The *glpP* and *glpF* genes of the glycerol regulon in *Bacillus subtilis*

LENA BEIJER,* RUNE-PÄR NILSSON, CHRISTINA HOLMBERG and LARS RUTBERG

Department of Microbiology, University of Lund, Sölvegatan 21, S-223 62 Lund, Sweden

(Received 3 August 1992; revised 5 October 1992; accepted 14 October 1992)

The *Bacillus subtilis* *glpP*FKD region contains genes essential for growth on glycerol or glycerol 3-phosphate (G3P). The nucleotide sequence of *glpP* encoding a regulatory protein and the previously unidentified *glpF* encoding the glycerol uptake facilitator was determined. *glpF* is located immediately upstream of *glpK* and the two genes were shown to constitute one operon which is transcribed separately from *glpP*. A σ A-type promoter and the transcriptional start point for *glpFK* were identified. In the 5' untranslated leader sequence (UTL) of *glpFK* mRNA a conserved inverted repeat is found. The repeat is believed to be involved in the control of expression of *glpFK* by termination/antitermination of transcription, a control mechanism previously suggested for the regulation of *glpD* encoding G3P dehydrogenase. Expression of *glpFK* and *glpD* requires the inducer G3P and the *glpP* gene product. A 2.9 kb *B. subtilis* chromosomal DNA fragment containing the *glpP* open reading frame was cloned to give plasmid pLUM7. pLUM7 contains a functional *glpP* gene as shown by its ability to complement various *glpP* mutants. Immediately upstream of *glpP* an open reading frame is found (ORF1). Disrupting ORF1 by plasmid integration in the *B. subtilis* chromosome does not affect the ability to grow on glycerol as sole carbon and energy source. With the present report all *B. subtilis* *glp* genes located at 75° on the chromosomal map have been identified.

Introduction

The *Bacillus subtilis* *glp* regulon contains genes specific for growth on glycerol or glycerol 3-phosphate (G3P) as sole carbon and energy source. Four *glp* genes have been identified. These genes are *glpT* encoding a G3P permease (Lindgren, 1978; R.-P. Nilsson and others, unpublished), *glpK* encoding glycerol kinase, *glpD* encoding G3P dehydrogenase and *glpP* encoding a regulatory protein (Lindgren & Rutberg, 1974, 1976). The *glpT* gene maps at 15° on the *B. subtilis* chromosomal map whereas the other three *glp* genes map at 75°. The *glpK*, *glpD* and most likely also the *glpT* gene are induced by glycerol or G3P, G3P being the true inducer. Induction of these genes also requires the *glpP* gene product and induction is repressed by glucose (Lindgren & Rutberg,

1974). The mechanism of induction of *glpD* has been studied in some detail. The *glpD* gene constitutes a monocistronic operon which is transcribed from a constitutive, glucose-insensitive promoter, but a full length *glpD* transcript is only found in induced cells. An inverted repeat is located in the 5' untranslated leader sequence (UTL) of *glpD* mRNA and this repeat is suggested to act as a transcriptional stop signal. Production of a full length *glpD* mRNA is thought to be controlled by termination/antitermination of a transcript initiated at the constitutive *glpD* promoter. In the presence of G3P, the *glpP* gene product is thought to enhance transcription across and beyond the inverted repeat (Holmberg & Rutberg, 1991, 1992).

We have previously reported the cloning of a 7 kb *EcoRI*–*PstI* *B. subtilis* chromosomal DNA fragment which carries wild type alleles of *glpP*, *glpK* and *glpD* mutations, and the nucleotide sequences of *glpK* and *glpD* have been presented (Holmberg *et al.*, 1990). The plasmid carrying the 7 kb fragment is a derivative of plasmid pHP13 and is called pLUM30. Here we report the nucleotide sequences of the two additional *glp* genes contained in pLUM30, *glpP* and the previously unidentified *glpF*. The latter gene is proposed to encode a membrane protein which facilitates diffusion of glycerol

* Author for correspondence. Tel. 46 108631; fax 46 157839.

Abbreviations: Ap, ampicillin; Cm, chloramphenicol; C23O, catechol-2,3-dioxygenase; Em, erythromycin; G3P, glycerol 3-phosphate; Pm, phleomycin.

The nucleotide sequence data reported in this paper have been submitted to GenBank and assigned the accession number M99611.

across the cell membrane. Transcriptional analysis has revealed that *glpP* and *glpFK* represent two separate operons. The presence of a conserved inverted repeat in the UTL of *glpFK* mRNA indicates that expression of these genes is controlled by a mechanism similar to that proposed for *glpD*. With the present report all *B. subtilis glp* genes located at 75° on the chromosomal map have been identified.

Methods

Strains and plasmids. The bacterial strains and plasmids used in this work are listed in Table 1. In LUG0402 a chromosomal DNA fragment starting at position 88 and ending at position 1234 (Fig. 2) is deleted and replaced with a DNA fragment from pUB110 containing the *ble* gene, using methods described by Fridén *et al.* (1987).

