A 35·7 kb DNA fragment from the *Bacillus subtilis* chromosome containing a putative 12·3 kb operon involved in hexuronate catabolism and a perfectly symmetrical hypothetical catabolite-responsive element

Carlo Rivolta,¹ Blazenka Soldo,¹ Vladimir Lazarevic,¹ Bernard Joris,² Catherine Mauël¹ and Dimitri Karamata¹

Author for correspondence: Dimitri Karamata. Tel: +41 21 3206075. Fax: +41 21 3206078.
e-mail: dimitri.karamata@igbm.unil.ch

¹ Institut de Génétique et de Biologie Microbiennes, Université de Lausanne, Rue César-Roux 19, CH-1005 Lausanne, Switzerland

² Centre d'Ingénierie des Protéines, Université de Liège, Institut de Chimie, B6, Sart Tilman, B-4000 Liège, Belgium

The *Bacillus subtilis* strain 168 chromosomal region extending from 109° to 112° has been sequenced. Among the 35 ORFs identified, *cotT* and *rapA* were the only genes that had been previously mapped and sequenced. Out of ten ORFs belonging to a single putative transcription unit, seven are probably involved in hexuronate catabolism. Their sequences are homologous to *Escherichia coli* genes *exuT*, *uidB*, *uxaA*, *uxaB*, *uxaC*, *uxuA* and *uxuB*, which are all required for the uptake of free D-glucuronate, D-galacturonate and β -glucuronide, and their transformation into glyceraldehyde 3-phosphate and pyruvate via 2-keto-3-deoxygluconate. The remaining three ORFs encode two dehydrogenases and a transcriptional regulator. The operon is preceded by a putative catabolite-responsive element (CRE), located between a hypothetical promoter and the RBS of the first gene. This element, the longest and the only so far described that is fully symmetrical, consists of a 26 bp palindrome matching the theoretical *B. subtilis* CRE sequence. The remaining predicted amino acid sequences that share homologies with other proteins comprise: a cytochrome P-450, a glycosyltransferase, an ATP-binding cassette transporter, a protein similar to the formate dehydrogenase α -subunit (FdhA), a protein similar to NADH dehydrogenases, and three homologues of polypeptides that have undefined functions.

Keywords: *Bacillus subtilis*, hexuronate catabolism, genome sequencing, catabolite repression

INTRODUCTION

Bacillus subtilis, a free-living soil bacterium, is able to use a large variety of natural compounds as sources of carbon and energy. In rich media containing sugars that can be rapidly metabolized, transport across the cell membrane and the degradation of alternative carbon sources are usually inhibited by the complex molecular mechanism of catabolite repression (Saier *et al.*, 1996).

Abbreviations: CRE, catabolite-responsive element; KDG, 2-keto-3-deoxygluconate; LR PCR, Long-Range PCR.

The GenBank accession number for the nucleotide sequence reported in this paper is AF015825.

Certain operons involved in degradation of these secondary carbon sources are regulated at the transcriptional level by a common repressor, CcpA, which binds to a conserved operator sequence, designated CRE for catabolite-responsive element (Hueck & Hillen, 1995).

As part of the European *B. subtilis* genome sequencing project, we determined the nucleotide sequence of the chromosomal region located between genes *cotT* (109·4°) (Aronson *et al.*, 1989) and *rapA* (112·4°), formerly known as *gsiAA* or *spo0L* (Mueller *et al.*, 1992) (Fig. 1). Computer analysis of this region revealed the presence of a putative operon involved in transport and degradation of extracellular hexuronate, preceded by a perfect CRE.

109°

Fig. 1. Genetic organization of the sequenced *B. subtilis* chromosome region. Arrows indicate ORFs and their orientation. Striped bars represent plasmid inserts rescued from chromosomal walking (pPS110, pPS303, pPS350 and pPS360), as well as the LR PCR product (subclones are not shown). Plasmids pPS120 and p6102.2, used as start points for chromosomal walking, are indicated as open bars. The positions of putative terminators are represented by the stem-loop symbol. The question mark indicates the uncertain terminator function attributed to the stem-loop structure downstream of *yjmA* (see text), and CRE indicates the location of the putative CRE. The scale, in kb, starts from the first nucleotide of the AF015825 sequence. The position of the region is reported in degrees. Restriction sites for *Bam*HI (B), *Eco*RI (E), *Sac*I (S) and *Sal*I (L) are shown.

