Microbiology (1994), 140, 1829–1838 Printed in Great Britain

Analysis of a ribose transport operon from *Bacillus subtilis*

Karen Woodson[1,2]† and Kevin M. Devine[2]

Author for correspondence: Kevin M. Devine. Tel: +353 1 7021872. Fax: +353 1 6798558.
e-mail: KDEVINE@VAX1.TCD.IE

National Pharmaceutical Biotechnology Centre,[1] and Department of Genetics,[2] Trinity College, Dublin 2, Ireland

The *csa-15* locus of *Bacillus subtilis* corresponds to an operon encoding proteins which display features characteristic of the ABC group of transporters. Sequence analysis reveals a very high level of identity to the ribose transport operon of *Escherichia coli*. This hypothesis is supported by the observation that strains carrying mutagenic insertions in this operon are unable to grow on ribose as sole carbon source. Expression of this operon is directed by a single SigA-type promoter which is negatively regulated by Spo0A during the late-exponential/transition state of the growth cycle. Expression is also subject to catabolite repression and this mode of regulation is dominant to control of expression by Spo0A.

Keywords: *Bacillus subtilis*, ribose transport operon, Spo0A, catabolite repression

INTRODUCTION

Cell walls and membranes are effective barriers against the influx and efflux of solutes and metabolites. Bacterial transporters have been classified into osmotic shock resistant and osmotic shock sensitive systems (Ames, 1986; Higgins *et al.*, 1990). This classification is based on the observation that a large number of transport systems have a substrate-binding protein located in the periplasm of the cell. This is released upon osmotic shock, and renders the cell incapable of transporting the metabolite. In addition to the periplasmic substrate-binding protein, this group of transporters is characterized by four protein cassettes: two hydrophobic membrane-located domains and two ATP-binding cassettes. This group is called the ABC (ATP-binding cassette) group of transporters (Higgins *et al.*, 1990). In some systems, the four cassettes are located on separate proteins as in the case of the oligopeptide transport system of *Salmonella typhimurium*. In other systems, some of the domains are fused into a single protein. The two ATP-binding domains are fused into a single protein in the ribose transport system of *Escherichia coli* while the two hydrophobic domains are fused into a single protein in the p69 system of *Mycoplasma*. Although first identified in bacteria, transport systems with these characteristics are also present in eukaryotes.

†**Present address:** Department of Laboratory Medicine, Washington University Medical School, 660 S. Euclid, St Louis, MO 63110, USA.

The GenBank accession number for the sequence reported in this paper is Z25798.

ABC-type transport systems for oligopeptides (Perego *et al.*, 1991; Rudner *et al.*, 1991) and dipeptides (Mathiopoulos *et al.*, 1991) have been identified in *Bacillus subtilis*. In addition to the proteins containing the four cassettes, these systems have a protein homologous to the periplasmic substrate-binding protein, even though this bacterium does not have a periplasm. Nevertheless, Perego *et al.* (1991) have demonstrated that the periplasmic oligopeptide binding protein OppA is cell wall associated in exponentially growing cells. Thus this protein is probably functionally equivalent to its Gram-negative counterpart in that it is located outside, but anchored to, the cell membrane and binds substrate with high affinity.

O'Reilly *et al.* (1994) describe a method which can be used to identify operons whose expression is regulated by any particular regulator protein. Using this strategy, 28 strains of *B. subtilis* were identified (designated CSA for *c*ontrol *b*y *S*po0A), each of which harbours an operon–*lacZ* fusion which is negatively regulated by Spo0A. Analysis revealed that in strain CSA8 the citrulline biosynthetic operon *argC–F* is fused to *lacZ* and that expression of this operon is negatively regulated by Spo0A during cell growth on solid medium, but not during growth in liquid culture (O'Reilly *et al.*, 1994). In this communication, we report that the operon to which the *lacZ* gene is fused in strain CSA15 displays a high level of identity to the ribose transport operon identified in *E. coli*. This putative ribose transport operon from *B. subtilis* displays many of the features characteristic of the ABC group of transporters. Expression of this operon is directed by a single SigA-

0001-8769 © 1994 SGM

Table 1. Strains and plasmids

Strain or plasmid	Genotype/description	Source/reference
E. coli		
tg1	K12 Δ(*lac–pro*) *supE thi hsd*R F′ *traD36 proAB lacI lac*ZΔM15	Amersham
B. subtilis		
JH642	*trpC2 pheA1*	BGSC*
JH646	*trpC2 pheA1 spo0A12*	BGSC*
KD882	*trpC2 pheA1*::pCSA15	This study
KD883	*trpC2 pheA1 spo0A12*::pCSA15	This study
KD887	*trpC2 pheA1*::pKW3	This study
KD888	*trpC2 pheA1 spo0A12*::pKW3	This study
KD889	The *Eco*RI–*Cla*I (nt 1–945) fragment ligated into pJM783 and transformed into JH642	This study
KD890	The *Eco*RI–*Cla*I (nt 1–945) fragment ligated to pJM783 and transformed into JH646	This study
KD891	The *Eco*RI–*Eco*RV (nt 1–493) fragment ligated to pDG268 and transformed into JH642	This study
KD892	The *Eco*RI–*Eco*RV (nt 1–493) fragment ligated to pDG268 and transformed into JH646	This study
Plasmids		
pGEM-7Zf(+)	ColE1 derived cloning vector	Promega, WI
pJM783	An integrating plasmid containing Cm^r and a promoterless *lacZ*	J. A. Hoch
pDG268	An integrating plasmid containing Cm^r and a promoterless *lacZ* inserted into an α-amylase gene	Antoniewski *et al.* (1990)
pCSA15	The *Eco*RI–*Sau*3A (nt 2495–3487) fragment of the ribose transport operon directionally located 5′ to *lacZ* of pJM783	This study
pKW1	The 3020 bp *Pst*I fragment cloned into pGEM7Zf(+)	This study
pKW2	The 1133 bp *Eco*RV fragment cloned into pGEM5Zf(+)	This study
pKW3	The *Eco*RV–*Taq*I (nt 1736–1832) fragment of the ribose transport operon ligated directionally, 5′ to *lacZ* of pJM783	This study
pKW5	The *Eco*RI–*Taq*I fragment (nt 1–1832) subcloned in pGEM7Zf(+)	This study

* *Bacillus* Genetic Stock Centre, Ohio State University, Columbus, Ohio, USA.

type promoter and is subject to carbon catabolite repression which operates at the level of transcription.

METHODS

Bacterial strains and growth conditions. The strains used in this study are described in Table 1. Strain CSA15 was identified and isolated as described in O'Reilly *et al.* (1994). *B. subtilis* strains were grown on SM (Schaeffer sporulation medium; Schaeffer *et al.*, 1965) or LB (Luria Bertani medium; Miller, 1972). Solid medium was made with LB or SM medium containing 1·5% (w/v) agar (Difco). Chloramphenicol (3 μg ml^{-1}), X-Gal (5-bromo-4-chloro-3-indolyl β-D-galactopyranoside, 40 μg ml^{-1}) or IPTG (isopropyl β-thiogalactopyranoside, 1 mM) were added as appropriate. *E. coli* strains were grown in LB broth or on agar with selection for ampicillin resistance at 100 μg ml^{-1}. To test for growth of *B. subtilis* integrant strains on ribose as sole carbon source, strains were plated on minimal medium (Anagnostopoulos & Spizizen, 1961) containing the auxotrophic requirements (Trp and Phe, 40 μg ml^{-1} each) and ribose (1%, w/v).