Media. Bacterial strains were kept on tryptose blood agar base (TBAB) plates. For preparation of plasmid DNA the bacteria were grown in Luria broth. Minimal salts solution was that of Anagnostopoulos & Spizizen (1961). Glucose or glycerol was added to a concentration of 5 g l⁻¹ (30 mM and 50 mM, respectively). G3P was added to a concentration of 40 mM. Casamino acids were added to a concentration of 5 g l⁻¹. Antibiotics were added to the following concentrations: ampicillin (Ap) 50 mg l⁻¹, chloramphenicol (Cm) 5 mg l⁻¹ (*B. subtilis*) or 12.5 mg l⁻¹ (*Escherichia coli*), erythromycin (Em) 5 mg l⁻¹ (*B. subtilis*) or 150 mg l⁻¹ (*E. coli*) and phleomycin (Pm) 0.15 mg l⁻¹.

Transformation. Competent *B. subtilis* cells were prepared as described by Arwert & Venema (1973). *E. coli* cells were made competent as described by Mandel & Higa (1970).

DNA techniques. Chromosomal DNA was prepared by standard techniques. Plasmid DNA was prepared by the alkaline lysis method of Ish-Horowitz & Burke (1981). When preparing plasmid DNA from *B. subtilis*, the cells were treated with lysozyme (5 g l⁻¹) prior to lysis with NaOH/SDS. Plasmid DNA from *E. coli* XL1-Blue was prepared by the boiling method of Ausubel *et al.* (1987). DNA fragments separated on agarose gels were purified using GeneClean (Bio101).

DNA sequence analysis. Nucleotide sequences were determined by the dideoxy chain-termination method of Sanger *et al.* (1977), using Sequenase (USB) and [α -³⁵S]dATP. Templates used were DNA fragments cloned into pHP13 or pBluescript II KS(-). Primers used were universal or specific to the *B. subtilis* chromosomal DNA being sequenced. Analysis of nucleotide sequences was done using the GCG sequence software package (Devereux *et al.*, 1984).

Growth of bacteria, preparation of cell free extracts and enzyme assays. Cell cultures were grown in minimal salts medium supplemented with Casamino acids and required amino acids. When appropriate glucose, Cm or Em was added. The cultures were grown to early exponential phase (OD₆₀₀ = 0.5) and then divided into two parts. To one part G3P was added and to the other part no addition was made. The cultures were then incubated at 37 °C for 2 h. Cell free extracts were prepared as described by Lindgren & Rutberg (1974). G3P dehydrogenase activity was measured as described by Lin *et al.* (1962). Catechol-2,3-dioxygenase (C23O) activity was measured as described by Sala-Trepat & Evans (1971). The extinction coefficient of 2-hydroxymuconic semialdehyde at pH 8.0 (33.9 mM⁻¹ cm⁻¹) was

estimated from the data reported by Bayly *et al.* (1966). The amount of protein was determined by the Lowry method.

RNA preparation. Cell cultures were grown at 37 °C on a rotary shaker in minimal salts solution supplemented with Casamino acids. When appropriate Cm was added. The cultures were grown to an OD₆₀₀ of 0.3 and then divided into two parts. To one part glycerol was added and to the other part no addition was made. The cultures were then incubated at 37 °C for 1 h. Total RNA was extracted as described by Resnekov *et al.* (1990).

Primer extension. Primer extension was performed as described by Ayer & Dynan (1988). For each reaction 40 µg total RNA, 0.2 pmol ³²P-end-labelled primer and 1 U avian myeloblastosis virus reverse transcriptase (Pharmacia) were used. RNA and primer were annealed by incubation at 80 °C for 5 min followed by incubation on ice for 5 min. The primer was end-labelled as described by Sambrook *et al.* (1989) using T4 polynucleotide kinase (Boehringer Mannheim) and [γ -³²P]ATP [> 5000 Ci mmol⁻¹, Amersham (1 Ci = 37 GBq)]. Primer extension products were separated on polyacrylamide (6%, w/v, acrylamide, 0.3%, w/v, bisacrylamide) gels containing 7 M-urea.

Northern analysis. Northern blot and hybridization analysis was performed according to the protocols of Amersham. RNA was blotted onto Hybond-N filters (Amersham). ³²P-labelled probes were synthesized using a random-primed DNA labelling kit (Boehringer Mannheim) and [α -³²P]dCTP (> 3000 Ci mmol⁻¹, Amersham). RNA molecular mass markers (BRL) were used to determine the size of transcripts.

Results and Discussion

Nucleotide sequences of *glpP* and *glpF*

A schematic representation of the *glpP*, *glpF*, *glpK* and *glpD* region of the *B. subtilis* chromosome is shown in Fig. 1. The nucleotide sequences of *glpP*, *glpF* and adjacent regions are shown in Fig. 2.

