ELSEVIER

Gene 175 (1996) 59–63

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

# Analysis of DNA flanking the *treA* gene of *Bacillus subtilis* reveals genes encoding a putative specific enzyme II$^{Tre}$ and a potential regulator of the trehalose operon

Frieder Schöck, Michael K. Dahl *

*Lehrstuhl für Mikrobiologie, Institut für Mikrobiologie, Biochemie und Genetik, Friedrich-Alexander-Universität Erlangen-Nürnberg, Staudtstr. 5, 91058 Erlangen, Germany*

Received 9 October 1995; revised 2 January 1996; accepted 5 January 1996

**Abstract**

Nucleotide sequencing revealed the genes *treP* encoding a putative specific enzyme II$^{Tre}$ upstream from *treA* and *treR* encoding a potential regulator downstream from *treA* of *Bacillus subtilis* 168. The *treP* gene encodes a 470-amino acid (aa) protein (50 kDa) showing high similarities to several different specific enzymes II of phospho*enol*pyruvate-dependent phosphotransferase systems. *treR* encodes a 238-aa protein (28 kDa) with high homologies in its N-terminal part to DNA-binding proteins including a helix-turn-helix motif. Homologies in its C-terminal part place it in the family of FadR-GntR transcriptional regulators. The three genes, *treP-treA-treR*, are probably organized in one operon expressed by a $\sigma^A$-dependent promoter 53 bp upstream from *treP* and a $\rho$-independent terminator 28 bp downstream from *treR*.

*Keywords:* Sugar transport; PTS; Disaccharides; Transcription; Repressor

## 1. Introduction

The utilization of the abundant carbon source trehalose by *Bacillus subtilis* involves the transport across the cytoplasmic membrane and an intracellular phospho-α-(1,1)-glucosidase (TreA) hydrolyzing trehalose 6-phosphate into glucose 6-phosphate and glucose (Kennett and Sueoka, 1971; Helfert et al., 1995; Gotsche and Dahl, 1995). TreA is also able to cleave *p*-nitrophenyl α-D-glucopyranoside (PNPG) in vitro (Kennett and Sueoka, 1971; Helfert et al., 1995). TreA activity is induced by trehalose and subject to catabolite repression (CR) (Helfert et al., 1995). CR in *B. subtilis* is in general mediated by a *cis*-acting element CRE (Weickert and Chambliss, 1990; Grundy et al., 1994), a *trans*-acting element CcpA (Henkin et al., 1991; Hueck et al., 1994) and a component of the phospho*enol*pyruvate:sugar phosphotransferase system HPr (Deutscher et al., 1994) (for a recent review see Hueck and Hillen, 1995). Several genes underlying CR contain a CRE located within 200 bp upstream or downstream from the promoter (Hueck et al., 1994). However, a *ccpA* deletion has no obvious effect on CR by glucose or mannitol on TreA activity, whereas CR by fructose is abolished (Helfert et al., 1995).

The aim of this study was to clone and sequence the entire trehalose operon. The putative function of the encoded genes is proposed on the basis of sequence comparisons. The data suggest two additional genes, *treP* and *treR*, encoding a potential specific enzyme II$^{Tre}$ and the putative regulator, respectively.

* Corresponding author. Tel. +49 9131 858094; Fax +49 9131 858082; e-mail: mdahl@biologie.uni-erlangen.de

Abbreviations: aa, amino acid(s); *B.*, *Bacillus*; bp, base pair(s); CcpA, catabolite controlling protein A; Cm, chloramphenicol; CR, catabolite repression; CRE, catabolite responsive element; *E.*, *Escherichia*; GCG, Genetics Computer Group (Madison, WI, USA); I, identity; IEP, isoelectric point; kb, kilobase(s) or 1000 bp; $M_r$, molecular weight; nt, nucleotide(s); ORF, open reading frame; PNPG, *p*-nitrophenyl α-D-glucopyranoside; S, similarity; SD, Shine-Dalgarno; TM, transmembrane (domain); *tre* or Tre, trehalose; *treA*, gene encoding phospho-α-(1,1)-glucosidase of *B. subtilis*; *treB*, gene encoding the enzyme II$^{Tre}$ of *E. coli*; *treP*, gene encoding the putative enzyme II$^{Tre}$ of *B. subtilis*; *treR*, gene encoding the potential regulator of the *tre* operon; wt, wild type.