Lambda bank of *B. subtilis* chromosomal DNA, plasmids and plasmid construction. The lambda bank of *B. subtilis* chromosomal DNA was constructed as previously described (Wood *et al.*, 1990). Chromosomal DNA from the *csa-15* locus was cloned as follows: chromosomal DNA from CSA15 was digested with *Eco*RI, ligated and transformed into *E. coli*. The resultant transformants contained a plasmid pCSA15 which contained a 992 bp fragment of chromosomal DNA (Fig. 1, nucleotides 2495–3487). pCSA15 was used to probe a lambda bank of chromosomal DNA and λ_{rbs} hybridized with this plasmid. From λ_{rbs}, a 3020 bp *Pst*I fragment was isolated (Fig. 1, nucleotides

```
EcoR1
GAATTCTCTAACATAATTAAACATTTTCTGGGATGATAGTCTTTTCTGTTTCTCCCCATTTACAGGTCTAAACGCATGACTTTGAAACAA-  90

TTTTAATAAAACTTAATATTTGTTCAAGAAATCTTCATCCATATTTGTGAAGACTTTGTCAAAAAAAGAGTGAAAACCTTAAATTTTTCA  180
                                                       -35                  -10       +1
ATTATATATACAATTTACAATTAGATTTCTTTTGATATTTTTATTGCTAACTTCGGATTGTTCATGATAATCTATCTATGTAAACGGTTA  270
                        rbsR->  <------------helix-turn-helix-------------->
                                   L  A  T  I  K  D  V  A  G  A  A  G  V  S  V  A  T  V  S  R  N  L  N
CATAAACAAGGAGGAGCTGTTTTGGCTACAATTAAAGATGTCGCCGGAGCGGCGGGCGTTTCAGTTGCGACGGTTTCCCGCAACCTGAAT  360

 D  N  G  Y  V  H  E  E  T  R  T  R  V  I  A  A  M  A  K  L  N  Y  Y  P  N  E  V  A  R  S
GATAACGGCTATGTACACGAAGAAACGCGAACGCGTGTCATAGCGGCGATGGCGAAGCTGAACTATTACCCGAATGAAGTCGCCAGATCT  450
                                               EcoRV
 L  Y  K  R  E  S  R  L  I  G  L  L  L  P  D  I  T  N  P  F  F  P  Q  L  A  R  G  A  E  D
CTATATAAAAGAGAATCCCGACTGATCGGACTTTTGCTCCCGGATATCACAAACCCTTTCTTCCCCCAGCTTGCCCGCGGTGCGGAGGAT  540

 E  L  N  R  E  G  Y  R  L  I  F  G  N  S  D  E  E  L  K  K  E  L  E  Y  L  Q  T  F  K  Q
GAATTGAACCGGGAAGGCTATCGCCTTATTTTCGGCAACAGTGACGAGGAATTGAAAAAAGAACTTGAATACCTTCAAACCTTTAAGCAA  630

 N  H  V  A  G  I  I  A  Q  R  I  T  R  I  S  R  E  Y  S  G  M  N  Y  P  V  V  F  L  D  R
AATCATGTCGCAGGCATTATTGCGCAACGAATTACCCGGATCTCGAGGGAATACAGCGGCATGAATTATCCAGTTGTTTTTCTGGACAGA  720

 T  L  E  G  A  P  S  V  S  S  D  G  Y  T  G  V  K  L  A  A  Q  A  I  I  H  G  K  S  Q  R
ACGCTTGAAGGGGCTCCTTCTGTGTCCAGTGACGGCTATACAGGAGTAAAGTTAGCCGCCCAGGCTATCATTCACGGAAAAAGCCAGCGC  810

 I  T  L  L  R  G  P  A  H  L  P  T  A  Q  D  R  F  N  G  A  L  E  I  L  K  Q  A  E  V  D
ATTACGCTCTTGAGAGGACCCGCTCACCTGCCGACCGCTCAAGACCGATTTAACGGCGCTTTGGAAATCTTAAAGCAGGCTGAAGTTGAT  900
                                                    ClaI
 F  Q  V  I  E  T  A  S  F  S  I  K  D  A  Q  S  M  A  K  E  L  F  A  S  Y  P  A  T  D  G
TTTCAGGTCATTGAGACAGCTTCATTTTCAATTAAAGATGCACAATCGATGGCGAAGGAGCTGTTTGCCTCTTATCCAGCGACAGATGGT  990

 V  I  A  S  N  D  I  Q  A  A  A  V  L  H  E  A  L  R  A  K  K  R  A  G  D  I  Q  I  I  G
GTGATCGCGAGTAATGATATTCAAGCCGCTGCCGTTTTACATGAAGCATTGCGCGCGAAAAAACGTGCCGGAGACATTCAAATTATCGGC 1080

 Y  D  D  I  P  Q  S  G  L  L  F  P  P  L  S  T  I  K  Q  P  A  Y  D  M  G  K  E  A  A  K
TATGATGACATTCCGCAAAGCGGACTTCTGTTCCCGCCACTTTCTACAATTAAACAGCCCGCATATGATATGGGGAAAGAAGCTGCCAAA 1170

 L  L  L  G  I  I  K  K  Q  P  L  A  E  T  A  I  Q  M  P  V  T  Y  I  G  R  K  T  T  R  K
CTGCTTTTGGGTATCATTAAAAAACAGCCGCTGGCAGAAACGGCAATTCAAATGCCTGTCACCTATATAGGAAGAAAGACGACAAGAAAG 1260
         rbsK  ->
 E  D  *  M  R  N  I  C  V  I  G  S  C  S  M  D  L  V  V  T  S  D  K  R  P  K  A  G  E  T
GAAGATTAAGATGCGTAATATTTGTGTGATTGGTAGCTGTTCTATGGATTTAGTGGTCACCTCGGACAAACGCCCAAAAGCCGGTGAAAC 1350

 V  L  G  T  S  F  Q  T  V  P  G  G  K  G  A  N  Q  A  V  A  A  A  R  L  G  A  Q  V  F  M
GGTTCTTGGCACGTCATTTCAGACTGTGCCGGGCGGCAAAGGAGCAAATCAGGCCGTTGCCGCAGCCAGACTTGGGGCGCAAGTGTTTAT 1440

 V  G  K  V  G  D  D  H  Y  G  T  A  I  L  N  N  L  K  A  N  G  V  R  T  D  Y  M  E  P  V
GGTCGGCAAAGTTGGAGACGATCATTATGGAACAGCTATTTTGAATAATCTCAAAGCTAACGGCGTTCGCACTGATTATATGGAACCGGT 1530

 T  H  T  E  S  G  T  A  H  I  V  L  A  E  G  D  N  S  I  V  V  V  K  G  A  N  D  D  I  T
TACACATACGGAAAGCGGAACCGCTCATATTGTGCTTGCTGAAGGCGACAACAGCATTGTCGTTGTCAAAGGCGCGAACGATGACATCAC 1620

 P  A  Y  A  L  N  A  L  E  Q  I  E  K  V  D  M  V  L  I  Q  Q  E  I  P  E  E  T  V  D  E
CCCAGCGTATGCGTTAAACGCGCTCGAACAGATTGAAAAGGTCGATATGGTGCTGATTCAGCAGGAAATCCCAGAAGAAACGGTTGATGA 1710
                         EcoRV
 V  C  K  Y  C  H  S  H  D  I  P  I  I  L  N  P  A  P  A  R  P  L  K  Q  E  T  I  D  H  A
GGTGTGCAAATACTGCCATTCTCACGATATCCCGATCATACTAAACCCAGCACCGGCCCGCCCGCTCAAGCAGGAAACAATTGATCACGC 1800
                           Taq1
 T  Y  L  T  P  N  E  H  E  A  S  I  L  F  P  E  L  T  I  S  E  A  L  A  L  Y  P  A  K  L
TACCTATCTGACGCCGAATGAACACGAGGCTTCGATTTTATTCCCTGAGCTTACCATTTCCGAAGCTCTGGCGCTTTATCCGGCAAAGCT 1890

 F  I  T  E  G  K  Q  G  V  R  Y  S  A  G  S  K  E  V  L  I  P  S  F  P  V  E  P  V  D  T
CTTTATTACGGAAGGAAAACAAGGCGTCCGCTATTCGGCTGGCAGCAAAGAGGTGCTCATCCCGTCCTTCCCTGTAGAACCCGTTGATAC 1980

 T  G  A  G  D  T  F  N  A  A  F  A  V  A  L  A  E  G  K  D  I  E  A  A  L  R  F  A  N  R
AACCGGAGCCGGTGATACGTTTAACGCAGCATTTGCCGTGGCCCTGGCTGAGGGAAAAGACATTGAAGCCGCATTGCGGTTTGCCAATCG 2070
                                                                          rbsD  ->
 A  A  S  L  S  V  C  S  F  G  A  Q  G  G  M  P  D  K  K  *              M  K  K  H
TGCCGCTTCCCTTTCTGTCTGCTCCTTCGGTGCCCAAGGCGGAATGCCCGACAAGAAATGAAGTAGAGGAGCTGCTGTCATGAAAAAACA 2160
```