An open reading frame, designated ORF1, ends at position 300 which is 27 bp upstream of the proposed start codon of *glpP* (see below). Sequencing of most of both strands of ORF1 (which is fully contained within the cloned *EcoRI*-*PstI* fragment) shows that it can encode a protein of 50 kDa. The predicted amino acid sequence of the ORF1 protein shows 33% identity with that of the *B. sphaericus* BioA protein and has several conserved motifs characteristic for other aminotransferases (Fig. 3). It should be pointed out, however, that the *B. subtilis* *bioA* gene has been mapped at 268° (Piggot *et al.*, 1990) whereas ORF1 is located at 75° on the *B. subtilis* chromosomal map. To test if ORF1 encodes a protein essential for glycerol catabolism, a gene disruption experiment was performed. An ORF1-internal 470 bp *SphI* fragment was cloned in the *B. subtilis* integration plasmid pHV32 (Niaudet *et al.*, 1982) to give plasmid pLUM316 (Fig. 1). pLUM316 can only replicate in *E. coli* but carries a *cat* gene which is expressed in both *E. coli* and *B. subtilis*. *B. subtilis* BR95 was transformed with pLUM316 and transformants resistant to Cm were selected. These transformants have

Table 1. Bacterial strains and plasmids

Strain/plasmid	Genotype/phenotype	Source/reference
Strain		
<i>B. subtilis</i>		
BR95	<i>ilvC1 pheA1 trpC2</i>	Our collection
LUG2506	<i>ilvC1 trpC2 glpP6</i>	Our collection
LUG2509	<i>ilvC2 trpC2 glpP9</i>	Our collection
LUG2512	<i>ilvC1 trpC2 glpP12</i>	Our collection
LUG0401	<i>ilvC1 pheA1 trpC2</i> $\Delta glpPFKD$ Pm ^R	Holmberg <i>et al.</i> (1990)
LUG0402	<i>ilvC1 pheA1 trpC2</i> $\Delta glpP$ Pm ^R	This work
<i>E. coli</i>		
JM83	<i>ara</i> $\Delta(lac-proAB)$ <i>rpsL</i> $\phi 80$ <i>lacZ</i> $\Delta M15$	Yanisch-Perron <i>et al.</i> (1985)
JM109	<i>endA1 recA1 hsdR17 supE44 gyrA96</i> <i>relA1 thi</i> $\Delta(lac-proAB)$ F' [<i>traD36</i> <i>proAB</i> ⁺ <i>lacF</i> ⁺ <i>lacZ</i> $\Delta M15$]	Yanisch-Perron <i>et al.</i> (1985)
XL1-Blue	<i>recA1 endA1 gyrA96 hsdR17</i> <i>supE44 relA1 thi lac F</i> ' [<i>proAB</i> ⁺ <i>lacF</i> ⁺ <i>lacZ</i> $\Delta M15$ Tn10(Tc ^R)]	Stratagene
Plasmid		
pHP13	Cm ^R Em ^R	Haima <i>et al.</i> (1987)
pUC18	Ap ^R	Yanisch-Perron <i>et al.</i> (1985)
pHV32	Ap ^R Cm ^R Tc ^R	Niaudet <i>et al.</i> (1982)
pAMB22	Cm ^R Tc ^R XylE ⁻	Zukowski & Miller (1986)
pUB110	Km ^R Pm ^R	Gryczan <i>et al.</i> (1978)
pBluescript II KS(-)	Ap ^R	Stratagene
pLUM30	Cm ^R Em ^R	Holmberg <i>et al.</i> (1990)
pLUM304	Cm ^R Em ^R	This work
pLUM309	Cm ^R	Holmberg <i>et al.</i> (1990)
pLUM316	Ap ^R Cm ^R Tc ^R	This work
pLUM7	Em ^R	This work
pLUM52	Cm ^R Tc ^R XylE ⁺	Holmberg & Rutberg (1991)
pLUM610	Cm ^R Tc ^R XylE ⁺	This work

Fig. 1. Schematic representation of the *glpP*, *glpF*, *glpK* and *glpD* region of the *B. subtilis* chromosome. Inverted repeats are indicated with hairpin symbols. Arrows indicate direction of transcription. The chromosomal DNA fragment present in each pLUM plasmid is shown. pLUM304, pLUM316, pLUM309 and pLUM610 are derivatives of pLUM30. C, *ClaI*; D, *DnaI*; E, *EcoRI*; EV, *EcoRV*; H, *HindIII*; N, *NcoI*; P, *PstI*; S, *SphI*. Restriction endonuclease sites shown are not necessarily unique.

no additional nutritional requirements and they grow with glycerol as sole carbon and energy source. To confirm that pLUM316 had integrated close to the *glp*

region, chromosomal DNA from one of the transformants was used to transform a *glpP* mutant, LUG2506, to Cm^R. Of 180 Cm^R transformants tested, 177 were also Glp⁺ confirming that pLUM316 had integrated close to the *glp* genes at 75°. This was also verified by Southern blot analysis (data not shown).