## 2. Experimental and discussion

### 2.1. Cloning of DNA containing treP and treR

We used a *B. subtilis* 168 genomic library described previously (Helfert et al., 1995). 2 μg DNA from the


*PII* S0378-1119(96)00120-5

genomic library was transformed to *B. subtilis* 1A167 (*tre-12*) which exhibits a trehalose minus phenotype (Kennett and Sueoka, 1971). Transformants were selected on minimal medium plates containing 10 mM trehalose and kanamycin (5 μg/ml) (Helfert et al., 1995). One of the clones carried a plasmid of 12 kb (designated pCH1, Helfert et al., 1995), harboring an insert of 5 kb and was used as template in sequencing reactions using synthesized oligonucleotides as specific primers to determine *treP* and sequences located upstream. Furthermore a 2018 bp *Nsi*I fragment of pCH1 was cloned into the compatible *Pst*I site of pBluescript SK$^-$ (Stratagene, La Jolla, CA, USA) yielding plasmid pSG3, which contains the *treR* gene and downstream sequences. This construct was used as template to determine the *treR* sequence, using synthesized oligonucleotides.

### 2.2. Nt sequence and analysis

The nt sequence of the *tre* operon was determined on both strands (Fig. 1, Helfert et al., 1995) and analyzed by the GCG program. The sequence upstream from *treA* contained one complete ORF of 1410 bp starting with an ATG start codon at nt position 167 in Fig. 1 ending with a TAA stop codon at position 1579 (corresponding to nt position 444 in Helfert et al., 1995). The ATG codon is preceded at an appropriate distance by the putative ribosome-binding site GGAGG between nt positions 155 and 159. The sequence downstream from *treA* contained an ORF of 714 bp starting with ATG (position 3352 in Fig. 1, corresponding to nt position 2217 in Helfert et al., 1995), preceded by the putative ribosome-binding site GGTGG, ending with a stop codon at position 4068. It is followed by a $\rho$-independent terminator-like structure. The sequences 5′-TACAAT-3′ (−10 region) 53 bp upstream from the start codon of *treP* and 5′-TTGACT-3′ (−35 region) 76 bp upstream match in each region with one mismatch to the consensus promoter sequence of $\sigma^A$ (Moran et al., 1982) and have an ideal spacing of 17 bp. This is in agreement with the highest TreA activity found at the end of the exponential growth phase (Helfert et al., 1995). No transcription terminator-like structures were identified between *treP* and *treA* or *treA* and *treR*. This suggests that all three genes may be located within an operon and that the promoter is located upstream from *treP*.

The *treR* ORF encodes a protein of 238 aa residues (Fig. 1) with a calculated $M_r$ of 28 kDa and a calculated IEP of 7.3. The translated sequence of *treR* has identities to the FadR-GntR family of transcriptional regulators (Haydon and Guest, 1991) (Table 1). Using the TreR aa sequence from *E. coli* (sequence taken from Horlacher et al. (1994) from EMBL data library accession No. U07790/U14003) in an alignment with TreR from *B. subtilis*, no significant homologies or homologous regions were detected (data not shown).

The *treP* ORF encodes a protein of 470 aa residues (Fig. 1) with a calculated $M_r$ of 50 kDa and a calculated IEP of 8.9. The deduced aa sequence revealed high similarities to several enzymes II (a comparison with the two enzymes of highest homology is shown in Fig. 2). TreP is highly similar to TreB of *E. coli*, which has been identified as the specific enzyme II$^{Tre}$ (Boos et al., 1990; Klein et al., 1995). Interestingly, a higher similarity was found to enzyme II$^{Scr}$ of *Staphylococcus xylosus* (Wagner et al., 1993). This may indicate that TreP could also transport sucrose. Nevertheless, these similarities indicate that TreP belongs to the family of enzyme II, the permeases of phospho*enol*pyruvate-dependent phosphotransferase systems.