Fig. 1. For legend see page 1833.

```
 G  I  L  N  S  H  L  A  K  I  L  A  D  L  G  H  T  D  K  I  V  I  A  D  A  G  L  P  V  P
CGGTATACTGAACAGCCATCTTGCCAAGATTTTAGCCGACCTTGGCCACACTGATAAAATTGTCATCGCGGATGCCGGACTGCCGGTTCC 2250

 D  G  V  L  K  I  D  L  S  L  K  P  G  L  P  A  F  Q  D  T  A  A  V  L  A  E  E  M  A  V
TGACGGCGTTTTGAAAATTGATCTTTCACTGAAGCCGGGCCTTCCGGCTTTCCAAGATACAGCGGCAGTACTGGCTGAGGAAATGGCGGT 2340

 E  K  V  I  A  A  A  E  I  K  A  S  N  Q  E  N  A  K  F  L  E  N  L  F  S  E  Q  E  I  E
CGAAAAAGTCATTGCTGCAGCTGAAATAAAAGCATCCAATCAGGAGAATGCGAAATTTCTAGAAAATCTTTTCTCTGAACAAGAGATTGA 2430
                                                                      EcoR1
 Y  L  S  H  E  E  F  K  L  L  T  K  D  A  K  A  V  I  R  T  G  E  F  T  P  Y  A  N  C  I
ATACCTTTCTCACGAGGAGTTTAAGCTGCTGACAAAAGATGCAAAGGCAGTCATAAGAACAGGAGAATTCACACCATATGCCAACTGCAT 2520
 PstI                                     rbsA ->
 L  Q  A  G  V  L  F  *                   M  Q  I  E  M  K  D  I  H  K  T  F  G  K  N  Q  V
CCTGCAGGCAGGTGTACTTTTCTAGAAAGGAAGATGAAACATGCAGATTGAAATGAAAGACATTCATAAAACATTCGGAAAAAATCAGGT 2610
                                                        ATP BINDING MOTIF
 L  S  G  V  S  F  Q  L  M  P  G  E  V  H  A  L  M  G  E  N  G  A  G  K  S  R  L  M  N  I
GCTGTCAGGCGTTTCCTTTCAGCTCATGCCTGGCGAGGTTCACGCATTAATGGGAGAAAACGGCGCCGGCAAGTCACGGCTTATGAACAT 2700

 L  T  G  L  H  K  A  D  K  G  Q  I  S  I  N  G  N  E  T  Y  F  S  N  P  K  E  A  E  Q  H
TTTGACAGGCCTGCACAAAGCAGATAAAGGTCAAATCAGCATAAACGGAAACGAAACGTATTTTTCCAATCCGAAAGAAGCGGAACAGCA 2790
                                                                                 EcoRV
 G  I  A  F  I  H  Q  E  L  N  I  W  P  E  M  T  V  L  E  N  L  F  I  G  K  E  I  S  S  K
TGGAATAGCCTTTATCCATCAGGAATTGAATATCTGGCCGGAAATGACCGTTCTTGAGAATCTATTTATCGGTAAAGAGATATCCTCCAA 2880
                                                                                  Sau3A
 L  G  V  L  Q  T  R  K  M  K  A  L  A  K  E  Q  F  D  K  L  S  V  S  L  S  L  D  Q  E  A
GCTGGGCGTTTTACAAACAAGAAAAATGAAAGCGCTAGCAAAAGAGCAATTTGACAAACTTTCCGTCTCTCTTTCTCTTGATCAAGAAGC 2970

 G  E  C  S  V  G  Q  Q  Q  M  I  E  I  A  K  A  L  M  T  N  A  E  V  I  I  M  D  E  P  T
CGGCGAATGTTCCGTCGGACAGCAGCAAATGATCGAAATTGCAAAAGCGCTTATGACAAATGCCGAGGTAATCATTATGGATGAACCGAC 3060

 A  A  L  T  E  R  E  I  S  K  L  F  E  V  I  T  A  L  K  K  N  G  V  S  I  V  Y  I  S  H
CGCAGCGTTGACTGAACGTGAAATCAGCAAGCTCTTTGAGGTCATTACAGCGTTAAAAAAGAACGGCGTCTCCATTGTCTATATTTCGCA 3150

 R  M  E  E  I  F  A  I  C  D  R  I  T  I  M  R  D  G  K  T  V  D  T  T  N  I  S  E  T  D
TCGCATGGAAGAAATTTTTGCGATTTGCGACAGAATTACCATCATGCGTGACGGAAAAACGGTAGATACAACAAACATCTCAGAAACTGA 3240

 F  D  E  V  V  K  K  M  V  G  R  E  L  T  E  R  Y  P  K  R  T  P  S  L  G  D  K  V  F  E
TTTTGATGAAGTCGTCAAAAAAATGGTCGGACGGGAGCTGACTGAACGATATCCAAAACGCACTCCTTCTCTCGGTGACAAAGTATTCGA 3330
                                                                                ATP BINDING
 V  K  N  A  S  V  K  G  S  F  E  D  V  S  F  Y  V  R  S  G  E  I  V  G  V  S  G  L  M  G
GGTGAAAAATGCTTCCGTAAAAGGGAGTTTTGAGGACGTCAGCTTTTATGTGCGTTCCGGTGAGATCGTCGGTGTTTCAGGATTAATGGG 3420
  MOTIF                                                 csa-15 insertion site
 A  G  R  T  E  M  M  R  A  L  F  G  V  D  R  L  D  T  G  E  I  W↓I  A  G  K  K  T  A  I
AGCCGGCCGGACAGAAATGATGAGAGCGCTGTTCGGCGTTGACAGGCTGGACACGGGTGAGATATGGATCGCTGGGAAAAAAACGGCTAT 3510

 K  N  P  Q  E  A  V  K  K  V  S  A  L  L  Q  R  I  A  R  M  K  G  S  C  S  T  A  S  P  E
TAAGAACCCGCAGGAAGCCGTAAAAAAGGTCTCGGCTTTATTACAGAGAATCGCAAGGATGAAGGGCTCCTGCTCGACAGCATCACCGGA 3600

 T  H  A  R  H  L  S  G  G  K  P  G  K  K  V  V  I  A  K  W  I  G  I  G  P  K  V  L  I  L
AACGCACGCACGCCATTTATCAGGAGGCAAACCAGGCAAAAAAGTGGTGATAGCCAAGTGGATCGGCATCGGACCGAAAGTGCTTATCTT 3690

 D  E  P  T  R  G  V  D  V  G  A  K  R  E  I  Y  T  L  M  N  E  L  T  E  R  G  V  A  I  I
GGATGAGCCAACCAGAGGTGTGGATGTTGGCGCCAAACGAGAGATTTATACGCTGATGAATGAGCTGACCGAACGCGGTGTCGCTATCAT 3780

 M  V  S  S  E  L  P  E  I  L  G  M  S  D  R  I  I  V  V  H  E  G  R  I  S  G  E  I  H  A
CATGGTGTCATCAGAGCTTCCTGAAATTCTGGGAATGAGCGATCGGATTATCGTTGTCCATGAAGGCAGAATCAGCGGCGAAATCCATGC 3870
                                                      rbsC ->
 R  E  A  T  Q  E  R  I  M  T  L  A  T  G  G  R  *  M  K  T  E  Q  L  Q  T  E  Q  K  R
GCGAGAAGCAACACAAGAACGAATTATGACACTTGCCACGGGACGGCGGTAATATGAAAACGGAACAACTGCAAACAGAACAAAAACGGA 3960

I  R  F  D  G  V  M  Q  K  L  G  P  F  L  G  L  F  I  L  V  I  I  V  S  I  L  N  P  S  F
TTCGCTTCGACGGAGTCATGCAAAAACTCGGCCCGTTTCTTGGTTTATTTATTCTCGTTATCATTGTATCTATTTTAAATCCCAGCTTTC 4050

L  E  P  L  N  I  L  N  L  L  R  Q  V  A  I  N  G  L  I  A  F  G  M  T  F  V  I  L  T  G
TTGAACCGCTGAATATTTTAAACCTGCTTCGCCAGGTCGCCATTAACGGATTAATCGCGTTCGGGATGACCTTTGTTATTTTGACAGGCG 4140
```

