Nucleotide sequence of the *Bacillus subtilis* temperate bacteriophage SP β c2

Vladimir Lazarevic,¹ Andreas Düsterhöft,² Blazenka Soldo,¹
Helmut Hilbert,² Catherine Mauël¹ and Dimitri Karamata¹

Author for correspondence: Dimitri Karamata. Tel: +41 21 320 60 75. Fax: +41 21 320 60 78.
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 Qiagen GmbH, Max-Volmer-Straße 4, D-40724 Hilden, Germany

The *Bacillus subtilis* 168 chromosomal region extending from 184° to 195°, corresponding to prophage SP β , has been completely sequenced using DNA of the thermoinducible SP β c2 mutant. This 134416 bp segment comprises 187 putative ORFs which, according to their orientation, were grouped into three clusters. Compared to its host, SP β c2 is characterized by a lower G+C content, shorter mean ORF length, as well as a different usage of start codons. Nearly 75% of predicted ORFs do not share significant homologies to sequences in available databases. The only highly similar proteins to SP β c2-encoded ones are host paralogues. SP β c2 promoter regions contain SOS box consensus sequences and a repeated motif, designated SP β repeated element (SPBRE), that is absent from the host genome. Gene *sspC*, encoding the small acid-soluble protein C, that has been previously sequenced and mapped to the vicinity of the SP β region, was found to be part of the prophage.

Keywords: *Bacillus subtilis*, SP β c2 prophage, temperate bacteriophage, genome sequencing

INTRODUCTION

Two entities residing in the chromosome of the *Bacillus subtilis* reference strain 168 are considered to be remnants of ancestral temperate bacteriophages: the *skin* element, whose excision leads to the reconstitution of the *sigK* gene encoding the late-sporulation specific σ^K factor (Stragier *et al.*, 1989; Kunkel *et al.*, 1990), and defective phage PBSX (Seaman *et al.*, 1964), which directs the formation of phage particles mainly containing chromosomal DNA (Okamoto *et al.*, 1968). Strain 168 is also lysogenic for phage SP β (Warner *et al.*, 1977), whose genome is inserted close to the host replication terminus, between *ilvA* and *kauA* (Zahler *et al.*, 1977). In cultures of SP β lysogenic strains, phage induction can occur at a very low frequency ($\sim 10^{-6}$), while more extensive induction can be achieved by treatment with mitomycin C or *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (Warner *et al.*, 1977). In addition, strains lysogenic for SP β c2 (Rosenthal *et al.*, 1979), a mutant carrying a thermoinducible repressor gene, can be massively induced by a brief heat shock.

Previous studies of SP β led to the identification of genes involved in DNA methylation (Trautner *et al.*, 1980), control of phage lysogeny and immunity (McLaughlin *et al.*, 1986), and endolysin production (Regamey & Karamata, 1998). We report the complete nucleotide sequence of the SP β c2 prophage and summarize the main features of identified ORFs.

METHODS

Bacteria and plasmids. *Escherichia coli* strains DH5 α (Hanahan, 1983) and TP610 (Glaser *et al.*, 1993) were used as hosts for plasmids obtained by cloning of SP β c2 *EcoRI* fragments in double-stranded vectors pMTL20EC (Chambers *et al.*, 1988) and pDIA5304 (Glaser *et al.*, 1993). *B. subtilis* strains CU1147 (Rosenthal *et al.*, 1979) and CU1050 (= *su*⁺) (Warner *et al.*, 1977) were used for SP β c2 induction and amplification, respectively.

DNA purification. For manual sequencing, plasmid DNA was prepared by the boiling method (Del Sal *et al.*, 1988). For automated sequencing, plasmids were purified using QIA-GEN-tip 100 columns and the QIAwell 8 Plasmid Kit (Qiagen). PCR products were purified using the QIAquick PCR Purification Kit (Qiagen).

Transformation. *E. coli* DH5 α competent cells were prepared and transformed by the procedure of Chung & Miller (1988).

Abbreviation: SPBRE, SP β repeated element.

The GenBank/EMBL/DBJ accession number for the nucleotide sequence reported in this paper is AF020713.

Competent cells of the *E. coli* strain TP610 were prepared and transformed by the calcium chloride method (Sambrook *et al.*, 1989).

DNA sequencing. The nucleotide sequence of plasmids pPS097, pPS554, pPS103 and pPS610 was determined by dideoxy chain termination with the Sequenase Version 2.0 Kit (USB) and [³⁵S]dATP α -S (Amersham). For other plasmids, ABI dye-terminators and *Taq* FS DNA polymerase (Perkin Elmer) were used for cycle sequencing reactions. Automated set-up of sequencing reactions was carried out on the BioRobot 9600 laboratory workstation (Qiagen). Automated sequencing was done on an ABI 377 DNA sequencer (Perkin Elmer). Plasmid walking was initiated with M13 forward and reverse primers and continued with custom-made oligonucleotides, and was applied to all of the clones with the exception of pPS232, which was sequenced with the shotgun strategy. Assembly, editing and finishing of data were carried out with the SeqMan II module of the Lasergene software package (DNASTAR). The final sequence was analysed by the University of Wisconsin Genetics Computer Group software (Devereux *et al.*, 1984).

