Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 8773–8777, December 1987
Biochemistry

Bacillus subtilis sucrose-specific enzyme II of the phosphotransferase system: Expression in *Escherichia coli* and homology to enzymes II from enteric bacteria

(membrane protein/carbohydrate transport)

AGNES FOUET, MARYVONNE ARNAUD, ANDRE KLIER, AND GEORGES RAPOPORT

Unité de Biochimie Microbienne, Département des Biotechnologies, Institut Pasteur, 25, rue du Dr Roux F75724, Paris Cedex 15, France

Communicated by Saul Roseman, July 29, 1987

ABSTRACT Sucrose is transported into *Bacillus subtilis* cells by way of a phosphotransferase system, which consists of a specific enzyme II, a nonspecific enzyme I, and a histidine-containing phosphocarrier protein. Mutations in the *sacP* locus abolish the specific transport of sucrose. The *B. subtilis sacP* gene was cloned and expressed in *Escherichia coli*, and transformed cells could transport and phosphorylate sucrose. This indicates that the *sacP* gene product is enzyme II of the sucrose phosphotransferase system of *B. subtilis*. The nucleotide sequence of the *sacP* gene was determined and was found to overlap with the *sacA* gene at the tetranucleotide ATGA, which may allow a translational coupling between *sacP* and *sacA*. The two genes are therefore probably organized in an operon structure with the promoter located 5′ to *sacP* gene. The deduced amino acid sequence gave a M_r of 48,945 for the sucrose-specific enzyme II polypeptide. The amino acid sequence was compared to that of three other known enteric bacterial enzymes II (β-glucoside-specific enzyme II, mannitol-specific enzyme II, and glucose-specific enzyme II). Homology was found with β-glucoside enzyme II, and well conserved regions were identified through the comparison of the proteins.

The sucrose metabolic system has been proposed as a model for studying the regulation of gene expression in *Bacillus subtilis*, and biochemical and genetical studies have established that at least nine different loci are involved (1, 2). Three of these loci, *sacT, sacP*, and *sacA*, are clustered; sucrose induces the expression of *sacP* and *sacA*. *sacA* codes for sucrase, an endocellular sucrose-6-phosphate hydrolase. It has been cloned, sequenced, and expressed in *Escherichia coli*, which contains no endogenous sucrase (3, 4).

In *B. subtilis*, sucrose is transported exclusively by way of a phospho*enol*pyruvate-dependent phosphotransferase system (PTS) (5), which involves two general proteins and one specific protein. Enzyme I of the PTS (enzyme I) and histidine-containing phosphocarrier protein, which are cytoplasmic proteins, are required as general phosphocarrier proteins for the transport and phosphorylation of all PTS sugars. Sugar specificity is determined by enzyme II, which is a specific integral membrane protein of the PTS. Cells harboring mutations in the *sacP* locus are unable to transport sucrose but are able to transport other PTS sugars, indicating that mutations in *sacP* act directly or indirectly on sucrose-specific enzyme II (II^{Suc}). The gene for enzyme I has been mapped on the chromosome, and it is not linked to the *sacT sacP sacA* cluster. The *sacT* locus has been defined by a single mutation giving rise to the constitutive high level synthesis of both the PTS and sucrase.

In this paper, we describe a plasmid carrying the *sacP* gene and the complementation of a *sacP* mutation in a Rec^- strain of *B. subtilis*. Expression of *sacP* in *E. coli* was also obtained, demonstrating that *sacP* corresponds to the structural gene for II^{Suc}. The first nucleotide sequence*, to our knowledge, of the gene for an enzyme II from a Gram-positive bacteria is reported, and the deduced amino acid sequence is presented. Sequence comparisons were made with enzyme II sequences from Gram-negative bacteria, and homology with *E. coli* β-glucoside-specific enzyme II (II^{Bgl}) is shown.

MATERIALS AND METHODS

Bacterial Strains, Phages, and Plasmids. *B. subtilis* QB976 (*trpC2, sacP1, recE4*) was constructed in this laboratory. *E. coli* strain TG1 (6) was used as a host for pUC9 derivatives (7), for pMK4 derivatives (8), and for the sequencing vector M13mp19 (9) and its derivative phages. *E. coli* strain HB101 (10) was used for subcloning experiments. *E. coli* strain Ts19 ($PtsI^{ts}$) (11) was used for assay of sucrose uptake.

DNA Preparation and Plasmid Constructions. Plasmids were isolated from *E. coli* as described by Birnboim and Doly (12). Single-stranded DNA was purified as described by Messing (7). Vectors were linearized and treated with calf intestine phosphatase, and fragment ends were made blunt by using the Klenow fragment of *E. coli* polymerase I.

Transformation and Selection of Recombinants. Transformation and selection of *E. coli* and *B. subtilis* were performed as described (13, 14). MacConkey plates, on which sugar fermentation gives rise to red colonies, were used.

Assay for Sucrose Uptake. The sucrose uptake was measured as described by Delobbe *et al.* (15).