At 27 bp downstream of the stop codon of ORF1 a second open reading frame (ORF2) starts with an ATG at position 328 and ends with a TGA stop codon at position 906 (Fig. 2). ORF2 can encode a protein of 192 amino acids with a molecular mass of 22 kDa. ORF2 is preceded by a typical ribosome binding site (rbs; AAAGGAG, $\Delta G = -56$ kJ). Upstream of the proposed ATG start codon there is a canonical σA -10 sequence separated by 16 bp from a possible -35 sequence (Fig. 2). Overlapping this possible promoter region from positions 266 to 279 is a sequence where 13/14 bp conform to the 'glucose repression sequence', -T-G-A/T-N-A-N-C-G-N-T-N-A/T-C-A-, proposed by Weickert & Chambliss (1990). ORF2 is contained within a 4 kb *EcoRI*-*ClaI* fragment, (Fig. 1). This fragment, which contains wild type alleles of several *glpP* muta-

Fig. 2. For legend see facing page.

tions, was subcloned from pLUM30 in plasmid pHP13 to give plasmid pLUM304. The *glpP* gene product is known to act *in trans* (Holmberg & Rutberg, 1991). To

test for *glpP* complementing activity by pLUM304 the plasmid was introduced into strains LUG0401/pLUM52 and LUG0402. In LUG0401 the whole *glp* region from

1021 GTGGTTTTTTTATTCTCTTCTTCTCTATCATGCTTTTAAATCGTGACTTTTAGGAGGAATG
rbs

1081 TGCTATGACAGCATTGTTGGGGAGAAGTCATCGGTACGATGCTGCTTATCATTTTGGTGC
M T A F W G E V I G T M L L I I F G A 19
glpF

1141 AGGTGTTTGTGTCAGGTGTCAATTTAAAGAAATCGCTTTCATTCCAGTCTGGCTGGATTGT
G V C A G V N L K K S L S F Q S G W I V 39
DraI

1201 TGTGTATTGTGATGGGGATTGGGTGTTGCCATGGCGGCATACGCGGTGGCGGCATCAG
V V F G W G L G V A M A A Y A V G G I S 59
NcoI

1261 CGGTGCCCATTGTAATCCGGCGCTAACGATAGCGCTTGCATTGTAGGAGATTTCCTTG
G A H L N P A L T I A L A F V G D F P W 79

1321 GAAAGAGGTTCCGGTTTATATTGCGGCGCAAATGATCGGAGCGATCATCGGTGCGGTGAT
K E V P V Y I A A Q M I G A I I G A V I 99

1381 TATTTATCTGCATTACCTCCCGCACTGGAAGTCAACGGATGATCCCGCTGCCAAGCTGGG
I Y L H Y L P H W K S T D D P A A K L G 119

1441 TGTTTTCTCAACAGGGCCCAGCATTCCGCATACATTGCAAAACGTTTAAAGCGAAGTGAT
V F S T G P S I P H T F A N V L S E V I 139

1501 TGGGACATTGTCTTGTACTTGAATCTTGGCCATAGGTGCAAATCAATTTACAGAAGG
G T F V L V L G I L A I G A N Q F T E G 159

1561 ACTTAATCCTTTAATCGTCGGTTTCTCATTTGTAGCAATCGGTATTTCTCTGGGAGGCAC
L N P L I V G F L I V A I G I S L G G T 179

1621 CACCGGCTATGCTATCAATCTGCACGTGACTTAGGTCCGCGGATCGCCCACGCTTTTCT
T G Y A I N P A R D L G P R I A H A F L 199

1681 TCCGATTCCGGGGAAGGGCTCATCAAATGGAATACGCATGGGTTCGGTAGTCGGCCC
P I P G K G S S N W K Y A W V P V V G P 219

1741 GATTTTAGGCGGATCATTCGGCGGTGTATTTTACAACGCTGCATTTAAAGGGCACATCAC
I L G G S F G G V F Y N A A F K G H I T 239

1801 AAGCAGCTTCTGGATTGTAAGCGTTATATTGTTGTGGTATTGTTAGGACTCTATGTTTA
S S F W I V S V I L V V V L L G L Y V Y 259

1861 TACGAAATCATTCTGCTAAAACATTATCAAATCTAAATATATTTAATCAAGGGGAG
T K S H S A K T L S N S K Y I * rbs 274

1921 ACATCTTATGGAAACGTACATTTTATCCTTAGATCAGGGGACGACAAGTTCAAGAGCGAT
M E T Y I L S L D Q G T T S S R A 17
glpK

Fig. 2. Nucleotide sequence of part of ORF1, *glpP*, *glpF* and the beginning of *glpK* of *B. subtilis*. Both strands were sequenced and all restriction sites fully overlapped. Potential rbs and promoter sequences are underlined. Nucleotides conforming to the 'glucose repression sequence' are marked with $\wedge$. The C to T transition occurring in pLUM304 is indicated. ∇ shows the deletion startpoint in the two deleted plasmid derivatives of pLUM309. The inverted repeat preceding *glpFK* is indicated with horizontal arrows. The transcriptional startpoint of *glpFK* is indicated with a vertical arrow. The 17-mer used in primer extension experiments is indicated. Left-margin numbers refer to the nucleotide sequence, right-margin to amino acid sequences. Nucleotide position 1230 in the sequence published here corresponds to position 1 in the nucleotide sequence published by Holmberg *et al.* (1990).