METHODS

Cloning and sequencing. DNA for use as a template in the sequencing reactions was obtained either by chromosomal walking (Glaser *et al.*, 1993) on the *B. subtilis* 168 *trpC2* chromosome (Anagnostopoulos *et al.*, 1993) or by Long-Range PCR (LR PCR). Plasmids p6102.2 (Longchamp, 1995), containing a fragment of the *rapA* gene, and pPS120, carrying part of *cotT* obtained by PCR, were used as start points for chromosomal walking. From *rapA*, three successive rounds of plasmid rescue yielded pPS110, pPS350 and pPS360, covering about 15 kb chromosomal DNA (Fig. 1). From *cotT*, only pPS303 (4 kb) was recovered. Because the 16 kb fragment separating the two cloned regions was refractory to cloning, it was obtained by LR PCR (GeneAmp XL PCR kit; Perkin Elmer) with primers 5'-AATAGAAATCCGCATCCCGTG-AAGA-3' and 5'-CCTTTTCATCATTTTGTATTTCGTG-3'.

Plasmids used for rescue experiments were derivatives of either pDIA5304 (Glaser *et al.*, 1993) or pMTL20EC vectors (Oultram *et al.*, 1988; Chambers *et al.*, 1988). Rescued DNA was cloned in *Escherichia coli* TP610 (Hedegaard & Danchin, 1985), a strain that maintains plasmids at a low copy number, whereas subclones were propagated in strain JM83 (Yanisch-Perron *et al.*, 1985).

Restriction patterns of plasmids and the PCR product were

checked by Southern blot hybridization against strain 168 chromosomal DNA (data not shown).

Both strands were sequenced using an ABI 373 DNA sequencer, with a Taq Dye Terminator Cycle Sequencing kit (Perkin Elmer) and custom-synthesized DNA oligonucleotides.

Computer analysis. The final assembled sequence was analysed with computer tools contained in the GCG package (Devereux *et al.*, 1984). Ambiguous reading frame starts were checked with the GENEMARK program (Borodovsky & McIninch, 1993). RBS locations were automatically defined by complementarity with the *B. subtilis* 16S rRNA sequence. The predicted ORFs present in the sequenced region and their relative RBSs are shown in Table 1.

Stem-loop structures between genes transcribed in the same direction were considered as terminators only when the free energy of base pairing in the stems, calculated according to the rules of Freier *et al.* (1986), was below $-12 \text{ kcal mol}^{-1}$ ($-50.4 \text{ kJ mol}^{-1}$).

RESULTS AND DISCUSSION

Free D-glucuronate, D-galacturonate and β -glucuronide, as well as D-galacturonate resulting from the degra-