### 2.3. Putative structure of the predicted TreP protein

Two highly conserved aa sequence motifs typical for enzymes II suggested to be involved in the phosphoryl transfer have also been identified in TreP. The CATRLR aa motif is typically found in the B domain of enzyme II, where the cysteine reflects the phosphorylation site (Fig. 2) (Nuoffer et al., 1988). A second conserved motif (GITE) is typically localized in the C domain which is also important for phosphoryl transfer to the substrate (Lengeler, 1990; Ruijter et al., 1992). Investigations using a series of *lacZ* and *phoA* fusions to *ptsG* of *E. coli* have revealed that the enzyme contains eight TM regions (Buhr and Erni, 1993). These regions are also highly conserved in TreP (Fig. 2) and reflect those regions of hydrophobic character (Fig. 3) implying an analogous structure as proposed for PtsG of *E. coli*. A potential enzyme IIA domain as observed for enzyme II$^{Glc}$ in *B. subtilis* (Gonzy-Tréboul et al., 1991) is missing in TreP. These results show that TreP belongs to the enzyme IIBC type.

### 2.4. Possible role of TreR

TreA expression is under control of trehalose induction (Helfert et al., 1995). This indicates the existence of a regulatory protein involved in the regulation of transcription of the *tre* operon. The *treR* gene product is one candidate which might regulate *tre* operon expression since it exhibits a putative helix-turn-helix motif in its N-terminal part, homologous to several DNA-binding proteins. The nature of its function, repressor or activator, is not clear with the current sequence data alone. However, its homology to the FadR-GntR family (Table 1) implicates a function as a regulator.

### 2.5. Potential CRE

Since TreA activity is under CR we have searched for potential CRE sequences by homology to the