ORF3	1	MTAFWGEVIGTMLLIIFGAGVCAGVNLKKSLSFQSGWIVVVFPGWGLGVAMAAYAVGGISGAHLNPALTIALAFVGDFFWKEVPVYIAQMIGA
GlpF	1	MSQTSTLKGQCIAEFLGTGLLIFFGVGCVAALKVAGASFGQWEISVI...WGLGVAMAIYLTAGVSGAHLNPAVTIALWLFACFDKRRVIPPVIVSQVAGA
		-----E--GT--LLI--FG--G--A-----Q-----V-----WGLGVAMA--Y---G--SGAHLNPA--TIAL-----F-----V---I--Q--GA
ORF3	94	IIGAVIYYLHYLPHW.....KSTDDPAAKLGVFSTGSPHPTFANVLSEVIGTFVLVLGI...LAIGANQFTEG.LNPLIVGFLIVAIGISLGGT
GlpF	98	FCAALVYGLYNNLFFDFEQTHHIVRGSVESVDLAGTFSTYPNPHINFVQAFVEMVITAILMGLILALTVDGNGVPRGPLAPLLIGLLIAVIGASMGPL
		---A---Y--Y-----G--FST--P-----F-----G---L---N---G--L--PL--G--LI--IG--S--G--
ORF3	180	TGYAINPARDLGPRIAHAFLP...IPGKGSSNWKYAWVPVVGPIILGSGFQGVFYNAAFKGHITSSFWIVSVILVVVLLGLYVYTKSHSAKTLNSKYI
GlpF	198	TGFAMNPARDFGPKV.FAWLAGWGNVAFGTGGRDIPYFLVPLFGPIVGAIVGAFAYRKLIGRHLPCD.....ICVVEEKETPTTPEQKASL.....
		TG--A--NPARD--GP---A--L-----G-----Y--VP--GPI--G---G---Y-----H-----V-----T--S-----L-----

Fig. 4. Comparison of ORF3 and *E. coli* GlpF amino acid sequences. Identical amino acid residues are shown in the line below the two compared sequences. Residues that are conserved throughout the MIP family are underlined in the ORF3 sequence. Dots represent gaps inserted to optimize the protein alignment, using the default parameter of the GAP program (Devereux *et al.*, 1984).

glpK is located 18 bp downstream of this stop codon. ORF3 can encode a protein of 274 amino acids and with a molecular mass of 29 kDa. A hydropathy plot (Kyte & Doolittle, 1982) of the predicted amino acid sequence of the ORF3 protein reveals several hydrophobic stretches and the protein has the potential to form 6–7 membrane spanning segments. The ORF3 protein shows 35% identity with the *E. coli* GlpF protein (Fig. 4). The latter protein is a membrane protein involved in facilitated diffusion of glycerol across the cell membrane. The *E. coli* GlpF protein belongs to the so called MIP family of integral membrane proteins which includes transport proteins from bacteria, plants, insects and animals (Pao *et al.*, 1991; Bairoch, 1991). Amino acid residues that are conserved throughout the MIP family are also conserved in the ORF3 protein (Fig. 4). In *E. coli*, *glpF* is located immediately upstream of *glpK* and the two genes form one operon (Sweet *et al.*, 1990). This is also true for *B. subtilis* ORF3 and *glpK* (see below). We are confident that ORF3 encodes the GlpF protein of *B. subtilis*. The *glpF* gene is preceded by a typical rbs (AGGAGG, $\Delta G = -71$ kJ). A σ A-type promoter TTGACA–17 bp–TACAAT is found upstream of the start of *glpF* (Fig. 2). TGA of the proposed –35 sequence for *glpF* is the stop codon of *glpP*. Overlapping the promoter region from position 904 to 917 (Fig. 2) is a sequence with a perfect match (14/14) to the ‘glucose repression sequence’ (Weickert & Chambliss, 1990). An inverted repeat is found between the proposed promoter and the rbs of *glpF* (Fig. 2). Interestingly, this repeat and part of the upstream sequence show a high degree of base pair conservation when compared to the inverted repeat located between the *glpD* promoter and *glpD* (Fig. 5). Most likely the inverted repeat is an important control element for the expression of *glpF* and *glpK* as has been demonstrated previously for *glpD* (Holmberg & Rutberg, 1991, 1992).