Table 1. Predicted ORFs present in the sequenced region

ORF	End points (nt)*	Size [aa (kDa)]	RBS (underlined) and start codons (bold)
CotT	1 < 114	37† (-)	<u>AGGGAGGTTTTCCTCT</u> TTG
YjeA	368 > 1771	467 (55.6)	<u>GAAGGAGTTT</u> TTGTTG
YjfA	1811 < 2284	157 (16.2)	<u>AAGGAGGTCATAAA</u> TTG
Yjfb	2409 < 2576	55 (6.0)	<u>AGGAGGTGGAAC</u> CATG
Yjfc	2703 > 3632	309 (35.3)	<u>AGGAGGGGAATCAG</u> GATG
YjgA	3610 < 4008	132 (14.8)	<u>GGGGTATCAATT</u> CGCAATG
YjgB	4109 < 4684	191 (20.8)	<u>AGGAGGGATTAAC</u> CATG
YjgC	4830 > 7787	985 (109.8)	<u>AGAAAGGGGAATGG</u> GAGATG
YjgD	7780 > 8340	186 (21.2)	<u>AAAGGGGGATGGG</u> CCATAATG
Yjha	8537 > 9178	213 (24.0)	<u>GGGGGAACAAT</u> CATG
Yjhb	9371 > 9883	170 (19.4)	<u>GGAGAGGTACGAAG</u> AGCTAATG
YjiA	9914 < 10192	92 (10.5)	<u>AGGTGAAGGAGAGA</u> ACAACGTG
YjiB	10583 > 11773	396 (45.0)	<u>AGGGGAGTGACCG</u> TACATG
YjiC	11796 > 12974	392 (44.0)	<u>AAGGAGACTGGAG</u> ATTTCATG
YjiA	13377 > 14189	270 (29.8)	<u>GGAGAGGGAACGG</u> GTTATG
YjkA	14235 < 14987	250 (27.4)	<u>GAAGGGAGGCACT</u> CGGTAATG
Yjkb	14987 < 15739	250 (27.8)	<u>AGGTGATTATTGT</u> TAATG
Yjla	15859 < 16833	324 (35.6)	<u>GAGGGATAT</u> CATG
Yjlb	16965 > 17462	165 (18.3)	<u>AAGGGAGGTGAAT</u> TGATG
Yjlc	17851 > 18273	140 (15.6)	<u>AAAGGAGAATTGG</u> ATATG
Yjld	18313 > 19491	392 (42.0)	<u>GGAGGATATAC</u> GATG
Yjma	19689 > 21110	473 (54.6)	<u>GGTGAGGAC</u> CATG
Yjmb	21178 > 22557	459 (50.4)	<u>GAGGTGGAAC</u> CATG
Yjmc	22662 > 23675	337 (36.4)	<u>AGGAAGGAGGCA</u> CTTGCTTG
Yjmd	23681 > 24700	339 (36.8)	<u>AGGTGAGTAAA</u> AATG
Yjme	24725 > 25804	359 (41.0)	<u>GGAGGTGCGAG</u> CATG
Yjmf	25801 > 26637	278 (29.5)	<u>GAGGTGGGCAT</u> ATG
Yjmg	26685 > 27953	422 (45.3)	<u>GGGGTGAGGTTG</u> AGATG
YjmH	28041 > 29042	333 (37.4)	<u>GGAAGGGAATT</u> AGATG
Yjmi	29119 > 30561	480 (55.4)	<u>GAGGGAGGCAATA</u> AGGGTG
YjmJ	30558 > 32051	497 (54.9)	<u>GGGGGAGAGGTAC</u> GTTG
Yjna	32090 < 32854	254 (27.1)	<u>AGGAGGCCCTTTAT</u> CATG
YjoA	33079 < 33543	154 (17.8)	<u>GGAGGAGATCATA</u> AATG
YjoB	33692 > 34963	423 (48.8)	<u>GGAGTATAAAC</u> CATG
RapA	35108 > 35739	210† (-)	<u>AGGGGGGATTA</u> AATG

* The transcriptional direction of the ORFs is indicated by >.

† Partial.

dation of plant pectin, were found to be alternative carbon sources for some bacteria. It has been reported that in *B. subtilis* the utilization of D-glucuronate is not common to all strains (Durand *et al.*, 1979), whereas, to our knowledge, growth in the presence of D-galacturonate has never been tested. We observed that *B. subtilis* strain 168 can grow in MS minimal medium (Sargent, 1973) devoid of glucose but supplemented with either D-glucuronate or D-galacturonate (0.2%) as a carbon source. As shown in Fig. 2, in *E. coli* (Ashwell *et al.*, 1960; Lin, 1996) and in *Erwinia chrysanthemi* (Hugouvieux-Cotte-Pattat & Robert-Baudouy, 1987) these sugars are imported into the cell by specific transporters and transformed via 2-keto-3-deoxygluconate (KDG) into pyruvate and glyceraldehyde 3-phosphate.