***In Vitro* Transcription/Translation of Plasmid-Encoded Proteins.** Covalently closed circular plasmids were used as templates for a prokaryotic coupled transcription/translation system as recommended by the manufacturer (Amersham). Products were separated by $NaDodSO_4$/PAGE (10% acrylamide/0.27% bisacrylamide), and [^{35}S]methionine-labeled proteins were visualized by autoradiography.

DNA Sequencing. Nucleotide sequencing by the dideoxy chain termination method (16) was carried out using phage M13mp19 (9) and the synthetic 17-mer primer. dATP[α-^{35}S] (22,200 GBq/mmol) was supplied by Amersham. Overlapping deletions were obtained using the technique described by Dale *et al.* (17). The experiment was carried out as directed by the manufacturer (International Biotechnologies, New Haven, CT).

Abbreviations: PTS, phosphotransferase system; enzyme I, enzyme I of the PTS; enzyme II, specific integral membrane protein of the PTS; II^{Mtl}, mannitol-specific enzyme II; II^{Bgl}, β-glucoside-specific enzyme II; II^{Suc}, sucrose-specific enzyme II.

*The sequence reported in this paper is being deposited in the EMBL/GenBank data base (Bolt, Beranek, and Newman Laboratories, Cambridge, MA, and Eur. Mol. Biol. Lab., Heidelberg) (accession no. J03006).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

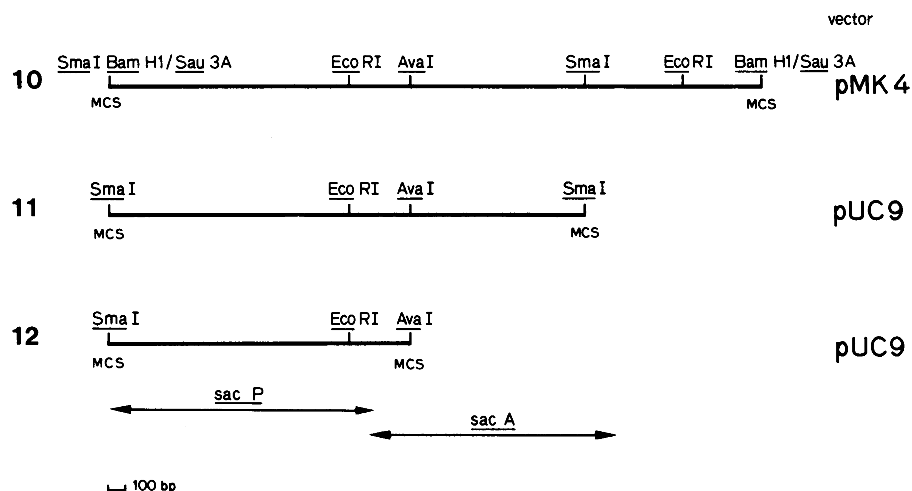

FIG. 1. Restriction endonuclease maps of plasmids carrying the *B. subtilis sacP* gene. The vectors are not displayed. The plasmids labeled 10, 11, and 12 represent pBSG8-10, pBSG8-11, and pBSG8-12, respectively. Plasmids pBSG8-11 and pBSG8-12 are pUC9 derivatives containing a 2.8-kb *Sma* I fragment of pBSG8-10 and a 1.7-kb *Sma* I–*Ava* I fragment of pBSG8-10, respectively. MCS, multicloning sites from pUC9.

RESULTS AND DISCUSSION

Isolation and Characterization of a DNA Fragment Harboring the *sacP* Gene. The 2.1-kilobase (kb) *Eco*RI DNA fragment harboring the *sacA* gene and part of the *sacP* gene (4) was used as a probe to screen the library constructed in *E. coli* using the shuttle vector pMK4 (18). One of the clones giving a positive signal was isolated, and the plasmid contained a 3.9-kb *B. subtilis* DNA fragment and was designated pBSG8-10 (Fig. 1).

The *B. subtilis* strain QB976 (*sacP1*, *recE4*) was transformed by pBSG8-10. All the chloramphenicol-resistant transformants tested grew when the sole carbon source was sucrose (0.1%). It has been shown that the 2.1-kb *Eco*RI fragment harbors the entire *sacA* gene and the 3′ end of the *sacP* gene (4). It can therefore be concluded that pBSG8-10 codes for a functional *sacP* gene product and for sucrase.