Preparation of SP β c2. A stock of SP β c2 was obtained by a heat shock of a strain CU1147 culture. The DNase-treated lysate was used to infect the SP β -cured strain CU1050, grown to an OD₅₉₅ of 0.3 in LB medium supplemented with 0.1% (w/v) glucose, 10 mM CaCl₂ and 10 mM MgCl₂. The incubation at 37 °C was continued until lysis. Phage DNA was isolated from the lysate with the Qiagen Lambda Midi Kit.

PCR amplification. PCRs were set up with 0.1 ng SP β c2 genomic DNA, 100 pmol each primer, 20 nmol each dNTP (Pharmacia) in 100 μ l reaction buffer (Pharmacia) containing 2.5 units *Taq* DNA polymerase (Pharmacia). Reactions were run with denaturation for 2 min at 95 °C, followed by 30 cycles of amplification (95 °C for 30 s, 45 °C for 1 min, 72 °C for 1 min per kb of the segment to be amplified) and a final extension for 10 min at 72 °C. The 6 kb region containing the segment encoding the phage cohesive ends was amplified using the primers 5' CGTTACTTTCTCTGAATCCTT 3' and 5' GTTTGAATAGTCATATCGCTTATCC 3', 1 ng of *B. subtilis* strain CU1147 chromosomal DNA and the Expand Long Template PCR System (Boehringer Mannheim), according to the manufacturer's recommendations.

RESULTS AND DISCUSSION

Sequencing strategy

Linear DNA extracted from SP β c2 phage particles was digested with *Eco*RI and separated through agarose gels. Non-overlapping fragments were cloned into plasmid vectors and sequenced. To determine the order of sequenced inserts, oligonucleotides corresponding to sequences relatively close to the extremities of each of the inserts were used to link neighbouring fragments by PCR, the products of which were directly sequenced. Gaps corresponding to uncloned SP β c2 *Eco*RI fragments represented nearly 15% of the phage genome. They were bridged by direct sequencing of PCR products generated from phage DNA. The sequence comprising the two extremities of linear phage DNA, released by *Eco*RI digestion, was determined from a single PCR product generated on chromosomal DNA of CU1147, the SP β c2 lysogenic strain. All sequenced fragments were

assembled in a contiguous DNA segment corresponding to the SP β c2 prophage.

DNA features

The 134 kb SP β c2 genome is one of the largest among phages, and to our knowledge the longest phage genome whose sequence has been deposited in the sequence databases. Comparison of the previously determined restriction map of SP β c2, obtained by DNA digestion with *Eco*RI, *Sal*I, *Sac*I, *Ava*I, *Ava*II and *Bam*HI (Fink & Zahler, 1982; Spancake *et al.*, 1987), with that deduced from the nucleotide sequence did not reveal important discrepancies. However, comparison of the theoretical to the experimentally determined *Pst*I restriction patterns revealed that, in DNA isolated from the phage particles, nearly half of the 40 *Pst*I sites were modified and protected from restriction (see below).

Several SP β c2 segments of more than 250 nt exhibit over 90% identity with *B. subtilis* chromosomal regions (Table 1). Their orientation with respect to the replication fork movement is always identical to that of their host counterparts. Homologies between the SP β prophage and chromosomal segments have been postulated to play a role in spontaneous deletions of parts of the *B. subtilis* chromosome as well as in the integration of integrase-deficient phage mutants, which would have to be forcibly inserted into the host chromosome by a Campbell-type mechanism (Zahler, 1993). Indeed, homologies to chromosomal regions near 556 and 1880 kb from the replication origin (Kunst *et al.*, 1997) could explain observed SP β insertions near the *dal* marker and between *thyA* and *glnA*, respectively (Zahler, 1982; Gardner *et al.*, 1982). Similarity to chromosomal segments located between 2057 and 2073 kb (Kunst *et al.*, 1997) would account for SP β integration near *gltAB* as well as for deletions extending from SP β to the region between *odh* and *gltAB* (Zahler, 1982; Rowe *et al.*, 1986).

The overall G + C content of the prophage is 34.6 mol%, lower than that of the *B. subtilis* genome, which amounts to 43.5 mol% overall (Kunst *et al.*, 1997). In a 5 kb region of SP β c2 containing ORFs *yolFGHIJK* (Table 2), the G + C content is further reduced to 28 mol%. Such unusually low G + C content is also characteristic of *ggaAB*, the minor host teichoic acid operon supposed to be acquired by horizontal transfer (Freymond, 1995; Lazarevic *et al.*, 1995). Intergenic regions, representing 14% of the phage genome, have a G + C content of 32.8 mol%, a figure slightly lower than the 35 mol% seen in the coding regions.