Transcription of *glpF* and *glpP*

As mentioned above a σ A-type *glpF* promoter is suggested from the nucleotide sequence. To test for promoter activity the following experiment was done. A 606 bp *EcoRV*–*DraI* fragment containing the proposed *glpF* promoter (positions 558 to 1164, Fig. 2) was cloned by blunt end ligation in the *SmaI* site of pUC18 in *E. coli* JM109. The identity and orientation of the insert was verified by DNA sequencing. Plasmid DNA was then cleaved with *EcoRI* and *HincII*, sites for which flank the insert. The insert was then ligated to *EcoRI*/*SmaI* cleaved pAMB22 DNA which places it upstream of the *xylE* reporter gene. The ligation mix was used to transform BR95 to Cm^R and the Cm^R transformants were tested for C23O activity. Plasmid DNA was extracted from one C23O-positive transformant and the identity of the plasmid, pLUM610 (Fig. 1), was verified by restriction enzyme analysis and by DNA sequencing of the insert. C23O activity was then measured in BR95/pLUM610 grown in the presence or absence of G3P and glucose. The results of these experiments (Table 3) show that the *EcoRV*–*DraI* fragment has promoter activity in *B. subtilis*. The C23O activity increased two- to threefold in cells grown with G3P and decreased when glucose was also present.

To determine the start point of the transcript initiated from the promoter in the *EcoRV*–*DraI* fragment the following experiment was done. BR95 and BR95/pLUM610 were grown with and without glycerol, total RNA was extracted and used in primer extension experiments. The primer used was a 17-mer corresponding to positions 973 to 989 (Fig. 2). The results of these experiments (Fig. 6) place the start of the transcripts at an A residue at position 938 (Fig. 2). This residue is located 10 nucleotides downstream of the centre of the proposed σ A –10 sequence. The conserved inverted

Fig. 5. The *glpFK* and *glpD* inverted repeats. The inverted repeats are presented as stem-loop structures. Highly conserved regions are boxed. The -10 regions and transcriptional startpoints are indicated.

Table 3. Activity of C23O in *B. subtilis* BR95 carrying different plasmids

Plasmid	Addition(s)	C23O activity*
pAMB22	-	< 1
	G3P	< 1
pLUM610	-	6
	G3P	15
	Glucose	3
	G3P + glucose	8

* C23O activity is expressed as nmol product (2-hydroxymuconic semialdehyde) formed min⁻¹ (mg protein)⁻¹.

repeat is located downstream of the transcriptional startpoint in the UTL of the *glpF* transcript. Probably the inverted repeat serves as a conditional stop signal for transcription as has been suggested previously for the inverted repeat in the UTL of *glpD* mRNA (Holmberg & Rutberg, 1991, 1992). The stop codon of *glpF* and the

Fig. 6. Determination of the start of the *glpFK* transcript by primer extension. The standard DNA sequencing reactions (A, C, T, G) were done with the same primer used in the primer extension reactions. Lane 1, 9 µg BR95 RNA; lane 2, 9 µg BR95/pLUM610 RNA.

start codon of *glpK* are separated by 18 bp suggesting that these two genes may constitute an operon. To test this suggestion a Northern blot analysis was done. BR95 was grown with and without glycerol, total RNA was extracted and probed with an 1100 bp *NcoI*-*ClaI* fragment (Fig. 1). RNA from LUG0401 grown in the presence of glycerol was used as a control. In uninduced BR95 a transcript of about 2400 nucleotides (nt) was detected which corresponds in size to that expected for a *glpFK* mRNA (Fig. 7). The relative amount of this transcript increased about fivefold when BR95 was grown with glycerol. A transcript of about 4400 nt was also found in glycerol-grown cells, which is thought to represent a *glpFKD* transcript. A transcript of similar size is also found in induced cells when using a probe specific for *glpD* (Holmberg & Rutberg, 1992). As expected, no transcript hybridizing with the probe was

Fig. 7. Northern analysis of *glpFK*-specific transcripts. Lanes 1, 2 and 3; 5, 10 and 20 μ g, respectively, of total RNA from uninduced BR95. Lanes 4, 5 and 6; 5, 10 and 20 μ g, respectively, of total RNA from induced BR95. Lane 7; 10 μ g total RNA from induced LUG0401. The probe used was a 32 P-labelled internal *glpFK* DNA fragment. The positions of the RNA size markers visualized by staining with methylene blue are indicated on the left.

found in LUG0401. From the above results we conclude that *glpF* and *glpK* constitute one operon.

The *xylE* reporter gene in pAMB22 was used to search for promoters upstream of *glpP*. However, no significant promoter activity was found within 240 bp upstream of the proposed start codon for *glpP*. Also no transcript initiated in the intergenic region between ORF1 and *glpP* has been observed in primer extension experiments. Northern blots were then done to search for a *glpP* transcript. The probes used were a 275 bp *HindIII*–*EcoRV* fragment which extends from 45 bp upstream of the start of *glpP* and into this gene, and a *glpP*-internal 210 bp *EcoRV*–*HinfI* fragment (Fig. 2). Using either probe a weak signal (compared to, e.g. uninduced *glpFK* transcript) corresponding to a transcript(s) of 1800 to 2100 nt was found (Fig. 8). This corresponds in size to a transcript initiated upstream of ORF1 and extending to the start of *glpF*. The fact that disrupting ORF1 by plasmid integration does not lead to a Glp-negative phenotype could be explained by *glpP* being transcribed from a new promoter located in the plasmid inserted into

ORF1. It seems clear from the above results, however, that *glpP* is not cotranscribed with *glpFK*.