Characterization of the *sacP* Gene Product in *E. coli*. Genetic and biochemical studies have shown that *sacP* mutants do not take up sucrose (1), but no direct evidence has been reported that *sacP* codes for II^{Suc}. In order to test this hypothesis, the *sacP* gene was expressed in *E. coli*, which is devoid of sucrose transport and saccharolytic activities, but contains enzyme I and histidine-containing phosphocarrier protein. The *sacP* gene was isolated on recombinant plasmid pBSG8-12 by subcloning the 1.7-kb *Sma* I–*Ava* I DNA fragment in pUC9 (Fig. 1). An *E. coli* Rec$^-$ strain (HB101) was transformed by pBSG8-10 and by pBSG8-12, and the capacity of the transformants to metabolize sucrose was tested. Cells transformed by the vector alone or by the recombinant plasmids containing either the sucrase gene or the *sacP* gene (pBSG8-12) remained white on MacConkey plates containing 2% sucrose, whereas those transformed by pBSG8-10, which carries both the *sacA* and *sacP* genes, grew as red colonies on MacConkey plates due to the hydrolysis of sucrose. These cells were also replica plated onto M9 agar plates containing 0.1% sucrose and ampicillin. The pBSG8-10 transformants grew well, thus demonstrating that they could use sucrose as the sole carbon source. Moreover, the growth curves obtained when strain TGI (pBSG8-10) was grown on M63 liquid medium supplemented with ampicillin (25 μg/ml), chloramphenicol (1.5 μg/ml), and either 0.1% glucose or 0.2% sucrose were identical. The generation time was 90 min at 37°C in both cases. The *sacP* gene product is therefore sufficient to allow the transport of sucrose into the *E. coli* cytoplasm.

The hypothesis that the *sacP* gene product is enzyme II was tested by assaying [U-^{14}C]sucrose transport in strains of *E. coli* containing *sacP* and either a thermosensitive enzyme

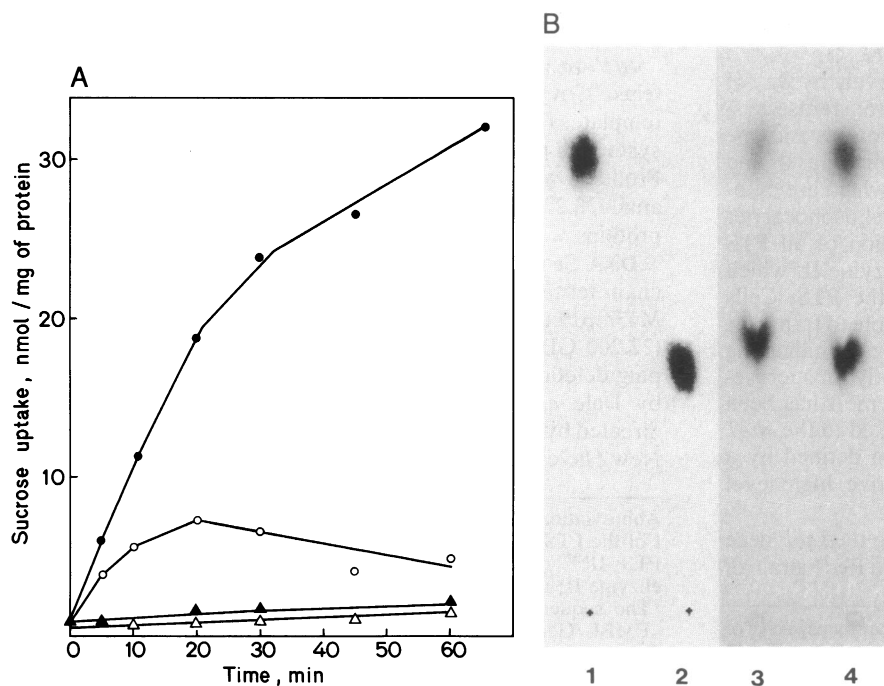

FIG. 2. Sucrose uptake by *E. coli* Ts19 recombinant cells. Growth was stopped by addition of chloramphenicol (250 μg/ml), and [U-^{14}C]sucrose was added to a final concentration of 0.3 mM (92.5 MBq/mmol) in a final volume of 10 ml. (*A*) [U-^{14}C] sucrose uptake. Aliquots (1 ml) were filtered through Millipore filters. ●, Ts19 (pBSG8-12) grown in L medium containing ampicillin (100 μg/ml) and further incubated at 30°C; ○, Ts19 (pBSG8-12) grown and further incubated at 42°C; ▲, Ts19 (pUC9) at 30°C; △, Ts19 (pUC9) at 42°C. (*B*) Uptake and concomitant phosphorylation of [U-^{14}C]sucrose. At 30 and 60 min a 1-ml aliquot from Ts19 (pBSG8-12) at 30°C was centrifuged, and the cytoplasmic content was analyzed by ascending paper chromatography. The solvent used was 1 M NH_4OAc, pH 3.8/ethanol, 2:5 (vol/vol). In lanes 1 and 2, sugars were revealed by treatment with 8 M urea in 4 M hydrochloric acid/acetone, 1:9 (vol/vol) after 5 min at 120°C. In lanes 3 and 4, labeled sugars were displayed by autoradiography. Lanes: 1, sucrose; 2, sucrose-6-phosphate; 3, Ts19 (pBSG8-12) at 30°C, after 30 min; 4, Ts19 (pBSG8-12) at 30°C, after 60 min.