Organization of ORFs and predicted functions of their products

Inspection of the SP β c2 nucleotide sequence revealed the presence of 187 putative ORFs starting with an NTG or an ATT codon, preceded at appropriate distance by a sequence complementary to the 3' end of 16S rRNA (3' UCUUCCUCCACUAG 5'). The usage of start codons (ATG, 65%; TTG, 24%; GTG, 6%), differs from that

Table 1. Similarity between SP β c2 and host regions

SP β c2 region (from-to; nt)	<i>B. subtilis</i> chromosome (from-to; nt)*	Identity (%)	Match length (nt)
7 590–7 849	1 879 954–1 880 213	90.0	260
8 359–10 331	2 073 482–2 071 508	92.5	1973
13 948–15 846	2 064 642–2 062 746	91.4	1899
38 560–39 862	555 931–557 228	97.0	1303
94 003–94 928	2 057 861–2 056 950	94.6	926
124 670–124 916	1 870 331–1 870 577	99.6	247

* Positions relative to the replication origin in assembled sequences with accession numbers from Z99104 to Z99124 (Kunst *et al.*, 1997).

calculated for the entire *B. subtilis* genome (ATG, 78 %; TTG, 13 %; GTG, 9 %) (Kunst *et al.*, 1997); in particular, the frequency of the standard ATG codon is notably reduced, while that of TTG is doubled. In addition, translation of several SP β c2 ORFs seems to be initiated by highly unusual start codons: ATT and CTG present six and three times, respectively (Table 2). These non-standard codons may reflect some specific regulation of phage gene expression at the translational level.

Based on the orientation of transcription, the prophage can be divided into three main clusters (Fig. 1). Contrary to the situation of clusters I and III – genes *yokA* to *yokK* and *yonR* to *yotN*, respectively – most ORFs forming cluster II, i.e. *bhlB* to *yonP*, are transcribed towards the bacterial origin of replication. In the excised and circularized phage configuration (not presented), genes of now contiguous clusters I and III have the same orientation. The mean length of a phage ORF is 622 nt, a figure appreciably lower than that of 890 nt calculated for the entire *B. subtilis* genome (Kunst *et al.*, 1997).

Cluster I comprises 23 ORFs. A putative site-specific recombinase encoded by *yokA*, located next to *attL*, may be involved in the insertion of the circularized SP β genome into the *B. subtilis* chromosome. The presence of the integrase gene in the vicinity of *att* has been reported for several bacteriophages (Stanley *et al.*, 1997). Phage integration into the chromosome of a nonlysogenic host results in the duplication of the attachment site (Fig. 2). The SP β attachment site (*attP*) is flanked on each side by an identical 16 base sequence but with reversed order. The role of this inverted repeat is not known, although its implication in phage excision seems likely. The plasmid pPS373, which has a 3.8 kb insert (Fig. 1) that contains both *attP* and ORF *yokA* encoding the putative site-specific recombinase, suffered rearrangements in the *E. coli* host DH5 α , possibly due to interplasmid recombination. Indeed, sequencing of each DNA strand of pPS373 revealed superimposed sequences beyond the middle part of *attP*. Such rearrangements were not observed when the sequencing template was pPS558 (Fig. 1), a plasmid containing *attP* and only part of *yokA*.

The presence within this cluster of *yokD*, an ORF encoding an aminoglycoside *N*^{3'}-acetyltransferase, is surprising, since this typically bacterial enzyme, conferring resistance to antibiotics, has no homologues in phage genomes sequenced so far. Neither is there an obvious role in the phage life cycle for YokG, a protein similar to the *Bacillus thuringiensis* delta-endotoxin.

Three putative proteins encoded by cluster I are endowed with a prokaryotic lipoprotein signal signature: YokB, which has an unpredicted function; a putative endonuclease, YokF; and the putative bacteriocin YolG. An SP β lysogen known to secrete betacin, a bacteriocin-like protein that is able to kill nonlysogenic bacteria, is rendered resistant to this protein by the product of the tolerance gene *tol* (Hemphill *et al.*, 1980). The latter most likely corresponds to *yolH*, an ORF following *yolG* and encoding an ABC-transporter similar to protein exporters. Sequence homologies, as well as this gene organization, strongly suggest the involvement of YolH in bacteriocin export. YokH, YokI and YolE exhibit 95, 91 and 94 % identity, respectively, to their host paralogues, to which no function has so far been assigned.

Cluster II may correspond to the late phage operon. At least three proteins encoded by this cluster may participate in host cell lysis: the *N*-acetylmuramoyl-L-alanine amidase BlyA (Regamey & Karamata, 1998), the holin BhlB (Regamey & Karamata, 1998) and a 252 kDa putative transglycosylase YomI.