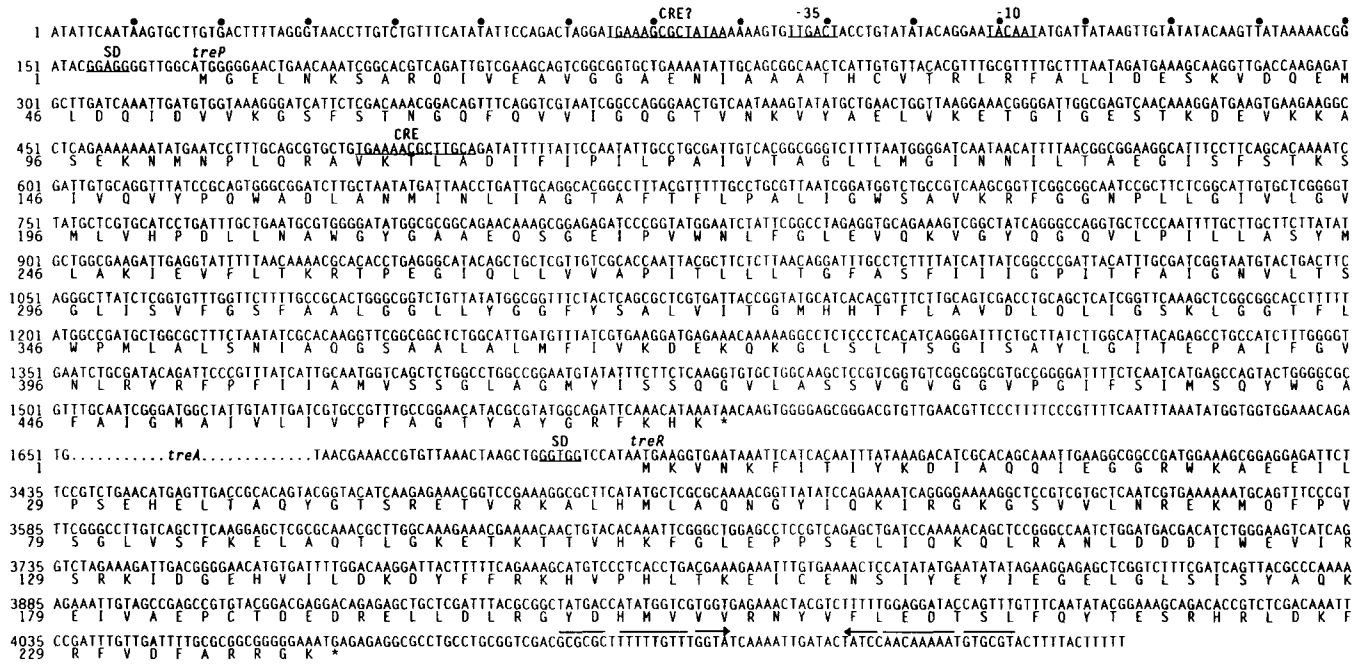

Fig. 1. Nt sequence of the *B. subtilis treP* and *treR* genes. The sequence presented contains the potential promoter, *treP* and downstream sequences up to position 1652 (corresponding to the ATG start codon of *treA* previously published by Helfert et al., 1995) and in addition *treR*, beginning with the stop codon of *treA* at position 3316 of the putative *tre* operon harboring the potential SD. A potential $\sigma^A$ promoter sequence (−35 (TTGACT) and −10 (TACAAT) regions) and the potential ribosome-binding sites (SD) are underlined. The potential $\rho$-independent terminator is indicated by arrows. The aa sequences derived from the nt sequences are depicted in one-letter code. Each stop codon of *treP* and *treR* is indicated by an asterisk. Potential CREs are underlined and designated. These data have been submitted to the EMBL Data Library under accession No. Z54245.

Table 1
Identities and similarities of TreR to the GntR-FadR family of regulators

| Regulators | Identities (%) | Similarities (%) |
|---|---|---|
| FadR | 27 | 51 |
| GenA | 22 | 49 |
| GntR | 19 | 50 |
| $HutC_{Ka}$ | 21 | 47 |
| $HutC_{Pp}$ | 24 | 48 |
| KorA | 18 | 43 |
| P30 | 28 | 49 |
| PhnF | 22 | 49 |

A comparison of the listed regulators with each other, the definition of the FadR-GntR family of regulators and the citations of the original publications is presented by Haydon and Guest (1991).

proposed consensus sequence (Hueck et al., 1994). One CRE matching the consensus was found within the coding sequence of *treP*, 319 bp downstream from the start codon. The localization of a functional CRE within the first structural gene of an operon has also been described for the *hut* (Wray et al., 1994) and *xyl* (Kraus et al., 1994) operons. Another potential CRE is localized upstream from the putative promoter at positions 65–78 (Fig. 1), with one mismatch at position 13 of the proposed consensus sequence. Previous studies with mutated CRE (Weickert and Chambliss, 1990; Martin-Verstraete et al., 1995) showed that this position is critical for a full function. As we have previously shown (Helfert et al., 1995), a strain carrying a *ccpA*::Tn917 insertion still exhibits CR of TreA activity by glucose and mannitol, but not by fructose. Since CcpA is supposed to interact with CRE (Fujita et al., 1995; Henkin et al., 1991; Kim et al., 1995) mediating CR one of the two or both potential CREs of the *tre* operon are involved in CR at the transcriptional level.

## 3. Conclusions

The high homology of TreP with several enzymes II, its suggested topology and its localization within the putative *tre* operon provide evidence for its function as the specific trehalose transporter.

TreR belongs to the FadR-GntR family and is a potential regulator of the *tre* operon.

The three genes *treP*, *treA* and *treR* are probably organized in the *tre* operon which is presumably expressed by a $\sigma^A$-dependent promoter and terminated by a $\rho$-independent terminator.

Two potential CREs are present in the *tre* operon which could be involved in CR.

