work was supported in part by U.S. Public Health Service grant GM19416 from the National Institutes of Health.

LITERATURE CITED

- Amory, A., 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.
- 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.
- Chambert, R., and M.-F. Petit-Glatron. 1984. Hyperproduction of exocellular levansucrase by *Bacillus subtilis*: examination of the phenotype of a *sacU^h* strain. *J. Gen. Microbiol.* **130**:3143-3152.
- Ferrari, E., D. J. Henner, M. Perego, and J. A. Hoch. 1987. Transcription of *Bacillus subtilis* subtilisin and expression of subtilisin in sporulation mutants. *J. Bacteriol.* **170**:289-295.
- Gaur, N. K., E. Dubnau, and I. Smith. 1986. Characterization of a cloned *Bacillus subtilis* gene that inhibits sporulation in multiple copies. *J. Bacteriol.* **168**:860-869.
- Gicquel-Sanzey, B., and P. Cassart. 1982. Homologies between different prokaryotic DNA-binding regulatory proteins and between their sites of action. *EMBO J.* **1**:591-595.
- Gussin, G. N., C. W. Ronson, and F. M. Ausubel. 1986. Regulation of nitrogen fixation genes. *Annu. Rev. Genet.* **20**:567-591.
- Higerd, T. B., J. A. Hoch, and J. Spizizen. 1972. Hyperprotease-producing mutants of *Bacillus subtilis*. *J. Bacteriol.* **112**:1026-1028.
- Hoch, J. A., M. Barat, and C. Anagnostopoulos. 1967. Transformation and transduction in recombination-defective mutants of *Bacillus subtilis*. *J. Bacteriol.* **93**:1925-1937.
- 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* 168. *Biochimie* **56**:1481-1489.
- 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.
- Lepesant, J.-A., F. Kunst, M. Pascal, J. Kejzlarová-Lepesant, M. Steinmetz, and R. Dedonder. 1976. Specific and pleiotropic regulatory mechanisms in the sucrose system of *Bacillus subtilis* 168, p. 58-69. In D. Schlessinger (ed.), *Microbiology—1976*. American Society for Microbiology, Washington, D.C.
- Lepesant, J.-A., J. Lepesant-Kejzlarova, M. Pascal, F. Kunst, A. Billault, and R. Dedonder. 1974. Identification of the structural gene of levansucrase in *Bacillus subtilis* Marburg. *Mol. Gen. Genet.* **128**:213-221.
- 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.
- Ribaudo, O., and M. Schwartz. 1984. Positive control of transcription initiation in bacteria. *Annu. Rev. Genet.* **18**:173-206.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463-5467.
- 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**:380-388.
- Stahl, M. L., and E. Ferrari. 1984. Replacement of the *Bacillus subtilis* subtilisin structural gene with an in vitro-derived deletion mutation. *J. Bacteriol.* **158**:411-418.
- 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.
- Tomioaka, 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.
- Yang, M., H. Shimotsu, E. Ferrari, and D. J. Henner. 1987. Characterization and mapping of the *Bacillus subtilis prtR* gene. *J. Bacteriol.* **169**:434-437.
- 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.