Homology between SP β c2 YomR and PBSX XkdV (Krogh *et al.*, 1996), the putative long tail-fibre protein (Junier, 1996), suggests that YomR is likely to fulfil the same function. This conclusion is supported by the presence in YomR of different repeated motifs (Fig. 3), a feature common to several Gram-negative bacteriophage long-fibre proteins (Sandmeier, 1994). The short ORFs *yomQ* and *yomP* that follow *yomR*, and their homologues in PBSX and *skin* elements, may code for the small tail-fibre proteins. Curiously, compared to its PBSX and *skin* counterparts, *yomS* seems to be truncated and located at the opposite extremity of the conserved

Table 2. Features of SP β c2 ORFs and their deduced products

ORF	Start/stop codon	Position*	G + C (mol %)	Deduced product		Similar protein					
				Size (aa)	Mass (kDa)	ORF name	(Possible) function	Organism	Match length (aa)	Identity (%)	Reference/ accession no.
Cluster I											
yokA	TTG/TAA	41 < 1678	34.1	545	63.14	cisA	DNA recombinase	B. subtilis	448	27	Sato et al. (1990)
yokB	ATG/TAA	1881 > 2591	34.3	236	27.34		Lipoprotein				
yokC	ATG/TAA	2798 > 3313	32.7	171	20.52						
yokD	ATG/TAG	3964 < 4782	36.0	272	30.91	aacC9	Aminoglycoside N ^{3'} - acetyltransferase IX	Micromonospora chalybeata	262	33	Salauze et al. (1991)
yokE	ATG/TAA	5086 < 5568	32.0	160	18.25						
yokF	TTG/TGA	5582 < 6472	38.9	296	33.00	yncB	Lipoprotein	B. subtilis	213	74	Kunst et al. (1997)
						nuc	Thermonuclease (DNase/RNase)	Staphylococcus aureus	193	43	Shortle (1983)
yokG	ATG/TAA	6774 > 7847	33.4	357	40.74	cryVIB	Delta-endotoxin	Bacillus thuringiensis	312	35	L07024
						ynzF		B. subtilis	87	83	Kunst et al. (1997)
yokH	TTG/TAG	8371 > 8928	34.7	185	21.53	yobM		B. subtilis	162	95	Kunst et al. (1997)
yokI	ATG/TAG	9028 > 10743	38.8	571	64.29	yobL		B. subtilis	460	91	Kunst et al. (1997)
yokJ	TTG/TAA	10752 > 11249	37.3	165	18.91						
yokK	ATG/TAA	11313 > 11891	34.3	192	22.05						
yokL	ATG/TAA	11927 > 12460	34.0	177	20.85	orf2		Bacteriophage ϕ 3T	112	100	U35638
yolA	ATG/TGA	13086 > 13553	37.6	155	17.23						
yolB	GTG/TAA	13559 > 13915	31.9	118	13.29	orf3†		Bacteriophage ϕ 3T	75	100	U35638
yolC	ATG/TAA	13958 < 14293	32.4	111	12.79	yozM		B. subtilis	111	82	Kunst et al. (1997)
yolD	ATG/TGA	14467 > 14799	33.3	110	13.00	yozL		B. subtilis	98	87	Kunst et al. (1997)
yolE	ATG/TAA	14792 > 16042	39.4	416	46.73	yobH	UmuC homologue	B. subtilis	201	94	Kunst et al. (1997)
						uvrA	UV resistance	Enterococcus faecalis	426	41	Ozawa et al. (1997)
yolF	ATG/TGA	16144 > 16461	26.1	105	12.25	ynzG		B. subtilis	68	32	Kunst et al. (1997)
yolG	ATG/TAA	16758 > 16928	42.1	56	5.98		Lipoprotein, bacteriocin precursor				
yolH	TTG/TGA	16986 > 19103	28.8	705	81.54		Bacteriocin exporter				
yolI	ATG/TAA	19100 > 19513	26.5	137	16.18	bdb	Protein thiol-disulfide oxidoreductase	Bacillus brevis	89	33	Ishihara et al. (1995)
yolJ	ATG/TGA	19513 > 20781	28.0	422	49.76						
yolK	ATG/TGA	20778 > 21224	27.7	148	17.14	yvgU		B. subtilis	139	34	Kunst et al. (1997)
						dsbB†	Disulfide bond formation protein	Vibrio alginolyticus	94	32	Nakamura et al. (1996)
Cluster II											
bhlB	ATG/TAG	21280 < 21546	31.4	88	10.12	xhlB	Holin	B. subtilis PBSX	83	59	Krogh et al. (1996)
bhlA	ATG/TAG	21557 < 21769	31.4	70	8.38	uviB	Bacteriocin secretion or immunity	Clostridium perfringens	64	42	Garnier & Cole (1988)
blyA	ATG/TAA	21857 < 22960	37.5	367	39.63		N-Acetylmuramoyl-L- alanine amidase				Regamey & Karamata (1998)
yomD	ATG/TAA	23188 < 24012	34.4	274	31.97						
yomE	ATG/TAA	24182 < 26116	37.0	644	71.10						
yomF	CTG/TAA	26153 < 26974	36.3	273	31.75						
yomG	TTG/TAA	26990 < 29617	36.6	875	98.88						
yomH	ATG/TAG	29629 < 30387	31.6	252	29.41						
yomI	TTG/TAA	30438 < 37295	37.3	2285	252.31	yjbJ		B. subtilis	154	62	Kunst et al. (1997)
yomJ	ATG/TAA	37349 < 38032	34.6	227	25.24		Immunity protein d				McLaughlin et al. (1986)
yomK	ATG/TAA	38114 < 38560	30.2	148	16.84						
yomL	ATG/TAA	38906 > 39082	27.6	58	6.51						
yomM	TTG/TAA	39107 > 39793	36.2	228	25.43	yddT		B. subtilis	228	98	Kunst et al. (1997)
yomN	GTG/TAA	39969 > 40298	36.0	109	12.36						
yomO	ATG/TAG	40301 < 41302	37.7	333	39.28	ripX	Integrase/recombinase	B. subtilis	286	22	U32685
yomP	GTG/TAA	41316 < 41735	37.8	139	15.37						
yomQ	ATG/TAA	41719 < 42219	31.9	166	19.20						
yomR	ATG/TAA	42269 < 42460	33.8	63	7.31	xkdX		B. subtilis PBSX	52	51	Krogh et al. (1996)
						yqcE		B. subtilis skin	52	51	Takemaru et al. (1995)
yomS	GTG/TGA	42457 < 42807	35.3	116	13.28	xkdW		B. subtilis PBSX	63	31	Krogh et al. (1996)
yomT	ATG/TAG	42818 < 44035	37.1	405	44.82	xkdV		B. subtilis PBSX	142	47	Krogh et al. (1996)
						yqcC		B. subtilis skin	120	30	Takemaru et al. (1995)
yomU	TTG/TAA	44037 < 44393	38.0	118	12.90	xkdY		B. subtilis PBSX	106	38	Krogh et al. (1996)
yomV	ATG/TAA	44457 < 44684	39.0	75	8.64	yqxG		B. subtilis skin	106	33	Takemaru et al. (1995)
yomW	ATT/TAA	44684 < 45295	37.4	203	21.71						
yomX	TTG/TAA	45313 < 46110	38.3	265	29.05						
yomY	ATG/TAA	46153 < 46863	34.5	236	27.78						
yomZ	ATG/TGA	46860 < 47366	35.1	168	19.33						