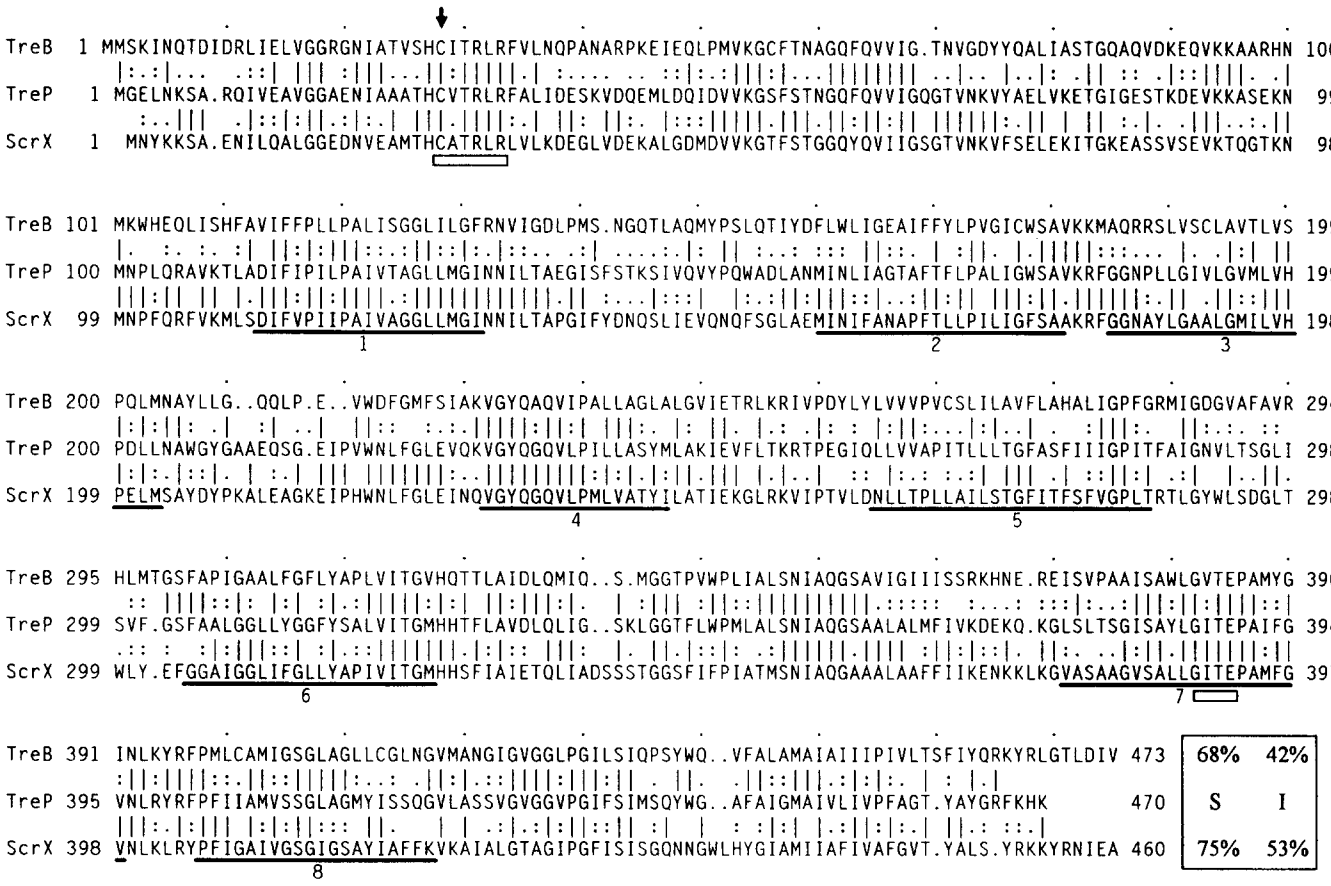


Fig. 2. Homologies of enzymes II. Presented are the similarities (dots between rows) and identities (lines between rows) of TreP of *B. subtilis* to TreB of *E. coli* (Klein et al., 1995) and to ScrA of *Staphylococcus xylosus* (Wagner et al., 1993) as obtained by sequence analysis using the GCG program. The percentage of similarities (S) and identities (I) to TreP is depicted in a box at the end of the last row. The arrow indicates the cysteine in domain B which is highly conserved in enzymes II and is phosphorylated during transport (Nuoffer et al., 1988). The open boxes indicate conserved regions which are thought to be important for the phosphoryl transfer to the incoming sugar (Lengeler, 1990; Ruijter et al., 1992). The lines numbered 1–8 indicate those regions which are supposed to be potential membrane spanning regions, determined in the PtsG protein of *E. coli* by a series of *phoA* and *lacZ* fusions (Buhr and Erni, 1993).