Conclusions

The present results together with previously published experiments have identified all *B. subtilis glp* genes located at 75° on the chromosomal map. These four genes represent three separate transcription units, namely, *glpP*, *glpFK* and *glpD*. The activities of *glpFK* and *glpD* are subject to several negative and positive controls involving, e.g. *GlpP*, the phosphoenolpyruvate:sugar phosphotransferase system (PTS; Reizer *et al.*, 1988) and glucose repression. Further work is now directed towards understanding the molecular nature of these controls.

This work was supported by grants from the Swedish Medical Research Council, Emil and Wera Cornells Stiftelse and Kungliga Fysiografiska Sällskapet i Lund. We thank Ingrid Stål for expert technical assistance.

Fig. 8. Northern analysis of *glpP*-specific transcripts. Lanes as described in the legend to Fig. 7. The ^{32}P -labelled probes used were synthesized from (a) the *Hind*III–*Eco*RV DNA fragment, and (b) the *Eco*RV–*Hinf*I DNA fragment, which is internal for *glpP*. The positions of the RNA size markers, visualized by staining with methylene blue, are indicated on the left.

References

- ANAGNOSTOPOULOS, C. & SPIZIZEN, J. (1961). Requirements for transformation in *Bacillus subtilis*. *Journal of Bacteriology* **81**, 741–746.
- ARWERT, F. & VENEMA, G. (1973). Transformation in *Bacillus subtilis*. *Molecular and General Genetics* **123**, 185–198.
- AUSUBEL, F. M., BRENT, R., KINGSTON, R. E., MOORE, D. D., SEIDMAN, J. G., SMITH, J. A. & STRUHL, K. (1987). Minipreps of plasmid DNA. In *Current Protocols in Molecular Biology* 1.6.4–1.6.5. New York: John Wiley & Sons.
- AYER, D. E. & DYNAN, W. S. (1988). Simian virus 40 major late promoter: a novel tripartite structure that includes intragenic sequences. *Molecular and Cellular Biology* **8**, 2021–2033.
- BAIROCH, A. (1991). PROSITE: a dictionary of sites and patterns in proteins. *Nucleic Acids Research* **19**, 2241–2245.
- BARTSCH, K., VON JOHN-MARTEVILLE, A. & SCHULZ, A. (1990). Molecular analysis of two genes of the *Escherichia coli* *gab* cluster: nucleotide sequence of the glutamate:succinic semialdehyde transaminase gene (*gabT*) and characterization of the succinic semialdehyde dehydrogenase gene (*gabD*). *Journal of Bacteriology* **172**, 7035–7042.
- BAYLY, R. C., DAGLEY, S. & GIBSON, D. T. (1966). The metabolism of cresols by species of *Pseudomonas*. *Biochemical Journal* **101**, 293–301.
- DEGOLS, G. (1987). Functional analysis of the regulatory region adjacent to the *cargB* gene of *Saccharomyces cerevisiae*. Nucleotide sequence, gene fusion experiments and cis-dominant regulatory mutation analysis. *European Journal of Biochemistry* **169**, 193–200.
- DEVEREUX, J., HEBERLI, P. & SMITHIES, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Research* **12**, 387–395.
- FRIDÉN, H., HEDERSTEDT, L. & RUTBERG, L. (1987). Deletion of the *Bacillus subtilis* *sdh* operon. *FEMS Microbiology Letters* **41**, 203–206.
- GLOECKLER, R., OHSAWA, I., SPECK, D., LEDOUX, C., BERNARD S., ZINSIUS, M., VILLEVAL, D., KISOU, T., KAMOGAWA, K. & LEMOINE, Y. (1990). Cloning and characterization of the *Bacillus sphaericus* genes controlling the bioconversion of pimelate into dethiobiotin. *Gene* **87**, 63–70.
- GRYCZAN, T. J., CONTENTE, S. & DUBNAU, D. (1978). Characterization of *Staphylococcus aureus* plasmids introduced by transformation into *Bacillus subtilis*. *Journal of Bacteriology* **134**, 318–329.
- HAIMA, P., BRON, S. & VENEMA, G. (1987). The effect of restriction on shotgun cloning in *Bacillus subtilis* Marburg. *Molecular and General Genetics* **209**, 335–342.
- HOLMBERG, C. & RUTBERG, B. (1991). Expression of the gene encoding glycerol-3-phosphate dehydrogenase (*glpD*) in *Bacillus subtilis* is controlled by antitermination. *Molecular Microbiology* **5**, 2891–2900.
- HOLMBERG, C. & RUTBERG, L. (1992). An inverted repeat preceding the *Bacillus subtilis* *glpD* gene is a conditional terminator of transcription. *Molecular Microbiology* **6**, 2931–2938.
- HOLMBERG, C., BEIJER, L., RUTBERG, B. & RUTBERG, L. (1990). Glycerol catabolism in *Bacillus subtilis*: nucleotide sequence of the genes encoding glycerol kinase (*glpK*) and glycerol-3-phosphate dehydrogenase (*glpD*). *Journal of General Microbiology* **136**, 2367–2375.
- ISH-HOROWICZ, D. & BURKE, J. F. (1981). Rapid and efficient cosmid cloning. *Nucleic Acids Research* **9**, 2989–2998.
- KYTE, J. & DOOLITTLE, R. F. (1982). A simple method for displaying the hydrophobic character of a protein. *Journal of Molecular Biology* **157**, 105–132.
- LIN, E. C. C., KOCH, J. P., CHUSED, T. M. & JORGENSEN, S. E. (1962). Utilization of L- α -glycerophosphate by *Escherichia coli* without hydrolysis. *Proceedings of the National Academy of Sciences of the United States of America* **48**, 2145–2150.
- LINDGREN, V. (1978). Mapping of a genetic locus that affects glycerol-3-phosphate transport in *Bacillus subtilis*. *Journal of Bacteriology* **133**, 667–670.
- LINDGREN, V. & RUTBERG, L. (1974). Glycerol metabolism in *Bacillus*