A search for proteins that are homologous to those deduced from the nucleotide sequence of the 35.7 kb region (Fig. 1) identified genes associated with the first steps of hexuronate-hexuronide metabolism, i.e. the uptake of these sugars (*yjmB* and *yjmG*) as well as their transformation into KDG (*yjmA*, *yjmE*, *yjmF*, *yjml* and *yjmJ*). The KDG transporter and the enzymes catalysing the subsequent steps of the metabolic pathway have previously been found to be encoded by genes located at 198° on the *B. subtilis* genetic map (Sorokin *et al.*, 1996; Pujic *et al.*, 1997).

A very high degree of homology allowed us to predict a function for the products of the genes reported above (Table 2). *YjmF* is an exception: it is very similar to a large number of generic oxidoreductases but, although

Fig. 2. Hexuronate, β -glucuronide and KDG degradation in *E. coli* (Lin, 1996). *E. coli* gene products are indicated as four-letter abbreviations, and *B. subtilis* homologues are in parentheses. Enzymes and transporters encoded by the sequenced putative ten-gene operon (Fig. 1) are in bold type. IN and OUT refer to intracellular and extracellular compartments, respectively; the question mark indicates the uncertain function attributed to YjmF. Encoded proteins: KdgT, KDG permease; ExuT, aldohexuronate permease; UidB, glucuronide permease; UidA, β -glucuronidase; UxaC, uronate isomerase; UxuB, mannonate oxidoreductase; UxaB, altronate oxidoreductase; UxuA, mannonate hydrolase; UxaA, altronate hydrolase; KdgK, KDG kinase; KdgA, 2-keto-3-deoxy-6-phosphogluconate aldolase.

it is likely to be a mannonate oxidoreductase, it shares only 19.6% identity and 46.7% similarity with that enzyme from *E. coli*.

Whereas the *E. coli* genes required for hexuronate conversion to KDG are organized in operons and regulons occupying different chromosomal positions (Stoeber *et al.*, 1974; Lin, 1996), in *B. subtilis* they all seem to belong to the *yjm* operon, a single transcription unit of ten ORFs. The expression of the latter is presumably modulated by the CRE-CcpA catabolite repression system (Henkin *et al.*, 1991; Fujita *et al.*, 1995). A 26 bp palindrome, located upstream of this operon and matching the predicted CRE sequence (Weickert & Chambliss, 1990), is probably the first example of a perfectly symmetrical CRE (Table 3). Experiments based on site-directed mutagenesis of the *amyE* gene operator demonstrated that the better the symmetry of the operator, the stronger the repression (Weickert & Chambliss, 1990). Unfortunately, base substitutions aimed at generating the putative CRE sequence presented here were not done. Therefore,

experimental evidence favouring the latter as an optimal CRE has not been obtained.

In addition to encoding homologues of the gene products of *exuT*, *uidB*, *uxaABC* and *uxuAB*, operon *yjm* also encodes YjmC, D and H. YjmC and YjmD are possibly two generic dehydrogenases; however, they do not seem to play any obvious role in uronic sugar catabolism. The putative transcriptional regulator YjmH, which belongs to the LacI family and is similar to the *B. subtilis* CcpA repressor, could potentially be involved in the regulation of the expression of the relevant operon.

The stem-loop structure immediately downstream of *yjmA* (Table 4), with a ΔG value of $-18.2 \text{ kcal mol}^{-1}$ ($-76.1 \text{ kJ mol}^{-1}$), is not necessarily a terminator because the distance separating it from the next downstream RBS (22 bp) is probably too short to accommodate a promoter sequence.