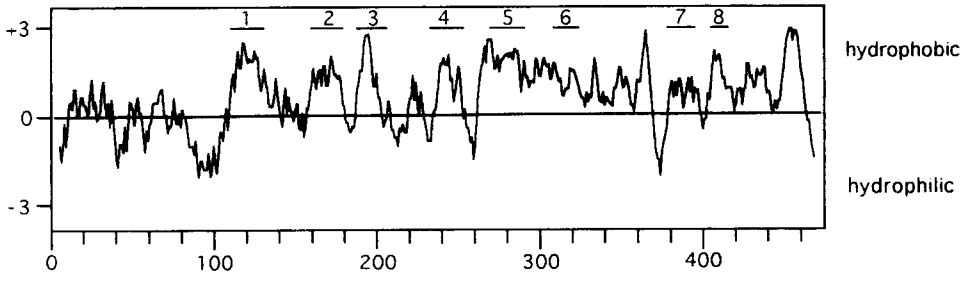


Fig. 3. Hydrophobicity profile of the predicted TreP protein. Hydropathy values for a window of 9 aa were calculated by the method of Kyte and Doolittle (1982) using the GCG program. The solid lines (numbered 1–8 corresponding to the underlined regions in Fig. 2) indicate the regions most likely to form TM α-helices. Hydrophobic regions correspond to the positive index numbers.

## Acknowledgement

We are grateful to U. Ehmann, V. Helbl and E. Küster for their interest in this work and many helpful discussions. This work was carried out in the laboratory of W. Hillen whose continuous interest in the project, his support and many helpful discussions are greatly appreciated. Financial support was obtained from the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

## References


Boos, W., Ehmann, U., Forkl, H., Klein, W., Rimmele, M. and Postma, P. (1990) Trehalose transport and metabolism in *Escherichia coli*. J. Bacteriol. 172, 3450–3461.

Buhr, A. and Erni, B. (1993) Membrane topology of the glucose transporter of *Escherichia coli*. J. Biol. Chem. 268, 11599–11603.

Deutscher, J., Reizer, J., Fischer, C., Galinier, A., Saier Jr., M.H. and Steinmetz, M. (1994) Loss of protein kinase-catalyzed phosphorylation of HPr, a phosphocarrier protein of the phosphotransferase system, by mutation of the *ptsH* gene confers catabolite repression resistance to several catabolic genes of *Bacillus subtilis*. J. Bacteriol. 176, 3336–3344.

Fujita, Y., Miwa, Y., Galinier, A. and Deutscher, J. (1995) Specific recognition of the *Bacillus subtilis gnt cis*-acting catabolite-responsive element by a protein complex formed between CcpA and seryl-phosphorylated Hpr. Mol. Microbiol. 17, 953–960.

Gonzy-Tréboul, G., De Waard, J.H., Zagorec, M. and Postma, P.W. (1991) The glucose permease of the phosphotransferase system of *Bacillus subtilis*: evidence for $II^{Glc}$ and $III^{Glc}$ domains. Mol. Microbiol. 5, 1241–1249.

Gotsche, S. and Dahl, M.K. (1995) Purification and characterization of the phospho-α-(1,1)-glucosidase (TreA) of *Bacillus subtilis* 168. J. Bacteriol. 177, 2721–2726.

Grundy, F.J., Turinsky, A.J. and Henkin, T.M. (1994) Catabolite regulation of *Bacillus subtilis* acetate and acetoin utilization genes by CcpA. J. Bacteriol. 176, 4527–4533.

Haydon, D.J. and Guest, J.R. (1991) A new family of bacterial regulatory proteins. FEMS Lett. 79, 291–296.