Table 2. (cont.)

ORF	Start/stop codon	Position*	G + C (mol %)	Deduced product		Similar protein					
				Size (aa)	Mass (kDa)	ORF name	(Possible) function	Organism	Match length (aa)	Identity (%)	Reference/ accession no.
<i>yomX</i>	ATG/TGA	47363 < 48013	35.6	216	24.18						
<i>yomY</i>	ATG/TAA	47997 < 48251	34.5	84	9.49						
<i>yomZ</i>	ATG/TGA	48248 < 48643	36.1	131	15.48						
<i>yona</i>	TTG/TAA	48658 < 49128	35.2	156	17.56						
<i>yonaB</i>	ATG/TAA	49164 < 50180	37.8	338	37.20						
<i>yonaC</i>	ATG/TAA	50219 < 50755	37.0	178	19.61						
<i>yonaD</i>	TTG/TAA	50780 < 52216	34.7	478	55.05						
<i>yonaE</i>	ATG/TAA	52247 < 53767	35.8	506	57.81						
<i>yonaF</i>	ATG/TAA	53785 < 55554	36.3	589	68.40						
<i>yonaG</i>	ATG/TAA	55541 < 56461	36.0	306	35.73						
<i>yonaH</i>	ATG/TAA	56564 < 57064	32.7	166	18.71						
<i>yoni</i>	ATG/TAA	57283 > 57693	31.1	136	15.66						
<i>yoniJ</i>	ATG/TAA	57727 < 58944	36.7	405	47.35						
<i>yoniK</i>	ATG/TAA	58961 < 59152	29.1	63	7.27						
<i>yoniN</i>	ATG/TAA	60834 < 61112	40.1	92	9.85	<i>hbs</i>	Histone-like protein II	<i>B. subtilis</i>	92	68	Micka <i>et al.</i> (1991)
<i>yoniO</i>	TTG/TAG	61356 < 63875	34.2	839	97.98						
<i>yoniP</i>	ATG/TAA	63915 < 64109	28.2	64	7.54						
Cluster III											
<i>yoniR</i>	ATG/TAG	65062 > 65388	35.1	108	12.37	<i>yqaE</i>	Transcriptional regulator	<i>B. subtilis</i> skin	110	50	Takemaru <i>et al.</i> (1995)
						<i>xre</i>	Repressor	<i>B. subtilis</i> PBSX	105	40	
<i>yoniS</i>	ATG/TAA	65503 > 66114	36.1	203	22.25		Lipoprotein				
<i>yoniT</i>	GTG/TGA	66489 > 66665	41.8	58	6.91						
<i>yoniU</i>	TTG/TAA	66980 > 67168	33.8	62	7.41						
<i>yoniV</i>	ATG/TAA	67250 > 68482	31.3	410	48.71						
<i>yoniX</i>	ATG/TAG	68810 > 69316	35.3	168	19.64						
<i>yopA</i>	TTG/TAA	69673 > 70989	32.4	438	50.87						
<i>yopB</i>	ATG/TGA	71250 > 71477	36.4	75	9.10						
<i>yopC</i>	ATG/TAA	71585 > 72913	31.5	442	52.15						
<i>yopD</i>	TTG/TAA	72971 > 73366	35.3	131	15.44	<i>yogG</i>		<i>B. subtilis</i>	125	45	Kunst <i>et al.</i> (1997)
<i>yopE</i>	ATG/TGA	73953 > 74204	32.1	83	9.54	<i>yndK</i>		<i>B. subtilis</i>	67	29	Kunst <i>et al.</i> (1997)
<i>yopF</i>	ATG/TAG	74208 > 74423	34.2	71	8.39						
<i>yopG</i>	GTG/TAA	74434 > 74565	40.1	43	5.48						
<i>yopH</i>	ATG/TAA	74604 > 75140	32.9	178	20.48						
<i>yopI</i>	TTG/TAA	75167 > 75700	30.1	177	20.49						