Of the remaining genes located in this region, only eight produced a BLAST score (Altschul *et al.*, 1990) above 100. Protein YjgC shares a very high homology with the

Table 2. Comparison of the ORF products to proteins in available databases

ORF*	Similar protein in database†	Database entry	BLAST score and probability
YjeA	Hypothetical 15.0 kDa protein from <i>Bacillus licheniformis</i> (Laoide <i>et al.</i> , 1989)	M26412	309 (1.3e-35)
YjgC	Formate dehydrogenase α -subunit from <i>Moorella thermoacetica</i>	U73807	398 (3.9e-238)
YjiB	Cytochrome P-450 109 (ORF405) from <i>B. subtilis</i> (Ahn & Wake, 1991)	P27632	721 (2.5e-112)
YjiC	Macrolide glycosyltransferase from <i>Streptomyces lividans</i> (Jenkins & Cundliffe, 1991)	M74717	233 (3.3e-64)
YjkA	Predicted coding region MJ0938 from <i>Methanococcus jannaschii</i> (Bult <i>et al.</i> , 1996)	U67537	224 (2.5e-40)
YjKB	Protein GlnQ; glutamine transport ATP-binding protein from <i>Bacillus stearothermophilus</i> (Wu & Welker, 1991)	P27675	118 (2.1e-47)
YjlA	Hypothetical 42.1 kDa protein from <i>Bacillus</i> sp.	P29156	376 (2.9e-101)
YjID	NADH dehydrogenase from <i>E. coli</i> (Oshima <i>et al.</i> , 1996)	P00393	127 (1.5e-32)
YjMA	Uronate isomerase from <i>E. coli</i>	D13328	901 (2.2e-176)
YjMB	Glucuronide permease from <i>E. coli</i>	P30868	153 (5.2e-62)
YjMC	Malate dehydrogenase from <i>Methanothermobacter fervidus</i> (Honka <i>et al.</i> , 1990)	X51840	263 (7.0e-71)
YjMD	Sorbitol dehydrogenase from <i>B. subtilis</i> (Ng <i>et al.</i> , 1992)	Q06004	145 (4.5e-55)
YjME	D-Mannonate hydrolase from <i>E. coli</i> (Burland <i>et al.</i> , 1995)	P24215	275 (1.1e-69)
YjMF	Gluconate oxidoreductase from <i>Gluconobacter oxydans</i> (Klasen <i>et al.</i> , 1995)	X80019	134 (2.9e-26)
YjMG	Aldohexuronate transport system from <i>E. coli</i>	D13328	262 (5.6e-31)
YjMH	RBS repressor from <i>E. coli</i> (Burland <i>et al.</i> , 1993)	L10328	301 (1.9e-57)
YjMI	Altronate oxidoreductase from <i>E. coli</i>	D13327	604 (1.3e-146)
YjMJ	Altronate dehydratase from <i>E. coli</i>	D13328	1378 (1.8e-198)

* Only ORFs producing BLAST scores above 100, from a search performed with default settings, are reported.

† The reported sequence is that showing the highest score among proteins with an experimentally defined function. In the absence of relevant biochemical data, the most similar putative polypeptide was retained.

Table 3. Sequence alignment of previously determined CRE motifs with the palindrome located upstream of the *yjMA* gene

Aligned sequences are from Hueck & Hillen (1995). Mismatches to the consensus sequence (Weickert & Chambliss, 1990) are indicated in lower-case type. CRE position refers to the distance between the first 'T' of the consensus sequence and the translation start site. The relative positions of the *yjMA* hypothetical promoter and of the putative *yjMA* CRE on sequence AF015825 correspond to nucleotides 19601 to 19629 and 19639 to 19664, respectively.

Gene	CRE sequence	Position (bp)
<i>acsA</i>	TGAAAGCGTTAcCA	+8
<i>acuA</i>	TGAAAACGCTTTat	-61
<i>amyE</i>	TGTAAGCGTTAACA	-124
<i>gntR</i>	TGAAAGCGGTAcCA	+107
<i>hutP</i>	TGAAACCGCTTcCA	+170
<i>licS</i>	aGAAAACGCTTTCA	-33
<i>xylA</i>	TGgAAGCGTaaACA	-93
<i>yjMA</i>	TCAAAATGTtAACGTTAACATTTTGA	-44
Consensus	TGWAANCGNTNWCA	

formate dehydrogenase α -subunit of *Moorella thermoacetica* and many other bacteria. The products of the probably co-transcribed *yjiB* and *yjiC* genes are very similar to P-450 cytochromes and glycosyltransferases, respectively. Interestingly, in *Streptomyces fradiae* a P-450 gene involved in the biosynthesis of the macrolide antibiotic tylosin is also located immediately upstream

of a gene encoding a hypothetical amino sugar transferase (Merson-Davies & Cundliffe, 1994).