Helfert, C., Gotsche, S. and Dahl, M.K. (1995) Cleavage of trehalose-phosphate in *Bacillus subtilis* is catalyzed by a phospho-α-(1,1)-glucosidase encoded by the *treA* gene. Mol. Microbiol. 16, 111–120.

Henkin, T.M., Grundy, F.J., Nicholson, W.L. and Chambliss, G.H. (1991) Catabolite repression of α-amylase gene expression in *Bacillus subtilis* involves a *trans*-acting gene product homologous to the *Escherichia coli lacI* and *galR* repressors. Mol. Microbiol. 5, 575–584.

Horlacher, R. et al. (1994) EMBL data library accession No. U07790/U14003.

Hueck, C. and Hillen, W. (1995) Catabolite repression in *Bacillus subtilis*: a global regulatory mechanism for the Gram-positive bacteria? Mol. Microbiol. 15, 395–401.

Hueck, C., Hillen, W. and Saier, M.H. (1994) Analysis of a *cis*-acting sequence mediating catabolite repression in Gram-positive bacteria. Res. Microbiol. 145, 503–518.

Kennett, R.H. and Sueoka, N. (1971) Gene expression during outgrowth of *Bacillus subtilis* spores. J. Mol. Biol. 60, 31–41.

Kim, J.H., Guvener, Z.T., Cho, J.Y., Chung, K.-C. and Chambliss, G.H. (1995) Specificity of DNA binding activity of the *Bacillus subtilis* catabolite control protein CcpA. J. Bacteriol. 177, 5129–5134.

Klein, W., Horlacher, R. and Boos, W. (1995) Molecular analysis of *treB* encoding the *Escherichia coli* enzyme II specific for trehalose. J. Bacteriol. 177, 4043–4052.

Kraus, A., Hueck, C., Gärtner, D. and Hillen, W. (1994) Catabolite repression of the *Bacillus subtilis xyl* operon involves a *cis* element functional in the context of an unrelated sequence, and glucose exerts additional *xylR*-dependent repression. J. Bacteriol. 176, 1738–1745.

Kyte, J. and Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105–132.

Lengeler, J.W. (1990) Molecular analysis of the enzyme II-complexes of the bacterial phosphotransferase system (PTS) as carbohydrate transport systems. Biochim. Biophys. Acta 1018, 155–159.

Martin-Verstraete, I., Stülke, J., Klier, A. and Rapoport, G. (1995) Two different mechanisms mediate catabolite repression of the *Bacillus subtilis* levanase operon. J. Bacteriol. 177, 6919–6927.

Moran, C.P., Jr., Lang, N., LeGrice, S.F.J., Lee, G. Stephens, M. Sonenshein, A.L., Pero, J. and Losick, R. (1982) Nucleotide sequence that signals the initiation of transcription and translation in *Bacillus subtilis*. Mol. Gen. Genet. 186, 339–346.

Nuoffer, C., Zanolari, B. and Erni, B. (1988) Glucose permease of *Escherichia coli*: the effect of cysteine to serine mutations on the function, stability and regulation of transport and phosphorylation. J. Biol. Chem. 263, 6647–6655.

Ruijter, G.J.G., Van Meurs, G., Verwey, M.A., Postma, P.W. and Van Dam, K. (1992) Analysis of mutations that uncouple transport from phosphorylation in enzyme $II^{Glc}$ of the *Escherichia coli* phosphoenolpyruvate-dependent phosphotransferase system. J. Bacteriol. 174, 2843–2850.

Wagner, E., Götz, F. and Brückner, R. (1993) Cloning and characterization of the *scrA* gene encoding the sucrose specific enzyme II of the phosphotransferase system from *Staphylococcus xylosus*. Mol. Gen. Genet. 241, 33–41.

Weickert, M.J. and Chambliss, G.H. (1990) Site-directed mutagenesis of catabolite repression operator sequence in *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA 87, 6238–6242.

Wray, L.V., Pettengill, F.K. and Fisher, S.H. (1994) Catabolite repression of *Bacillus subtilis hut* operon requires a *cis*-acting site located downstream of the transcription initiation site. J. Bacteriol. 176, 1894–1902.