I (Ts19) at 30°C and 42°C (Fig. 2) or a wild-type enzyme I. It can be seen that both the rate of sucrose uptake and the maximal intracellular level of sucrose were temperature dependent in strain Ts19, indicating that the transport of sucrose occurs by way of the temperature-sensitive enzyme I. The thermostability of the cloned *sacP* gene product was tested by assaying sucrose transport in strain TGI, which contains wild-type enzyme I. The initial rate of sucrose uptake by strain TGI (pBSG8-12) was 6 nmol/min per mg of protein at 30°C and 8 nmol/min per mg of protein at 42°C, values that represent 30% and 40% of the rate of wild-type *B. subtilis*. The initial rate in strain Ts19 at 30°C is only of 1 nmol/min per mg of protein, showing that, even at permissive temperature, the thermosensitive enzyme I is less active than the wild-type enzyme I. The initial rate values indicate that the complementation is very efficient, which contrasts with results reported earlier for *in vitro* complementation experiments by Simoni *et al.* (19). Two reasons may explain this difference. The first one comes from the *in vivo* situation where intact membrane may stabilize proteins or reactive intermediates. The second one may be that the stoichiometry could be shifted in favor of the number of II^{Suc} molecules since the gene is on a multicopy plasmid. The rate of sucrose uptake by the control strain Ts19 (pUC9) corresponds to diffusion. If sucrose transport occurs by way of the PTS, the sugar would be expected to be phosphorylated. This was confirmed by analyzing the nature of the intracellular product