YjKB belongs to the family of ATP-binding cassette transporters and is highly similar to GlnQ, a *Bacillus stearothermophilus* protein required for glutamine transmembrane transport. Although YjID shares rel-

Table 4. Putative ρ -independent transcription terminators found in the *cotT-rapA* region

Stem-loop structure	End points (nt)	ΔG [kcal mol ⁻¹ (kJ mol ⁻¹)]	Upstream ORF
<pre> G 5' AGAAAAGGCUGUCC U UUUUUUCCGACAGG A 3' UUUAUC C </pre>	1775–1806	-16.6 (-69.5)	<i>yjeA</i>
<pre> ACA 5' AAAAAGCGUC U UUUUUCGCGG A 3' UUAG CGAC </pre>	3602–3630	-9.7 (-40.6)	<i>yjfC</i>
<pre> - U 5' ACCCGAGC GGC U UGGGCUCC CCG C 3' UUUUU A A </pre>	8347–8373	-16.5 (-69.0)	<i>yigD</i>
<pre> U 5' UGCACGUUUUCA C ACGUGCAAAGGU U 3' UAUUUUUU G </pre>	9194–9223	-13.6 (-56.9)	<i>yjhA</i>
<pre> U 5' UAAAAAAGAAAGGCC G AUUUUUUCUUUCUCGG U 3' UGUAAU G </pre>	9881–9916	-17.5 (-73.2)	<i>yjhB</i>
<pre> A 5' AAAACCGAAUC A UUUUGGCUUAG U 3' UUUGUC UU </pre>	12 982–13 010	-12.6 (-52.7)	<i>yjiC</i>
<pre> AA CC 5' ACAA AGCCGCUUUU C UGUU UCGGCGAAAG U 3' UUACUGUC -U GC </pre>	14 195–14 231	-16.6 (-69.5)	<i>yjjA</i>
<pre> GA 3' UGUUUUUU UCCUCUUGUCC A 5' AGGAGAAUGGG A UG </pre>	15 824–15 851	-12.9 (-54.0)	<i>yjlA</i>
<pre> -A CU 5' GGCAUUGCCA AAGCAUUC U CUGUAAACGGU UUCGUGAG U 3' UUUUUUUU CU CC </pre>	19 508–19 552	-22.4 (-93.7)	<i>yjID</i>
<pre> - U 5' AGUG UCCGCUCAG A UCAC AGGCAGAGUC U 3' AGCCUA U GC </pre>	21 114–21 145	-18.2 (-76.1)	<i>yjmA</i>
<pre> UU 5' AAGGCUAUCC C UUCCGAUAGG A 3' UUACUUC CGA </pre>	32 058–32 084	-13.7 (-57.3)	<i>yjmJ</i>
<pre> GA 3' UUUCGCG UUUUUUUUCUAUGGA A 5' AGAAAAAGAUACCU U GU </pre>	33 035–33 070	-14.2 (-59.4)	<i>yjoA</i>
<pre> GUU 5' AAGAAAGCACGGGU U UUUUUUCGUGCCCA G 3' GUUG AAAA </pre>	34 964–35 000	-18.4 (-77.0)	<i>yjoB</i>

evant similarities with NADH dehydrogenases from *E. coli* and *Haemophilus influenzae*, it probably does not catalyse the same enzymic reaction.

Finally, proteins YjeA, YjkA and YjlA show high homologies with other putative bacterial polypeptides with unidentified biochemical functions.

ACKNOWLEDGEMENTS

We are grateful to Drs F. Denizot and J. Haiech for valuable help in automated sequencing, and to T. Junier for useful suggestions about the UNIX environment.