- subtilis*: gene-enzyme relationships. *Journal of Bacteriology* **119**, 431-442.
- LINDGREN, V. & RUTBERG, L. (1976). Genetic control of the *glp* system in *Bacillus subtilis*. *Journal of Bacteriology* **127**, 1047-1057.
- MANDEL, M. & HIGA, A. (1970). Calcium-dependent bacteriophage DNA infection. *Journal of Molecular Biology* **53**, 159-162.
- NIAUDET, B., GOZE, A. & ERLICH, S. D. (1982). Insertional mutagenesis in *Bacillus subtilis*: mechanism and use in gene cloning. *Gene* **19**, 277-284.
- OTSUKA, A. J., BUONCRISTIANI, M. R., HOWARD, P. K., FLAMM, J., JOHNSON, C., YAMAMOTO, R., UCHIDA, K., COOK, C., RUPPERT, J. & MATSUZAKI, J. (1988). The *Escherichia coli* biotin biosynthetic enzyme sequences predicted from the nucleotide sequence of the *bio* operon. *Journal of Biological Chemistry* **263**, 19577-19585.
- PAO, G. M., WU, L.-F., JOHNSON, K. D., HÖFTE, H., CHRISPEELS, M. J., SWEET, G., SANDAL, N. N. & SAIER, M. H., JR. (1991). Evolution of the MIP family of integral membrane transport proteins. *Molecular Microbiology* **5**, 33-37.
- PIGGOT, P. J., AMJAD, M., WU, J.-J., SANDOVAL, H. & CASTRO, J. (1990). Genetic and physical maps of *Bacillus subtilis* 168. In *Molecular Biological Methods for Bacillus*, pp. 493-543. Edited by C. R. Harwood & S. M. Cutting. Chichester: John Wiley & Sons.
- REIZER, J., SAIER, M. H., DEUTSCHER, J., GRENIER, F., THOMPSON, J. & HENGSTENBERG, W. (1988). The phosphoenolpyruvate:sugar phosphotransferase system in Gram-positive bacteria: properties, mechanism and regulation. *CRC Critical Reviews in Microbiology* **15**, 297-338.
- RESNEKOV, O., RUTBERG, L. & VON GABAIN, A. (1990). Changes in the stability of specific mRNA species in response to growth stage in *Bacillus subtilis*. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 8355-8359.
- SALA-TREPAT, J. M. & EVANS, W. C. (1971). The meta cleavage of catechol by *Azotobacter* species. *European Journal of Biochemistry* **20**, 400-413.
- SAMBROOK, J., FRITSCH, E. F. & MANIATIS, T. (1989). Preparation of radiolabeled DNA and RNA probes. In *Molecular Cloning: A Laboratory Manual* 2nd edn, 10.60-10.61. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5463-5467.
- SWEET, G., GANDOR, C., VOEGELE, R., WITTEKINDT, N., BEUERLE, J., TRUNIGER, V., LIN, E. C. C. & BOOS, W. (1990). Glycerol facilitator of *Escherichia coli*: cloning of *glpF* and identification of the *glpF* product. *Journal of Bacteriology* **172**, 424-430.
- WEICKERT, M. J. & CHAMBLISS, G. H. (1990). Site-directed mutagenesis of a catabolite repression operator sequence in *Bacillus subtilis*. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 6238-6242.
- YANISCH-PERRON, C., VIEIRA, J. & MESSING, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of M13mp18 and pUC19 vectors. *Gene* **33**, 103-119.
- ZUKOWSKI, M. M. & MILLER, L. (1986). Hyperproduction of an intracellular heterologous protein in a *sacU*^h mutant of *Bacillus subtilis*. *Gene* **46**, 247-255.