A B
$M_r \times 10^{-3}$
68
50
43
30
20.1
14.4
1 2 1 2

FIG. 3. Analysis of plasmid-encoded polypeptides. Circular plasmid DNA (1 μg) was used in our assay of plasmid-encoded gene products. Translational products were separated by $NaDodSO_4$/PAGE and were detected by autoradiography. (*A*) Lanes: 1, pUC9; 2, pBSG8-11. (*B*) Lanes: 1, pUC9; 2, pBSG8-12. The estimated molecular weights are indicated. Arrowheads indicate bands of approximate M_r values of 50,000 and 38,000–40,000.

```
                  MetAspTyrLysGluThrAlaLysArgLeuIleGluLeuLeuGlyGlyLysGluAsnIleIleSerAlaAla
TAAAAGGGGGAGAAATAGATGGATTACAAAGAGACTGCAAAACGCCTCATTGAGCTTCTCGGAGGGAAAGAAAATATTATCAGCGCGGCT
                  └─► sacP
HisCysAlaThrArgLeuArgLeuValMetLysAspGluSerLysIleAspGlnAlaGlnValGluGluLeuAspGlyValLysGlyAla
CATTGTGCAACAAGACTGCGTTTAGTGATGAAAGATGAATCAAAGATAGATCAAGCACAAGTAGAAGAGCTTGATGGCGTTAAAGGGGCT
      100
PheSerSerSerGlyGlnTyrGlnIleIlePheGlyThrGlyLeuValAsnLysValPheAspAlaPheSerLysGluAlaAspIleGlu
TTCAGCAGCTCTGGCCAGTACCAGATCATTTTCGGAACAGGCCTTGTGAATAAAGTATTCGATGCCTTTTCTAAAGAAGCTGATATCGAG
              200
ArgGluGluHisValAsnHisGlnAspAlaAlaLysGluLysLeuAsnProAlaAlaArgPheAlaLysThrLeuSerAsnIlePheVal
CGTGAAGAACACGTCAATCATCAGGATGCGGCAAAAGAAAAGCTGAATCCCGCTGCGAGATTTGCGAAAACGCTTTCTAATATTTTTGTT
                        300
ProIleIleProAlaIleValAlaSerGlyLeuLeuMetGlyLeuLeuGlyMetIleAsnAlaPheHisTrpMetSerLysAspSerAla
CCAATCATCCCAGCTATTGTAGCCAGCGGCCTATTAATGGGATTGCTGGGTATGATCAATGCGTTTCATTGGATGAGCAAGGATTCTGCG
                                 400
LeuLeuGlnLeuLeuAspMetPheSerSerAlaAlaPheIlePheLeuProIleLeuIleGlyValSerAlaSerLysGluPheGlySer
CTTTTGCAGCTGCTTGATATGTTTTCAAGTGCAGCATTCATTTTCTTGCCGATTCTAATCGGGGTCAGCGCTTCAAAAGAGTTTGGCAGC
                                          500
AsnProTyrLeuGlyAlaValIleGlyGlyIleMetIleHisProAsnLeuLeuAsnProTrpGlyLeuAlaGluGluGlnLeuIleThr
AATCCATACTTGGGAGCGGTCATCGGGGGAATCATGATTCATCCGAATCTTTTAAATCCATGGGGATTGGCGGAAGAACAACTGATTACA
                                                    600
CysIlePheSerAspLeuIleSerLeuPheSerAlaThrGlyAsnCysTyrProCysProAlaCysGlyValCysAspGluGlnGlyArg
TGCATCTTTTCGGATTTGATATCGCTCTTCTCGGCTACAGGGAACTGTTATCCCTGTCCTGCTTGCGGTGTATGTGATGAGCAAGGTAGA
                                                             700
GluMetAspGluLysSerGlySerThrCysGlyAspLeuLeuValThrProPheValThrValIleValThrGlyPheValAlaPheIle
GAAATGGACGAGAAAAGTGGTTCCACATGCGGTGATTTGCTTGTTACACCATTTGTTACTGTGATTGTCACCGGGTTTGTTGCTTTTATC
                                                                      800
AlaIleGlyProLeuGlyArgAlaLeuGlySerGlyIleThrValAlaLeuThrTyrValTyrAspHisAlaGlyPheValAlaGlyLeu
GCGATTGGCCCTCTAGGCAGAGCGCTTGGCTCCGGCATTACGGTTGCTTTAACTTATGTATATGACCATGCCGGTTTTGTCGCAGGTCTG
                                                                               900
IlePheGlyGlyThrTyrSerLeuIleValLeuThrGlyValHisHisSerPheHisAlaIleGluAlaGlyLeuIleAlaAspIleGly
ATTTTCGGGGGCACGTATTCACTCATCGTGCTGACAGGCGTTCATCACAGCTTCCATGCGATTGAGGCAGGGCTAATTGCAGATATAGGC

LysAsnTyrLeuLeuProIleTrpSerMetAlaAsnValAlaGlnGlyGlyAlaGlyLeuAlaValPhePheMetAlaLysLysAlaLys
AAAAACTACTTGCTGCCGATCTGGTCAATGGCGAATGTCGCACAGGGCGGGGCAGGTCTTGCCGTCTTCTTTATGGCGAAAAAGGCAAAG
      1000
ThrLysGluIleAlaLeuProAlaAlaPheSerAlaPheLeuGlyIleThrGluProValIlePheGlyValAsnLeuArgTyrArgLys
ACAAAAGAAATCGCGCTTCCGGCAGCATTTTCCGCTTTTCTCGGCATTACTGAGCCCGTCATATTCGGAGTCAATCTGCGCTACCGAAAA
              1100
ProPheIleAlaAlaMetIleGlyGlyAlaLeuGlyGlyAlaTyrValValPheThrHisValAlaAlaAsnAlaTyrGlyLeuThrGly
CCGTTTATCGCCGCGATGATCGGAGGCGCATTGGGTGGGGCATATGTCGTCTTTACACATGTAGCTGCAAACGCTTACGGTTAACGGGA
                       1200
IleProMetIleAlaIleAlaAlaProPheGlyPheSerAsnLeuIleHisTyrLeuIleGlyMetAlaIleAlaAlaValSerAlaPhe
ATTCCGATGATCGCCATTGCGGCTCCGTTTGGCTTCAGCAACTTGATTCACTATCTGATTGGCATGGCGATTGCGGCAGTGTCAGCGTTC
                                 1300
IleAlaAlaPheValMetLysIleAsnGluAspGluGluArgLysLys***
                                           MetThrAlaHisAspGlnGluLeuArgArgArgAlaTyrGlu
ATTGCAGCATTTGTCATGAAGATTAATGAGGATGAGGAGAGGAAAAAATGACAGCACATGACCAGGAGCTTCGTCGCCGGGCTTATGAAG
                                 1390      └─► sacA
GluValGluLysLysGluProIleAlaAsnSerAspProHisArgGlnHisPheHisIleMetProProValGlyLeuLeuAsnAspPro
AAGTGGAGAAAAAAGAGCCCATCGCTAACAGCGATCCGCACCGCCAGCATTTTCATATCATGCCGCCGGTTGGGCTGCTGAATGACCCG
                                                 1500
```

FIG. 4. Nucleotide sequence of a portion of the 2.8-kb *Sma* I fragment spanning *sacP* and showing the deduced amino acid sequence of the gene product and the N-terminal part of sucrase. The potential ribosome binding site of the *sacP* gene and the internal *Eco*RI site are underlined (thick line and thin line, respectively). The stop codon is symbolized by three asterisks. The *sacA* and the putative *sacP* initiation codons are indicated by arrows.

when transported into cells containing *sacP* but lacking *sacA* (Fig. 2*B*). Sucrose-6-phosphate was detected, which gives evidence that sucrose had been phosphorylated during transport. The presence of free sucrose was probably due to the presence of a nonspecific phosphatase as reported by Haguenauer and Képès (20).

The results presented above show that *E. coli* cells containing the *sacP* gene transport sucrose by way of a phosphotransferase system; the *sacP* gene product is therefore likely to be enzyme II of the sucrose PTS of *B. subtilis*. The II^{Suc} is most probably located in the inner membrane in *E. coli*, and sucrose presumably reaches the periplasm by way of diffusion through the outer membrane porins, such as OmpF and OmpC (21).