This work was supported by grant 96.0245 from the Office Fédéral de l'Éducation et de la Science (Switzerland) and grant BIO4-CT96-0655 from the European Commission.

REFERENCES

- Ahn, K. S. & Wake, R. G. (1991). Variations and coding features of the sequence spanning the replication terminus of *Bacillus subtilis* 168 and W23 chromosomes. *Gene* 98, 107–112.
- 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.
- Anagnostopoulos, C., Piggot, P. J. & Hoch, J. A. (1993). The genetic map of *Bacillus subtilis*. In *Bacillus subtilis and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, pp. 425–461. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology.
- Aronson, A. I., Song, H. Y. & Bourne, N. (1989). Gene structure and precursor processing of a novel *Bacillus subtilis* spore coat protein. *Mol Microbiol* 3, 437–444.
- Ashwell, G., Wahba, J. A. & Hickman, J. (1960). Uronic acid metabolism in bacteria. *J Biol Chem* 235, 1559–1565.
- Borodovsky, M. & McIninch, J. D. (1993). GENEMARK: parallel gene recognition for both DNA strands. *Comput Chem* 17, 123–133.
- Bult, C. J., White, O., Olsen, G. J. & 20 other authors (1996). Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* 273, 1058–1073.
- Burland, V., Plunkett, G., III, Daniels, D. L. & Blattner, F. R. (1993). DNA sequence and analysis of 136 kilobases of the *Escherichia coli* genome: organizational symmetry around the origin of replication. *Genomics* 16, 551–561.
- Burland, V., Plunkett, G., III, Sofia, H. J., Daniels, D. L. & Blattner, F. R. (1995). Analysis of the *Escherichia coli* genome VI: DNA sequence of the region from 92.8 through 100 minutes. *Nucleic Acids Res* 23, 2105–2119.
- Chambers, S. P., Prior, S. E., Barstow, D. A. & Minton, N. P. (1988). The pMTL *nic*-cloning vectors. I. Improved pUC polylinker regions to facilitate the use of sonicated DNA for nucleotide sequencing. *Gene* 68, 139–149.
- Devereux, J., Haeblerli, P. & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* 12, 387–395.
- Durand, M., Pichinoty, F., Job, C. & Mandel, M. (1979). Nutrition carbonée et étude taxonomique de *Bacillus subtilis* et *B. licheniformis*. *Can J Microbiol* 25, 491–498.
- Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T. & Turner, D. H. (1986). Improved free-energy parameters for predictions of RNA duplex stability. *Proc Natl Acad Sci USA* 83, 9373–9377.
- Fujita, Y., Miwa, Y., Galinier, A. & 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.
- Glaser, P., Kunst, F., Arnaud, M. & 14 other authors (1993). *Bacillus subtilis* genome project: cloning and sequencing of the 97 kb region from 325° to 333°. *Mol Microbiol* 10, 371–384.
- Hedegaard, L. & Danchin, A. (1985). The *cya* gene region of *Erwinia chrysanthemi* B374: organisation and gene products. *Mol Gen Genet* 201, 38–42.
- 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.
- Honka, E., Fabry, S., Niermann, T., Palm, P. & Hensel, R. (1990). Properties and primary structure of the L-malate dehydrogenase from the extremely thermophilic archaeobacterium *Methanothermus fervidus*. *Eur J Biochem* 188, 623–632.
- Hueck, C. J. & Hillen, W. (1995). Catabolite repression in *Bacillus subtilis*: a global regulatory mechanism for the gram-positive bacteria? *Mol Microbiol* 15, 395–401.
- Hugouvieux-Cotte-Pattat, N. & Robert-Baudouy, J. (1987). Hexuronate catabolism in *Erwinia chrysanthemi*. *J Bacteriol* 169, 1223–1231.
- Jenkins, G. & Cundliffe, E. (1991). Cloning and characterization of two genes from *Streptomyces lividans* that confer inducible resistance to lincomycin and macrolide antibiotics. *Gene* 108, 55–62.
- Klasen, R., Bringer-Meyer, S. & Sahm, H. (1995). Biochemical characterization and sequence analysis of the gluconate:NADP 5-oxidoreductase gene from *Gluconobacter oxydans*. *J Bacteriol* 177, 2637–2643.
- Laoide, B. M., Chambliss, G. H. & McConnell, D. J. (1989). *Bacillus licheniformis* alpha-amylase gene, *amyL*, is subject to promoter-independent catabolite repression in *Bacillus subtilis*. *J Bacteriol* 171, 2435–2442.
- Lin, E. C. C. (1996). Dissimilatory pathways for sugars, polyols, and carboxylates. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd edn, vol. 1, pp. 307–342. Edited by F. C. Neidhardt and others. Washington, DC: American Society for Microbiology.
- Longchamp, P. (1995). *Génétique des endolysines associées à l'induction des prophages défectifs de type PBSX chez Bacillus subtilis*. PhD thesis, University of Lausanne.
- Merson-Davies, L. A. & Cundliffe, E. (1994). Analysis of five tylosin biosynthetic genes from the *tylBA* region of the *Streptomyces fradiae* genome. *Mol Microbiol* 13, 349–355.
- Mueller, J. P., Bukusoglu, G. & Sonenshein, A. L. (1992). Transcriptional regulation of *Bacillus subtilis* glucose starvation-inducible genes: control of *gsiA* by the ComP–ComA signal transduction system. *J Bacteriol* 174, 4361–4373.
- Ng, K., Ye, R., Wu, X. C. & Wong, S. L. (1992). Sorbitol dehydrogenase from *Bacillus subtilis*. Purification, characterization, and gene cloning. *J Biol Chem* 267, 24989–24994.
- Oshima, T., Aiba, H., Baba, T. & 22 other authors (1996). A 718-kb DNA sequence of the *Escherichia coli* K-12 genome corresponding to the 12.7–28.0 min region on the linkage map. *DNA Res* 3, 137–155.
- Oultram, J. D., Peck, H., Brehm, J. K., Thompson, D. E., Swinfield, T. J. & Minton, N. P. (1988). Introduction of genes for leucine biosynthesis from *Clostridium pasteurianum* into *C. acetobutylicum* by cointegrate conjugal transfer. *Mol Gen Genet* 214, 177–179.
- Pujic, P., Dervyn, R., Sorokin, A. & Ehrlich, S. D. (1997). Analysis of the *kdgRKAT* operon involved in D-galacturonate metabolism in *Bacillus subtilis*. In *9th International Conference on Bacilli*, abstract F99. Lausanne: IGBM, University of Lausanne.
- Saier, M. H., Jr, Chauvaux, S., Cook, G. M., Deutscher, J., Paulsen,