Fig. 1. For legend see facing page.

2522–5542) and was cloned into pGEM7Zf(+) to generate pKW1 (Woodson, 1992). A plasmid bank of *Eco*RV fragments from λ_{rbs} was generated in pGEM5Zf(+) and was probed with pKW1. The plasmid pKW2 contained a 1133 bp fragment (Fig. 1, nucleotides 1736–2869). The overlap of pKW1 and pKW2 was confirmed by sequencing. To clone DNA upstream of

```
G  I  D  L  S  V  G  A  I  L  A  L  S  S  A  L  V  A  G  M  I  V  S  G  V  D  P  V  L  A
GCATTGATCTTTCTGTTGGCGCTATTCTTGCCCTGTCCAGTGCTTTAGTTGCGGGGATGATTGTGTCCGGTGTCGATCCGGTTCTCGCGA 4230

I  I  L  G  C  I  I  G  A  V  L  G  M  I  N  G  L  L  I  T  K  G  K  M  A  P  F  I  A  T
TCATCCTTGGCTGTATCATTGGTGCCGTACTAGGCATGATCAACGGATTATTGATTACTAAAGGAAAAATGGCGCCCTTTATCGCCACGC 4320

L  G  T  M  T  V  F  R  G  L  T  L  V  Y  T  D  G  N  P  I  T  G  L  G  T  N  Y  G  F  Q
TTGGCACCATGACTGTGTTTCGCGGACTGACGCTAGTGTATACAGATGGAAATCCGATTACCGGACTTGGCACAAACTACGGTTTTCAGA 4410

M  F  G  R  L  G  Y  F  L  G  I  P  V  P  A  I  T  M  V  L  A  F  V  I  L  W  V  L  L  H
TGTTCGGACGGCTCGGTTACTTTTTAGGCATTCCTGTACCGGCAATTACGATGGTTCTTGCCTTTGTCATCCTTTGGGTGCTTCTTCATA 4500

K T  P  F  G  R  R  T  Y  A  I  G  G  N  E  K  A  A  L  I   S  G  I  K  V  T  R  V  K  V
AAACACCATTCGGACGCCGAACGTACGCTATCGGCGGCAACGAAAAAGCCGCGCTCATTTCAGGCATCAAAGTGACGCGCGTGAAAGTGA 4590

M  I  Y  S  L  A  G  L  L  S  A  L  A  G  A  I  L  T  S  R  L  H  S  A  Q  P  T  A  G  E
TGATCTATTCTTTAGCCGGGCTTTTATCCGCTCTTGCAGGTGCCATATTGACTTCCCGCCTGCATTCGGCCCAGCCGACTGCGGGAGAAT 4680

S  Y  E  L  D  R  I  A  A  V  V  L  G  G  T  S  L  S  G  G  R  G  P  I  V  G  T  L  I  G
CGTACGAACTTGATCGTATCGCGGCAGTCGTCTTAGGAGGGACAAGTCTTTCCGGCGGCCGAGGACCGATTGTCGGCACGTTAATCGGGG 4770

V  L  I  I  G  T  L  N  N  G  L  N  L  L  G  V  S  S  F  Y  Q  L  V  V  K  G  I  V  I  L
TGCTGATCATCGGCACACTTAATAACGGACTTAATCTGCTTGGCGTCTCATCATTTTATCAGCTGGTTGTCAAAGGGATTGTTATCTTAA 4860
                                                                 rbsB  ->
I  A  V  L  L  D  R  K  K  S  A  *                                M  K  K  A  V  S  V  I  L  T  L  S  L  F  L
TTGCGGTATTGTTAGACCGCAAGAAGTCAGCTTAGGAGGGTTTTACATGAAAAAGGCTGTATCCGTCATTTTAACGTTATCATTATTTTT 4950
 ACYLATION  SITE
 L  T  A  C  S  L  E  P  P  N  G  K  R  S  N  S  G  N  K  K  E  F  T  I  G  L  S  V  S  T
GTTAACCGCCTGTTCGCTTGAGCCTCCCAATGGCAAACGATCAAACTCGGGGAACAAAAAGGAATTCACCATTGGCTTGTCCGTCTCAAC 5040

 L  N  N  P  F  F  V  S  L  K  K  G  I  E  K  E  A  K  K  R  G  M  K  V  I  I  V  D  A  Q
GCTTAATAATCCTTTTTTTGTCTCATTAAAAAAGGGTATCGAAAAAGAAGCTAAAAAACGGGGAATGAAAGTCATCATTGTTGATGCACA 5130

 N  D  S  S  K  Q  T  S  D  V  E  D  L  I  Q  Q  G  V  D  A  L  L  I  N  P  T  D  S  S  A
AAATGATTCATCGAAACAGACGAGTGACGTGGAAGATTTAATTCAGCAAGGTGTTGATGCATTATTAATCAACCCGACTGATTCTTCGGC 5220

 I  S  T  A  V  E  S  A  N  A  V  G  V  P  V  V  T  I  D  R  S  A  E  Q  G  K  V  E  T  L
GATCTCAACGGCAGTAGAATCTGCAAACGCCGTCGGTGTGCCCGTCGTAACAATCGATCGATCTGCGGAACAAGGAAAAGTTGAAACCCT 5310

V  A  S  D  N  V  K  G  G  E   M  A  A  A  F  I  A  D  K  L  G  K  G  A  K  V  A  E  L  E
CGTTGCTTCCGATAATGTAAAAGGCGGTGAAATGGCCGCGGCGTTTATTGCCGACAAACTTGGAAAAGGAGCAAAGGTGGCAGAGCTTGA 5400

 G  V  P  G  A  S  A  T  R  E  R  G  S  G  F  H  N  I  A  D  Q  K  L  Q  V  V  T  K  Q  S
AGGCGTCCCCGGCGCATCTGCCACACGGGAACGCGGCTCAGGATTCCATAACATCGCAGACCAAAAGCTCCAAGTTGTCACAAAACAATC 5490

 A  D  F  D  R  T  K  G  L  T  V  M  E  N  L  L  Q
AGCTGACTTTGACCGCACGAAAGGCCTGACTGTCATGGAAAACCTGCTGCAG
```

Fig. 1. DNA sequence and deduced amino acid sequence of the ribose transport operon from *B. subtilis*. The sequence of 5542 bp of the ribose transport operon is shown. There are six open reading frames in this operon (the sequence of *rbsB* is incomplete), each indicated by -> at the beginning and * at the end. Within the open reading frames the <--helix-turn-helix--> motif for RbsR, the ATP-binding sites for RbsA and the acylation site for RbsB are indicated above or below the sequence. The −35 and −10 regions of the SigA-type promoter are overlined; the catabolite repression associated sequence is underlined and the single initiation site of transcription within this sequence is marked by +1 over the G nucleotide shown in bold. The ribosome-binding site for the operon is shown overlined and in bold italics. The *csa-15–lacZ* fusion site is indicated by an arrow.

nucleotide 1736, an *Eco*RV–*Taq*I fragment (nucleotides 1736–1832, Fig. 2) was cloned into pJM783 (to give pKW3), and inserted into the chromosome of JH642. Chromosomal DNA from this integrant was digested with *Eco*RI, ligated and transformed into *E. coli* strain Tg1. Plasmids isolated from these transformants contained a 1832 bp *Eco*RI–*Taq*I fragment of chromosomal DNA (nucleotides 1–1832, Fig. 2) which was subcloned into pGEM7Zf(+) to give pKW5.