<i>yopJ</i>	ATG/TAG	75702 > 76118	36.2	138	16.02						
<i>yopK</i>	ATG/TAA	76295 > 77455	31.0	386	45.32						
<i>yopL</i>	ATG/TAA	77469 > 77594	40.4	41	4.21						
<i>yopM</i>	ATG/TAA	77921 > 78121	27.3	66	7.90						
<i>yopN</i>	ATG/TAA	78124 > 78441	28.6	105	12.13	<i>sigH</i>	Sigma factor H	<i>B. subtilis</i>	66	34	Kunst <i>et al.</i> (1997)
<i>yopO</i>	ATG/TAA	78489 > 78701	36.1	70	7.95	<i>sinR</i>		<i>B. subtilis</i>	64	28	Gaur <i>et al.</i> (1986)
<i>yopP</i>	ATG/TAA	78691 > 79767	33.9	358	41.60	<i>ripX</i>	Integrase/recombinase	<i>B. subtilis</i>	259	25	U32685
<i>yopQ</i>	ATG/TAA	79874 > 81256	32.3	460	53.50						
<i>yopR</i>	ATG/TAA	81280 > 82257	30.7	325	37.57						
<i>yopS</i>	ATG/TGA	82446 < 82670	35.1	74	8.58	<i>yotL</i>		<i>B. subtilis</i> SPβ	63	39	
						<i>yopO</i>		<i>B. subtilis</i> SPβ	60	38	
<i>yopT</i>	ATG/TAA	82853 > 83071	30.5	72	8.08						
<i>yopU</i>	TTG/TAA	83141 > 83338	30.3	65	7.66						
<i>yopV</i>	ATG/TGA	83450 > 83644	32.3	64	7.68						
<i>yopW</i>	TTG/TGA	83733 > 84068	32.1	111	12.55						
<i>yopX</i>	ATG/TGA	84065 > 84469	40.7	134	15.19						
<i>yopY</i>	ATG/TAA	84466 > 84744	30.8	92	10.94						
<i>yopZ</i>	GTG/TAA	84758 > 84961	40.1	67	7.96	gene 30		Bacteriophage SPP1	38	60	Alonso <i>et al.</i> (1997)
<i>yogA</i>	GTG/TGA	84974 > 85324	38.1	116	13.74						
<i>yogB</i>	ATG/TAA	85321 > 85659	35.6	112	13.43						
<i>yogC</i>	TTG/TGA	85666 > 86073	34.5	135	16.25						
<i>yogD</i>	ATG/TAG	86114 > 86869	35.7	251	28.59	<i>roi</i>	DNA-binding protein	Bacteriophage H-19B	234	32	Neely & Friedman (1998)
<i>yogE</i>	ATG/TAG	86923 > 87087	32.1	54	6.59						
<i>yogF</i>	ATG/TAA	87096 > 87299	34.8	67	8.20						
<i>yogG</i>	ATG/TAA	87344 > 87601	29.4	85	9.94						
<i>yogH</i>	ATG/TGA	87684 > 88136	38.6	150	16.55						
<i>yogI</i>	ATG/TAG	88185 > 88379	30.2	64	7.45						
<i>yogJ</i>	ATG/TGA	88590 > 89105	34.4	171	19.88						
<i>yogK</i>	CTG/TAG	89125 > 89328	32.8	67	8.38						
<i>yogL</i>	ATT/TAG	89368 > 90060	33.1	230	27.59	<i>orf4</i>		Bacteriophage φ31	223	31	Walker <i>et al.</i> (1998)
<i>yogM</i>	ATG/TGA	90185 < 90463	32.2	92	10.58						
<i>yogN</i>	TTG/TAA	90667 > 90885	38.3	72	8.38						
<i>yogO</i>	TTG/TAA	90902 < 91276	33.6	124	13.60						
<i>yogP</i>	ATG/TAA	91390 > 91731	35.9	113	13.18						
<i>yogR</i>	ATG/TGA	91691 > 92047	32.4	118	13.49						