I. T., Reizer, J. & Ye, J. J. (1996). Catabolite repression and inducer control in Gram-positive bacteria. *Microbiology* **142**, 217–230.

Sargent, M. G. (1973). Synchronous cultures of *Bacillus subtilis* obtained by filtration with glass fiber filters. *J Bacteriol* **116**, 736–740.

Sorokin, A., Azevedo, V., Zumstein, E., Galleron, N., Ehrlich, S. D. & Serror, P. (1996). Sequence analysis of the *Bacillus subtilis* chromosome region between the *serA* and *kdg* loci cloned in a yeast artificial chromosome. *Microbiology* **142**, 2005–2016.

Stoeber, F., Lagarde, A., Nemoz, G., Novel, G., Novel, M., Portalier, R., Pouyssegur, J. & Robert-Baudouy, J. (1974). Le métabolisme des hexuronides et des hexuronates chez *Escherichia coli* K12: aspects physiologiques et génétiques de sa régulation. *Biochimie* **56**, 199–213.

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.

Wu, L. & Welker, N. E. (1991). Cloning and characterization of a glutamine transport operon of *Bacillus stearothermophilus* NUB36: effect of temperature on regulation of transcription. *J Bacteriol* **173**, 4877–4888.

Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119.

Received 18 September 1997; revised 14 November 1997; accepted 18 November 1997.