DNA manipulations and sequencing. DNA manipulation was carried out according to Sambrook *et al.* (1989). Enzymes were purchased from commercial suppliers and used according to the manufacturers' instructions. Sequencing was performed with a sequencing kit (Promega) using universal forward and reverse primers. Sequencing templates were prepared by a combination of subcloning and sequential deletion (Erase-A-Base, Promega). Gaps were filled in using oligonucleotide primers synthesized on an Applied Biosystems PCR-Mate DNA synthesizer. Primer extension was performed as described by O'Reilly *et al.* (1994).

Bacterial transformation. *B. subtilis* DNA transformations were carried out according to the method of Anagnostopoulos & Spizizen (1961). *E. coli* transformations were performed according to the method of Sambrook *et al.* (1989).

***β*-Galactosidase assays.** Strains of *B. subtilis* harbouring transcriptional *lacZ* fusions were grown in SM broth or on SM agar as appropriate. Samples were harvested at regular intervals

Table 2. Percentage amino acid identity between the open reading frames of the *B. subtilis* ribose transport operon and their homologous proteins

B. subtilis Rbs protein	Homologue (strain)	Function*	Percentage identity
RbsR	CcpA (*B. subtilis*)	Repressor	32
	DegA (*B. subtilis*)	Repressor-like protein	30
	RbsR (*E. coli*)	Repressor	31
	PurR (*E. coli*)	Repressor	29
	CytR (*E. coli*)	Repressor	30
	GalR (*E. coli*)	Repressor	29
RbsK	RbsK (*E. coli*)	Ribokinase	37
	RBSK (*S. cerevisiae*)	Ribokinase	30
	ScrK (*K. pneumoniae*)	Frucktokinase	27
RbsD	RbsD (*E. coli*)	MTP (ribose)	48
RbsA	RbsA (*E. coli*)	ATP-BP (ribose)	47
	MglA (*E. coli*)	ATP-BP (galactose)	42
	AraG (*E. coli*)	ATP-BP (arabinose)	41
RbsC	RbsC (*E. coli*)	MTP (ribose)	50
	AraH (*E. coli*)	MTP (arabinose)	38
	MglC (*E. coli*)	MTP (galactose)	39
RbsB†	RbsB (*E. coli*)	PBP (ribose)	49
	MglB (*E. coli*)	PBP (galactose)	24
	MglB (*C. freundii*)	PBP (galactose)	28
	ORF2 (*V. polymyxa*)‡	Gene upstream of β-glucanase	27
	DegA (*B. subtilis*)	Repressor-like protein	28

* MTP, membrane substrate-binding protein; ATP-BP, ATP-binding protein; PBP, periplasmic substrate-binding protein.

† Partial sequence.

‡ GenBank accession number M33791.

throughout the growth cycle and were assayed for β-galactosidase activity according to the method of Ferrari *et al* (1986). The specific activity is expressed as Miller units (Miller, 1972).

Computer analysis. Amino acid sequences were deduced from the nucleotide sequence using ANALYSEQ of the STADEN package (Staden, 1982). The GenBank database was accessed using ACNUC (Gouy *et al.*, 1985). Homology searches of the GenBank database were carried out using the TBLASTN program (Altschul *et al.*, 1990). Multiple sequence alignments were performed using the CLUSTAL V package (Higgins *et al.*, 1992).

RESULTS

Cloning and sequencing of the *csa-15* locus

The strain CSA15 was identified by the strategy outlined by O'Reilly *et al.* (1994). It contains a transcriptional fusion, *csa-15*, whose expression is negatively regulated by Spo0A. This locus has not yet been mapped on the *B. subtilis* chromosome. To further characterize the *csa-15* locus, DNA spanning this region of the chromosome was isolated and sequenced as outlined in Methods. A total of 5542 bp of sequence was obtained from a series of overlapping DNA fragments (Fig. 1). There are six open reading frames located within this DNA segment, each of which displays a high level of identity at the amino acid level to one of the six open reading frames comprising the ribose transport operon of *E. coli* (Bell *et al.*, 1986; Buckle *et al.*, 1986). The order of the cistrons in the *B. subtilis* operon is (promoter proximal to promoter distal): *rbsR* (repressor), *rbsK* (ribokinase), *rbsD* (a membrane transport protein), *rbsA* (an ATP-binding transport protein), *rbsC* (a membrane transport protein) and *rbsB* (a periplasmic substrate-binding protein). The percentage amino acid identity between each cistron of the *B. subtilis* ribose transport operon and homologous proteins from other bacteria is shown in Table 2. It is evident that the proteins encoded by the *B. subtilis* operon are more similar to those of the ribose transport operon of *E. coli* than to other members of each group, with homologies ranging between 47% and 50% for the four structural genes of the operon. The similarity between these proteins from *B. subtilis* and the *E. coli* ribose transport proteins is particularly evident when the hydropathy profiles are compared. In the case of the RbsC protein for example, the hydropathy profiles are very similar, with each protein having eight hydrophobic domains, suggesting that these proteins must have a very similar structure and function (Fig. 2). The RbsR protein belongs to a group of repressor proteins which includes RbsR and GalR from *E. coli*, and CcpA and DegA from *B. subtilis*. The level of identity between RbsR and this group of regulatory genes is lower

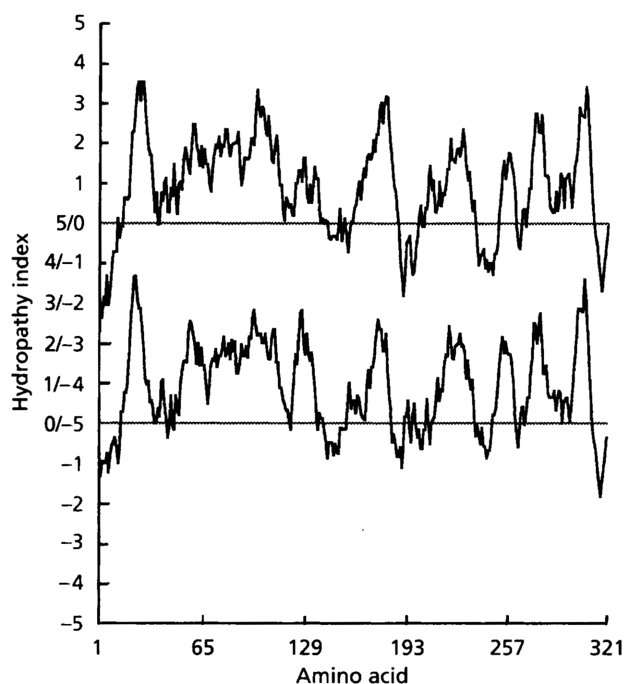

Fig. 2. Hydropathy plots of the RbsC proteins from the *B. subtilis* (upper trace) and *E. coli* (lower trace) ribose transport operons. Each trace is centred at zero (horizontal lines) and both are drawn on the same scale. The plots were calculated according to the method of Kyte & Doolittle (1982) using a window spacing of 10 amino acids. Positive and negative values represent hydrophobic and hydrophilic regions respectively. Double numbers (e.g. 4/–1, 3/–2) on the ordinate represent hydrophobic values for the lower trace and hydrophilic values for the upper trace. It is evident that each protein has eight hydrophobic domains whose extent and location within the two proteins are very similar.