In order to estimate the M_r of II^{Suc}, the 2.8-kb *Sma* I fragment encoding *sacP* and the 454 amino-terminal amino acids of *sacA* was cloned in pUC9 to give pBSG8-11 (Fig. 1). pBSG8-11 and pBSG8-12 were transcribed and translated *in vitro* (Fig. 3). Three major additional bands were observed with pBSG8-11 compared to the pattern obtained after transcription and translation of pUC9: a doublet with an approximate M_r of 38,000–40,000 and a weak band of M_r of ≈50,000. The pattern obtained with pBSG8-12 showed the presence of only the additional doublet (M_r of 38,000–40,000). This indicated that the doublet (M_r of 38,000–40,000) is due to the expression of the *sacP* gene and that the *sacP* gene product has an apparent M_r of 38,000–40,000. The presence of a doublet rather than a single band could be due to translation reinitiation at two ATG codons.

Nucleotide Sequence of the *sacP* Gene. The complete nucleotide sequence of the *sacP* gene is shown in Fig. 4. The 1.3-kb *Sma* I–*Eco*RI fragment was subcloned in both orientations in M13mp19, and 95% of the fragment was sequenced on both strands. A 750-base-pair DNA fragment overlapping the internal *Eco*RI site was cloned, and deleted phages enabled sequencing through the *Eco*RI site. The sequence data obtained were compiled by computer using the DB system of Staden (22) to give a continuous overlapping sequence. The sequence shown in Fig. 4 starts about 50 nucleotides from the *Sma* I site and harbors the region containing the entire *sacP* gene and the first 131 nucleotides of the *sacA* gene. The sequence 3′ to the *Eco*RI site (i.e., the last 270 nucleotides) has already been published (4).

The sequence was then examined for open reading frames in the six possible frames. On one strand, stop codons were found regularly in the three possible frames. In the other orientation, stop codons regularly interrupted two phases, and in the last frame, an open reading frame was found (starting at position 19 with an ATG start codon and ending at position 1398 at a TGA stop codon). In the same frame (at position 118) there is another ATG codon.

Seven nucleotides upstream of the first ATG codon (position 19) is a potential ribosome binding site (AAAGGGGGA) (23). The calculated free energy of interaction of this sequence with the 3′ end of *B. subtilis* 16S rRNA is $\Delta G = -20.8$ kcal/mol (-87 kJ/mol) (24). The distance between this putative ribosome binding site and the ATG is 11 nucleotides, which is in good agreement with that usually found in *B. subtilis* (mean value of 10.6) (25). These results strongly suggest that this ATG is the *in vivo* initiation codon.

The nucleotide sequences encoding II^{Suc} and sucrase overlap (Fig. 4) at the sequence ATGA, where ATG is the initiation codon of the sucrase gene and TGA is the stop codon of the II^{Suc} open reading frame. This result supports the hypothesis that *sacP* and *sacA* are part of an operon in which *sacA* is distal to *sacP* (4). Postma and Lengeler (28) pointed out that the structural genes for the enzymes II are not part of a *pts* operon or regulon but form units with the corresponding metabolic enzymes. The existence of translational coupling in *B. subtilis* has been demonstrated, as has the optimization of such coupling by the sequence ATGA (26, 27).

Comparison of II^{Suc} Sequence with That of Other Known Enzymes II. The open reading frame defined by the nucleo-