Table 2. (cont.)

ORF	Start/stop codon	Position*	G + C (mol %)	Deduced product		Similar protein					
				Size (aa)	Mass (kDa)	ORF name	(Possible) function	Organism	Match length (aa)	Identity (%)	Reference/ accession no.
<i>yoqS</i>	ATG/TAA	92049 > 92396	33.9	115	13.46						
<i>yoqT</i>	ATG/TAA	92473 > 92622	32.0	49	6.09						
<i>yoqU</i>	GTG/TGA	92788 > 93201	38.8	137	16.43						
<i>yoqV</i>	TTG/TAA	93267 > 94079	37.8	270	31.03	AF1725		<i>Archaeoglobus fulgidus</i>	265	32	Klenk <i>et al.</i> (1997)
<i>yoqW</i>	ATG/TAA	94149 > 94823	40.5	224	25.95	<i>yobE</i>		<i>B. subtilis</i>	224	93	Kunst <i>et al.</i> (1997)
						<i>yoaM</i>		<i>B. subtilis</i>	223	86	Kunst <i>et al.</i> (1997)
<i>yoqX</i>	ATG/TAA	94894 > 95115	33.7	73	5.95						
<i>yoqY</i>	ATG/TAA	95170 > 95565	37.3	131	15.08						
<i>yoqZ</i>	TTG/TGA	95664 > 96488	37.8	274	30.36	<i>yndI</i>		<i>B. subtilis</i>	244	49	Kunst <i>et al.</i> (1997)
						<i>ymaC</i>		<i>B. subtilis</i>	201	50	Kunst <i>et al.</i> (1997)
						<i>yjgB</i>		<i>B. subtilis</i>	189	44	Kunst <i>et al.</i> (1997)
						<i>orf37.1</i>		Bacteriophage SPP1	193	40	X67865
<i>yorA</i>	ATG/TAA	96485 > 98245	36.0	586	64.97	gene 33		Bacteriophage SPP1	563	43	Alonso <i>et al.</i> (1997)
						<i>yclG</i>		<i>B. subtilis</i>	559	32	Kunst <i>et al.</i> (1997)
<i>yorB</i>	ATG/TAA	98334 > 98630	34.3	98	11.24						
<i>yorC</i>	ATG/TGA	98693 > 99073	33.5	126	15.02						
<i>yorD</i>	GTG/TAA	99150 > 99464	35.8	104	12.11		Cold shock and salt stress-induced				Graumann <i>et al.</i> (1996)
						<i>xkdW</i>		<i>B. subtilis</i> PBSX	86	33	Krogh <i>et al.</i> (1996)
<i>yorE</i>	TTG/TAA	99638 > 100009	32.5	123	14.75						
<i>yorF</i>	ATG/TAA	100031 > 100945	34.2	304	34.20						
<i>yorG</i>	ATG/TAA	101028 > 101999	36.3	323	36.63						
<i>yorH</i>	TTG/TAG	102042 > 102512	34.6	156	18.62						
<i>yorI</i>	TTG/TGA	102527 > 104041	33.8	504	58.13		DNA helicase				
<i>yorJ</i>	TTG/TAA	104057 > 105193	34.3	378	44.42						
<i>yorK</i>	ATG/TAA	105193 > 106923	34.1	576	65.77	<i>yrvE</i>		<i>B. subtilis</i>	550	27	Kunst <i>et al.</i> (1997)
						<i>recJ</i>	ss-DNA-specific exonuclease	<i>Chlamydia trachomatis</i>	421	28	Stephens <i>et al.</i> (1998)
<i>yorL</i>	TTG/TAA	106936 > 110853	35.4	1305	150.54	<i>dnaE</i>	DNA polymerase III α subunit	<i>Aquifex aeolicus</i>	829	30	Deckert <i>et al.</i> (1998)
<i>yorM</i>	ATT/TAA	110881 > 111597	32.6	238	27.09	<i>yocH</i>		<i>B. subtilis</i>	135	42	Kunst <i>et al.</i> (1997)
						<i>yabE</i>		<i>B. subtilis</i>	160	38	Ogasawara <i>et al.</i> (1994)
<i>yorN</i>	TTG/TAA	111713 > 111862	36.6	49	5.61						
<i>yorO</i>	TTG/TAA	111898 > 112095	34.3	65	7.67						
<i>yorP</i>	TTG/TAA	112128 > 112343	36.5	71	8.02			Mycobacteriophage L5	45	40	Harfull & Sarkis (1993)
<i>yorQ</i>	ATG/TAA	112336 > 112491	35.2	51	5.95						