(29–32%) than that observed for the structural genes. The region of greatest identity between these proteins is located within the helix-turn-helix motif present in each protein. The ribokinase displays characteristics of kinases and is clearly homologous to ribokinase from *E. coli.* These similarities support the hypothesis that the *csa-15* locus encodes a ribose transport system in *B. subtilis.*

Sequence analysis revealed that there was a SigA-like promoter positioned between nucleotides 224 and 253 (Fig. 1). There is a very strong ribosome binding site [nucleotides 258 and 272, $\Delta G = -17{\cdot}8$ kcal mol^{-1} ($-74{\cdot}5$ kJ mol^{-1})] located between the putative promoter and the start codon of the *rbsR* cistron. The open reading frames of the *B. subtilis* operon do not overlap and each has a strong ribosome-binding site. These vary in strength from $\Delta G = -17{\cdot}8$ kcal mol^{-1} ($-74{\cdot}5$ kJ mol^{-1}) for the *rbsR* cistron to $\Delta G = -11{\cdot}9$ kcal mol^{-1} ($-49{\cdot}8$ kJ mol^{-1}) for the *rbsD* cistron (calculated according to the method of Tinoco *et al.*, 1973). The sequence TGTAAACGGTTACA (between nucleotides 258 and 272), conforms to the consensus sequence TGWNANCGNTNWCA (where N is any base and W is A or T) associated with catabolite repression in a number of *Bacillus* operons (Weickert & Chambliss, 1990). This sequence motif is located between the –10 region of the putative promoter and the ribosome-binding site, a location similar to that found in other catabolite-repressed *B. subtilis* operons. There is a helix-turn-helix motif positioned between amino acids 2 and 23 of the RbsR protein with a value of 5·45, calculated according to the method of Dodd & Egan (1990), and supporting the view that it is a DNA-binding protein. There are two regions of RbsA in which a high level of identity at the amino acid level is observed between the *E. coli* and *B. subtilis* proteins. Both these conserved regions contain the amino acid motif GXXGXGR/K associated with protein–ATP-binding sites. The amino acid sequence LTACSL, located in the amino terminus of RbsB, is found in many lipoprotein precursors, and is the position at which cleavage and acylation (of the C) of proteins occurs (Gilson *et al.*, 1988). It is evident that the putative ribose transport operon identified in *B. subtilis* exhibits many of the features associated with the ABC group of transporters identified in Gram-negative bacteria.

Location of the promoter of the ribose transport operon

Analysis of the *csa-15–lacZ* operon fusion in strains KD882 (wild-type) and KD883 (*spo0A*) indicates that expression of this operon is negatively regulated by Spo0A (O'Reilly *et al.*, 1994). A strategy using integrating plasmids was used to identify the operon promoter and to locate the site at which expression is regulated by Spo0A. Two DNA fragments, *Eco*RI–*Cla*I (nucleotides 1–945, Fig. 1) and *Eco*RV–*Taq*I (nucleotides 1736–1832, Fig. 1), were separately subcloned into the plasmid pJM783 and integrated into the chromosomes of JH642 and JH646 by a Campbell-type insertion. The transcriptional operon–*lacZ* fusions thus generated had points of fusion located at nucleotide 945 in strains KD889 (wild-type) and KD890 (*spo0A*), and at nucleotide 1832 in strains KD887 (wild-type) and KD888 (*spo0A*). The levels and patterns of *β*-galactosidase expression throughout the growth cycle for these integrants were the same as those observed with *csa-15–lacZ* in strains KD882 (wild-type) and KD883 (*spo0A*) respectively (data not shown). Thus the operon promoter and site of regulation by Spo0A is positioned upstream of nucleotide 945. To further locate the operon promoter, a transcriptional fusion between the *Eco*RI–*Eco*RV fragment (1–493, Fig. 1) and *lacZ* was generated in pDG268. This plasmid was linearized and transformed into the chromosomes of strains JH642 and JH646 generating KD891 and KD892 respectively. Integration of this linearized plasmid occurs by a double-crossover event at the *α*-amylase locus. The pattern and level of *lacZ* expression observed in strains KD891 and KD892 are shown in Fig. 3. It is evident that the level of expression in strains harbouring a *spo0A* mutation is higher than that observed in wild-type cells. This pattern and level of expression is also observed with strains carrying transcriptional fusions within the operon. These data demonstrate that the ribose transport operon promoter, and the sequences through which control of operon expression by

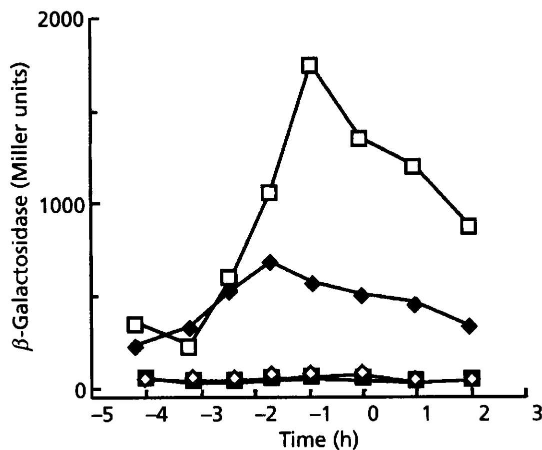

Fig. 3. Effect of glucose addition on the expression of *lacZ* in strains KD891 and KD892 during the growth cycle. Strains KD891 (wild-type) and KD892 (*spo0A*) were grown in SM broth in the presence and absence of 1% (w/v) glucose. Expression of the ribose transport operon was measured by determination of the levels of β-galactosidase (given as Miller units) present in cells at various stages of the growth cycle. Time is expressed as the number of hours before (minus) or after (plus) the entry of cells into the stationary phase (t_0). ◆, KD891, no glucose; □, KD892, no glucose; ◇, KD891, 1% glucose; ■, KD892, 1% glucose.

Spo0A is mediated, are located on the 493 bp *Eco*RI–*Eco*RV fragment. It is also evident that the transcription unit extends from within this fragment to the *Sau*3A site at nucleotide 3487, the site of the original transcriptional fusion in CSA15.

In order to verify that this operon encodes a ribose transport system, the ability of strains harbouring mutagenic and non-mutagenic insertions at this locus to grow on ribose as sole carbon source was tested. The *Eco*RI–*Cla*I fragment (nucleotides 1–945, Fig. 1), overlaps the promoter region of the operon; thus integration directed by this fragment (in strains KD889 and KD890) should be non-mutagenic. In contrast, insertions directed by the *Eco*RV–*Taq*I (nucleotides 1736–1832) fragment, as in strains KD887 and KD888, are mutagenic since it is located within the *rbsK* cistron of the operon. Thus the ability of wild-type strains (KD887 and KD889) and strains mutant in *spo0A* (KD888 and KD890) to grow on ribose as sole carbon source was tested. Strains JH642, JH646, KD889 and KD890, all of which have an intact ribose transport operon, grew on plates containing ribose as the sole carbon source. In contrast, strains KD887 and KD888, which do not have an intact ribose transport operon, did not grow under these conditions. These data support the hypothesis that this operon is involved in the transport of ribose into the cell.