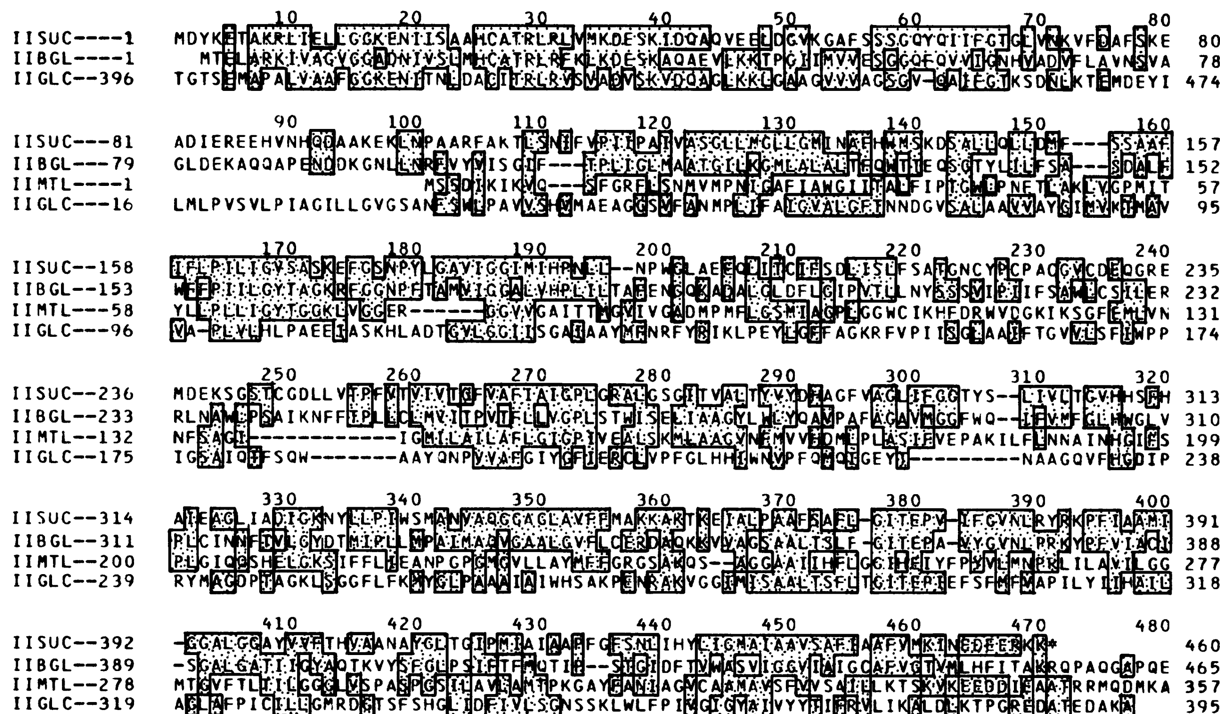

FIG. 5. Comparison of *B. subtilis* II^{Suc} with *E. coli* II^{Bgl}, *S. typhimurium* II^{Mtl}, and *E. coli* II^{Glc}. The amino acid sequences of the four polypeptides are given in the one-letter code and have been aligned by introducing gaps (hyphens) to maximize identities. Identical and similar residues are boxed and stippled (accepted conservative replacements are I, L, V, and M; D and E; A and G; R and K; S and T; and F and Y). Numbers at the ends of the lines indicate the number of the nearest residue in the sequence. Numbers above the lines indicate the position number; 1 corresponds to the N-terminal methionine from II^{Suc}.

tide sequence is 1380 nucleotides long and encodes a polypeptide of 460 residues with a calculated M_r of 48,945. This value is somewhat higher than that determined in the *in vitro* experiment by $NaDodSO_4$/PAGE (38,000–40,000). However, other sequenced enzymes II (29–32) also show lower apparent M_r values when analyzed by $NaDodSO_4$/PAGE than is indicated by the sequence data.

The amino acid sequence of the II^{Suc} was compared to that of three other enzymes II—namely, mannitol-specific enzyme II (II^{Mtl}) from *Salmonella typhimurium* (29), glucose-specific enzyme II (II^{Glc}) from *E. coli* (30), and II^{Bgl} from *E. coli* (31). The alignments are shown in Fig. 5.

II^{Suc} and II^{Glc} are 460 and 477 residues long, respectively, whereas II^{Bgl} and II^{Mtl} are 625 and 637 residues long, respectively. Gaps have been introduced to optimize alignment.

When II^{Suc} and II^{Glc} were compared, it appeared that the C-terminal part of II^{Glc} shared similarity with the N-terminal part of II^{Suc}, and a weaker similarity could be displayed encompassing the remaining parts of the sequences.

The greatest similarity is between II^{Suc} and II^{Bgl} (most of the gaps introduced in these two proteins were at the same positions, in order to allow comparison with II^{Mtl}). These two enzymes II, therefore, share homology, and their genes most probably derive from a common ancestor. The same conclusion may probably be drawn for the genes for the other enzymes II because of the blocks of homology found, although the similarity is not as great.

Indeed, the comparison between the four proteins shows four major blocks of similarity (Fig. 6). One is located in the N-terminal region of two proteins (II^{Suc} and II^{Bgl}) and in the C-terminal part of II^{Glc}, a second is near position 160, a third is near position 270, and the last one is near position 370. The last three regions have previously been described by Bramley and Kornberg (31). The four proteins share common properties, such as interactions with the cytoplasmic membrane and the phosphorylated histidine-containing phosphocarrier protein as well as transport and phosphorylation of a sugar. This may explain the regions of sequence conservation.

In conclusion, the results presented here demonstrate that *sacP* codes for *B. subtilis* II^{Suc}, which shares homology with enzymes II from enteric bacteria. The *sacP* and *sacA* genes are probably organized in an operon structure in which the promoter is expected to be located 5′ to *sacP*. The regulatory locus *sacT*, which is contiguous to *sacP*, is probably located in the vicinity of the promoter. Cloning the entire operon would provide useful tools to study the regulation of the gene expression in the *B. subtilis* sucrose regulon.