Catabolite repression of the *rbs* operon in *B. subtilis*

Sequence analysis of the ribose transport operon reveals a motif often found in operons which are catabolite-repressed in *B. subtilis* (Weickert & Chambliss, 1990). In order to investigate expression of the *rbs* operon in the presence and absence of glucose, the levels of β-galactosidase in strains KD891 and KD892 were examined. Addition of glucose reduced the level of accumulated β-galactosidase activity to less than 50 U for both strains (Fig. 3). These low levels were observed at all stages of the growth cycle in both wild-type and *spo0A* backgrounds. The level of operon transcript was determined by primer extension analysis, using total RNA prepared from wild-type (JH642) and *spo0A* (JH646) strains grown in the presence and absence of glucose. Transcripts initiating at nucleotide 260 were observed in both JH642 and JH646 cells grown in the absence of glucose. However, transcripts were not observed in either strain at any stage of the growth cycle upon the addition of 1% glucose (data not shown). Thus transcription of the operon initiating at nucleotide 260, is subject to catabolite repression.

DISCUSSION

The sequence of the *csa-15* locus of *B. subtilis* (identified by O'Reilly *et al.*, 1994) indicates that it encodes a transport system with many of the characteristics of the ABC group of transporters originally identified in Gram-negative organisms (Ames, 1986; Higgins *et al.*, 1990). Among members of this group of transport operons, the highest level of identity is observed with the ribose transport operon identified from *E. coli* (see Table 2). Strains harbouring non-mutagenic and mutagenic insertions into the ribose transport operon were tested for growth on media containing ribose as sole carbon source. Only those strains in which this operon is intact showed detectable growth under these conditions. These growth patterns and the observed sequence homologies strongly suggest that this operon encodes proteins involved with ribose transport. The organization of the cistrons of the *B. subtilis* operon differs from that of its *E. coli* homologue. In *B. subtilis* the cistrons encoding the proteins involved in the transport of ribose, i.e. *rbsD*, *rbsA*, *rbsC* and *rbsB* are arranged in this order, which is the same as that in *E. coli*. However, the cistrons encoding the regulatory protein RbsR, and ribokinase RbsK, are located distal to *rbsB* in *E. coli* whereas they are encoded by the first two cistrons of the operon in *B. subtilis*. In addition to their different positioning relative to the structural genes of the operon, the order of the *rbs*R and *rbs*K cistrons in *B. subtilis* is reversed relative to that in *E. coli*.

A periplasmically located substrate-binding protein is characteristic of many members of the ABC group of transporters (Ames, 1986; Higgins *et al.*, 1990). This protein is not involved in the actual transport of substrate into the cell, but functions to bind substrate with very high affinity and deliver it to the transport complex. The RbsB protein from *B. subtilis* displays a high level of identity at the amino acid level, and has a similar hydropathy profile, to members of this family of proteins. Therefore RbsB from *B. subtilis* is probably functionally equivalent to its *E. coli* homologue. It is interesting that such proteins exist in *B. subtilis*, an organism which does

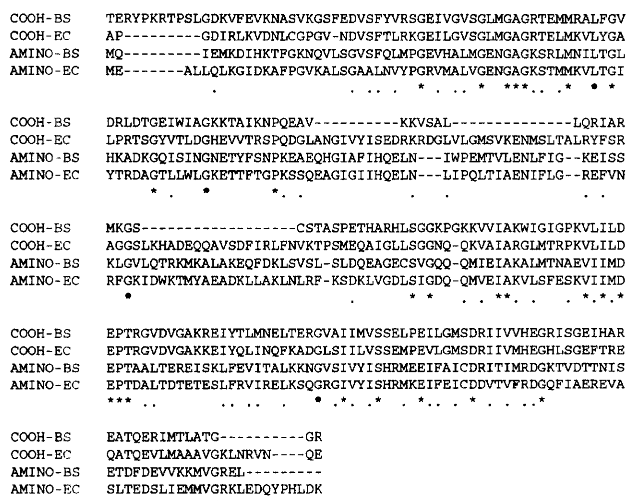

Fig. 4. Amino acid alignment of the amino and carboxy (COOH) terminal halves of the RbsA proteins from *B. subtilis* (BS) and *E. coli* (EC). The alignment was carried out using the CLUSTAL V program. Dashes within the amino acid sequence are placed to optimize the alignment. An asterisk below an amino acid denotes that it is present in all four sequences. A dot below a sequence denotes that a conserved amino acid is placed at this position in all four sequences.

not have a periplasmic space. Periplasmic substrate-binding proteins are also found in the dipeptide (DppA) and oligopeptide (OppA) transport systems of *B. subtilis* (Perego *et al.*, 1991; Rudner *et al.*, 1991; Mathiopoulos *et al.*, 1991). The RbsB protein from *B. subtilis* has a signal peptide and a sequence (L-T-A-C-S-L) in the vicinity of the amino terminus which is typical of proteins modified by acylation, features also found in OppA and DppA. A model proposed by Perego *et al.* (1991) suggests that this group of proteins in *B. subtilis* is secreted from the cytoplasm but remains attached to the membrane through this lipid anchor. Therefore, these proteins appear to be equivalent to their *E. coli* counterparts both in cellular location and in function.

The ABC group of transporters is defined by one or more of the component proteins having an ATP-binding cassette (Higgins *et al.*, 1990). Two ATP-binding sites are required for substrate transport. They can be located on separate proteins, as in the oligopeptide transport system of *S. typhimurium*, or on the same protein, as in the ribose transport system of *E. coli* (Ames, 1986). In addition, the amino- and carboxy-terminal halves of the RbsA protein are homologous to each other and can be aligned, a feature already noted by Ames (1986). The RbsA protein from *B. subtilis* shares many features with its *E. coli* homologue. There is a motif GXXGXGK/R which is associated with ATP-binding sites located in each half of RbsA from *B. subtilis*. In Fig. 4, the amino- and carboxy-terminal halves of the RbsA proteins from *E. coli* and *B. subtilis* are aligned. There is 45·5 % identity at the amino acid level between the amino-terminal halves, and 47 % identity between the carboxy-terminal halves of RbsA from *E. coli* and *B. subtilis*. A comparison of the amino- and carboxy-terminal halves of each protein reveals 26 % identity for the RbsA from *E. coli* and 23 % identity for the RbsA protein from *B. subtilis*. Thus the proposed duplication which gave rise to the mature RbsA protein probably occurred before the divergence of *E. coli* and *B. subtilis*.

Expression of the ribose transport operon from *B. subtilis* is directed by a single SigA-type promoter and is negatively regulated by Spo0A ~ P during the late exponential phase of the growth cycle (O'Reilly *et al.*, 1994). This regulation is not mediated indirectly through the transition state regulator AbrB (O'Reilly *et al.*, 1994). Expression of the dipeptide transport system is also regulated by Spo0A ~ P but in this case regulation is mediated indirectly through AbrB and begins at the transition stage of the growth cycle (Slack *et al.*, 1991). Thus there are significant differences in the mechanism and timing of Spo0A ~ P-regulated expression of the two transport operons. How then does Spo0A ~ P regulate expression of the *rbs* operon? A recognition sequence, TGNCGAA, is found in the vicinity of promoters whose expression is either negatively regulated by Spo0A ~ P (as in the case of *abrB*: Perego *et al.*, 1988; Strauch *et al.*, 1989) or positively regulated by Spo0A ~ P (as in the case of *spoIIE* and *spoIIG*: York *et al.*, 1992; Satola *et al.*, 1991). This sequence is not found in the promoter region of the *rbs* operon, suggesting that regulated expression mediated by Spo0A is indirect. The indirect regulation is not mediated through AbrB, however, which suggests that a novel pathway exists through which expression can be regulated by Spo0A.

Expression of the *rbs* operon in *B. subtilis* is catabolite-repressed. This has been demonstrated using *rbs–lacZ* fusions and by primer extension analysis. In the presence of 1 % glucose, levels of β-galactosidase in strains carrying *rbs–lacZ* fusions are reduced up to 40-fold, and transcripts could not be observed by primer extension analysis. This was observed at all stages of the growth cycle in both wild-type and *spo0A* cells. These observations suggest that the catabolite repression effect is dominant to regulated expression by Spo0A. A sequence motif frequently found in the vicinity of the transcriptional start site of *B. subtilis* operons which are catabolite-repressed (Weickert & Chambliss, 1990) overlaps the +1 of the *rbs* operon. It is noteworthy that the RbsR protein encoded by the first cistron of the ribose transport operon is homologous to members of the Gal repressor protein family. It may be significant that the CcpA protein, shown to be involved in catabolite repression of *amyE* (Henkin *et al.*, 1991), is also a member of this family and shows significant homology to the RbsR protein. It will thus be interesting to establish whether the RbsR protein plays a role in catabolite repression of the operon. It must also be noted that expression of *spo0A* is directed by two promoters, one of which is catabolite-repressed (Yamashita *et al.*, 1989; Chibazakura *et al.*, 1991; Strauch *et al.*, 1992). It will thus be interesting to investigate the interrelationships between Spo0A, RbsR, catabolite repression and temporal regulation of the *rbs* operon in *B. subtilis*.

ACKNOWLEDGEMENTS

K.W. was supported by BioResearch Ireland through the National Pharmaceutical Biotechnology Centre at Trinity College Dublin. We thank Paul Sharp and our colleagues at the Irish National Centre for Bioinformatics for assistance with sequence analysis. This work was funded by EOLAS grant SC/90/126 to K.M.D.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403–410.

Ames, G. F.-L. (1986). Bacterial periplasmic transport systems: structure, mechanism and evolution. *Ann Rev Biochem* **55**, 397–425.

Anagnostopoulos, C. & Spizizen, J. (1961). Requirements for transformation in *Bacillus subtilis*. *J Bacteriol* **81**, 741–746.

Bell, A. W., Buckel, S. D., Groarke, J. M., Hope, J. N., Kingsley, D. H. & Hermodson, M. A. (1986). The nucleotide sequences of the *rbsD*, *rbsA* and *rbsC* genes of *Escherichia coli* K12. *J Biol Chem* **261**, 7652–7658.

Buckle, S. D., Bell, A. W., Mohana Rao, J. K. & Hermodson, M. A. (1986). An analysis of the structure of the product of the *rbsA* gene of *Escherichia coli* K12. *J Biol Chem* **261**, 7659–7662.

Chibazakura, T., Kawamura, F. & Takahashi, H. (1991). Differential regulation of *spo0A* transcription in *Bacillus subtilis*: glucose represses promoter switching at the initiation of sporulation. *J Bacteriol* **173**, 2625–2632.

Dodd, I. B. & Egan, J. B. (1990). Improved detection of helix-turn-helix DNA binding motifs in protein sequences. *Nucleic Acids Res* **18**, 5019–5026.

Ferrari, E., Howard, S. M. H. & Hoch, J. A. (1986). Effect of stage 0 sporulation mutations on subtilisin expression. *J Bacteriol* **166**, 173–179.

Gilson, E., Alloing, G., Schmidt, T., Claverys, J.-P., Dudler, R. & Hofnung, M. (1988). Evidence for high affinity binding-protein dependent transport systems in Gram-positive bacteria and in *Mycoplasma*. *EMBO J* **7**, 3971–3974.

Gouy, M., Gautier, C., Attimonelli, M., Lanave, C. & diPaola, G. (1985). ACNUC – a portable retrieval system for nucleic acid sequence databases: logical and physical designs and usage. *CABIOS* **1**, 167–172.

Henkin, T. M., Grundy, F. J., Nicholson, W. L. & Chambliss, G. H. (1991). Catabolite repression of alpha-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.

Higgins, C. F., Hyde, S. C., Mimmack, M. M., Gileadi, U., Gill, D. R. & Gallagher, M. P. (1990). Binding protein-dependent transport systems. *J Bioenerg Biomembr* **22**, 571–591.

Higgins, D. G., Bleasby, A. J. & Fuchs, R. (1992). CLUSTAL V: improved software for multiple sequence alignment. *CABIOS* **8**, 189–191.

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

Mathiopoulos, C., Mueller, J. P., Slack, F. J., Murphy, C. G., Patankar, S., Bukusoglu, G. & Sonenshein, A. L. (1991). A *Bacillus subtilis* dipeptide transport system expressed early during sporulation. *Mol Microbiol* **5**, 1903–1913.

Miller, J. H. (1972). *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

O'Reilly, M., Woodson, K., Dowds, B. C. A. & Devine, K. M. (1994). The citrulline biosynthetic operon *argC–F* and a ribose transport operon *rbs* from *Bacillus subtilis* are negatively regulated by Spo0A. *Mol Microbiol* **11**, 87–98.

Perego, M., Spiegelman, G. B. & Hoch, J. A. (1988). Structure of the gene for the transition state regulator, *abrB*: regulator synthesis is controlled by the *spo0A* sporulation gene in *Bacillus subtilis*. *Mol Microbiol* **2**, 689–699.

Perego, M., Higgins, C. F., Pearce, S. R., Gallagher, M. P. & Hoch, J. A. (1991). The oligopeptide transport system of *Bacillus subtilis* plays a role in the initiation of sporulation. *Mol Microbiol* **5**, 173–185.

Rudner, D. Z., LeDeaux, J. R., Ireton, K. & Grossman, A. D. (1991). The *spo0K* locus of *Bacillus subtilis* is homologous to the oligopeptide permease locus and is required for sporulation and competence. *J Bacteriol* **173**, 1388–1398.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Satola, S., Kirchman, P. A. & Moran, C. P. (1991). Spo0A binds to a promoter used by sigA RNA polymerase during sporulation in *Bacillus subtilis*. *Proc Nat Acad Sci USA* **88**, 4533–4537.

Schaeffer, P., Miller, J. & Aubert, J. (1965). Catabolic repression of bacterial sporulation. *Proc Natl Acad Sci USA* **54**, 701–711.

Slack, F. J., Mueller, J. P., Strauch, M. A., Mathiopoulos, C. & Sonenshein, A. L. (1991). Transcriptional regulation of a *Bacillus subtilis* dipeptide transport operon. *Mol Microbiol* **5**, 1915–1925.

Staden, R. (1982). Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. *Nucleic Acids Res* **10**, 4731–4761.

Strauch, M. A., Spiegelman, G. B., Perego, M., Johnson, W. C., Burbulys, D. & Hoch, J. A. (1989). The transition state regulator *abrB* of Bacillus subtilis is a DNA binding protein. *EMBO J* **8**, 1615–1621.

Strauch, M. A., Trach, K. A., Day, J. & Hoch, J. A. (1992). Spo0A activates and represes its own synthesis by binding at its dual promoters. *Biochimie* **74**, 619–626.

Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. & Gralla, J. (1973). Improved estimation of secondary structure in ribonucleic acids. *Nature New Biol* **246**, 40–41.

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

Wood, H., Dawson, M., Devine, K. M. & McConnell, D. J. (1990). Characterisation of PBSX, a defective prophage of *Bacillus subtilis*. *J Bacteriol* **172**, 2667–2674.

Woodson, K. (1992). *Identification and characterisation of Spo0A controlled operons in Bacillus subtilis*. PhD thesis, University of Dublin.

Yamashita, S., Kawamura, F., Yoshikawa, H., Takahashi, H., Kobayashi, Y. & Saito, H. (1989). Dissection of the expression signals of the *spo0A* gene of *Bacillus subtilis*; glucose represses sporulation specific expression. *J Gen Microbiol* **135**, 1335–1345.

York, K., Kenney, T. J., Satola, S., Moran, C. P., Poth, H. & Youngman, P. (1992). Spo0A controls the SigA-dependent activation of *Bacillus subtilis* sporulation-specific transcription unit *spoIIE*. *J Bacteriol* **174**, 2648–2658.

Received 1 November 1993; revised 31 January 1994; accepted 15 February 1994.