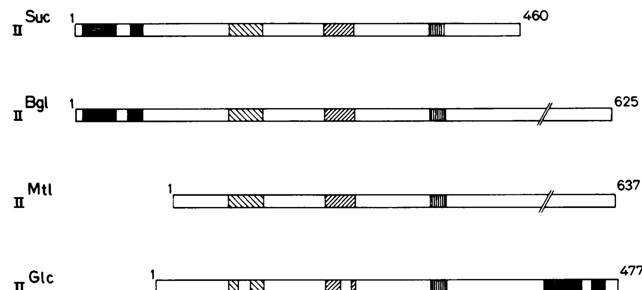

FIG. 6. Schematic representation of the similarities in the amino acid sequences of the polypeptides shown in Fig. 5. Hatched areas correspond to conserved regions. Numbers indicate the number of amino acid residues in the protein.

We thank Dr. R. Dedonder in whose laboratory this work was carried out. We are indebted to Dr. H. de Reuse and Dr. M. Débarbouillé for the gifts of strain Ts19 and the *B. subtilis* DNA bank, respectively. The authors wish to thank Dr. A. Edelman and Dr. G. N. Cohen for critical reading of the manuscript. This work was supported by research grants from Centre National de la Recherche Scientifique and from Fondation pour la Recherche Médicale.

1. Lepesant, J.-A., Kunst, F., Pascal, M., Kejzlarova-Lepesant, J., Steinmetz, M. & Dedonder, R. (1976) in *Microbiology 76*, ed. Schlessinger, D. (Am. Soc. Microbiol., Washington, DC), pp. 58–69.
2. Kunst, F., Steinmetz, M., Lepesant, J.-A. & Dedonder, R. (1977) *Biochimie* **59**, 287–292.
3. Fouet, A., Klier, A. & Rapoport, G. (1982) *Mol. Gen. Genet.* **186**, 399–404.
4. Fouet, A., Klier, A. & Rapoport, G. (1986) *Gene* **45**, 221–225.
5. Lepesant, J.-A. & Dedonder, R. (1968) *C.R. Acad. Sci.* **267**, 1109–1112.
6. Gibson, T. J. (1984) Dissertation (Cambridge Univ., Cambridge, U.K.).
7. Messing, J. (1983) *Methods Enzymol.* **101**, 20–77.
8. Sullivan, M. A., Yasbin, R. E. & Young, F. E. (1984) *Gene* **29**, 21–26.
9. Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene* **33**, 103–119.
10. Bolivar, F. & Backman, K. (1979) *Methods Enzymol.* **68**, 245–267.
11. Bourd, G. I., Bolshakova, T. N., Saprikina, T. P., Klyutchova, V. V. & Gershanovitch, V. N. (1971) *Mol. Biol. (Moscow)* **5**, 384–389.
12. Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res.* **7**, 1513–1523.
13. Rapoport, G., Klier, A., Billault, A., Fargette, F. & Dedonder, R. (1979) *Mol. Gen. Genet.* **176**, 239–245.
14. Lepesant, J.-A., Kunst, F., Lepesant-Kejzlarova, J. & Dedonder, R. (1972) *Mol. Gen. Genet.* **118**, 135–160.
15. Delobbe, A., Haguenauer, R. & Rapoport, G. (1971) *Biochimie* **53**, 1015–1021.
16. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
17. Dale, R. M. K., McClure, B. A. & Houchins, J. P. (1985) *Plasmid* **13**, 31–40.
18. Débarbouillé, M., Kunst, F., Klier, A. & Rapoport, G. (1987) *FEMS Microbiol. Lett.* **41**, 137–140.
19. Simoni, R. D., Nakazawa, T., Hays, J. B. & Roseman, S. (1973) *J. Biol. Chem.* **248**, 932–940.
20. Haguenauer, R. & Képès, A. (1972) *Biochimie* **54**, 505–512.
21. Nakae, T. (1976) *J. Biol. Chem.* **251**, 2176–2178.
22. Staden, R. (1982) *Nucleic Acids Res.* **10**, 4731–4751.
23. Shine, J. & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 1342–1346.
24. Tinoco, I., Jr., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. & Gralla, J. (1973) *Nature (London) New Biol.* **246**, 40–41.
25. Hager, P. W. & Rabinowitz, J. C. (1985) in *The Molecular Biology of the Bacilli*, ed. Dubnau, D. (Academic, New York), Vol. 2, pp. 1–32.
26. Zaghloul, T. I., Kawamura, F. & Doi, R. H. (1985) *J. Bacteriol.* **164**, 550–555.
27. Sprengel, R., Reiss, B. & Schaller, H. (1985) *Nucleic Acids Res.* **13**, 893–909.
28. Postma, P. W. & Lengeler, J. W. (1985) *Microbiol. Rev.* **49**, 232–269.
29. Lee, C. A. & Saier, M. H., Jr. (1983) *J. Biol. Chem.* **258**, 10761–10767.
30. Erni, B. & Zanolari, B. (1986) *J. Biol. Chem.* **261**, 16398–16403.
31. Bramley, H. F. & Kornberg, H. L. (1987) *J. Gen. Microbiol.* **133**, 563–573.
32. Bouma, C. L., Meadow, N. D., Stover, E. W. & Roseman, S. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 930–934.