<i>yorR</i>	ATG/TAG	112491 > 112988	35.7	165	19.06	<i>cdk</i>	Cytidylate kinase	<i>Sulfolobus acidocaldarius</i>	130	32	Y13049
<i>yorS</i>	ATG/TAA	112997 > 113515	31.9	172	20.42						
<i>yorT</i>	GTG/TAA	113547 > 113666	30.0	39	4.70						
<i>mtbP</i>	TTG/TAA	113715 > 115046	36.1	443	50.51		Modification methylase Bsu				Tran-Betcke <i>et al.</i> (1986)
<i>yorV</i>	TTG/TGA	115090 > 115308	35.1	72	8.46						
<i>yorW</i>	ATG/TAA	115311 > 115676	34.1	121	13.77						
<i>yorX</i>	TTG/TAG	115716 > 115943	35.9	75	8.50						
<i>yorY</i>	ATT/TAA	115956 > 116138	30.6	60	7.11						
<i>yorZ</i>	ATG/TAA	116205 > 116417	30.5	70	8.00						
<i>yosA</i>	ATG/TAA	116521 > 116640	44.1	39	4.09						
<i>yosB</i>	ATG/TAA	116771 > 116950	35.5	59	6.78						
<i>yosC</i>	ATG/TAA	116995 > 117537	35.3	180	21.46						
<i>yosD</i>	GTG/TAA	117576 > 117971	38.3	131	15.62						
<i>yosE</i>	TTG/TAA	117986 > 118333	31.8	115	12.89						
<i>yosF</i>	ATG/TAA	118347 > 118472	32.5	41	4.41						
<i>yosG</i>	ATG/TGA	118515 > 118877	33.6	120	13.73						
<i>yosH</i>	ATG/TAA	118938 > 119408	38.0	156	17.51						
<i>yosI</i>	ATG/TAA	119440 > 119574	33.3	44	5.01						
<i>yosJ</i>	ATG/TAG	119595 > 119789	38.4	64	7.82						
<i>yosK</i>	ATG/TGA	119834 > 120034	42.2	66	8.15						
<i>yosL</i>	ATG/TAG	120122 > 120475	35.8	117	13.14						
<i>bndI</i>	ATT/TAA	120481 > 120870	34.3	129	14.65	<i>nrdI</i> (<i>ymaA</i>)		<i>B. subtilis</i>	124	41	Scotti <i>et al.</i> (1996)
<i>bndE</i>	TTG/TAA	120833 > 121552					Ribonucleotide reductase, large subunit				Lazarevic <i>et al.</i> (1998)
<i>bndF</i>	TTG/TGA	121805 > 124339	36.2	1084	124.62		Ribonucleotide reductase, small subunit				Lazarevic <i>et al.</i> (1998)
		126058 > 126466	37.5	329	38.18						
<i>yosQ</i>	ATG/TAG	125361 > 125882	34.4	173	20.12		Homing endonuclease				Lazarevic <i>et al.</i> (1998)
<i>yosR</i>	ATG/TAA	126463 > 126705	37.0	80	9.09	<i>trx</i>	Thioredoxin	<i>B. subtilis</i>	76	40	Chen <i>et al.</i> (1989)
<i>yosS</i>	ATG/TAA	126751 > 127179	41.4	142	16.15	<i>yncF</i>	dUTPase homologue	<i>B. subtilis</i>	142	92	Kunst <i>et al.</i> (1997)

Table 2. (cont.)

ORF	Start/stop codon	Position*	G + C (mol %)	Deduced product		Similar protein					
				Size (aa)	Mass (kDa)	ORF name	(Possible) function	Organism	Match length (aa)	Identity (%)	Reference/ accession no.
<i>yosT</i>	ATG/TAA	127 274 > 127 723	33.5	149	17.02						
<i>yosU</i>	TTG/TAA	127 763 < 128 008	27.2	81	9.36						
<i>yosV</i>	TTG/TGA	128 328 > 128 618	31.6	96	10.96						
<i>yosW</i>	TTG/TGA	128 765 > 129 106	34.2	113	12.93						
<i>yosX</i>	TTG/TGA	129 337 > 129 690	31.3	117	14.01						
<i>yosZ</i>	ATG/TAA	129 734 > 129 799	40.9	21	2.23						
<i>sspC</i>	ATG/TAA	129 990 < 130 208	37.8	72	7.76		Small, acid-soluble spore protein				Connors & Setlow (1985)
<i>yotB</i>	ATG/TAG	130 327 > 131 154	35.9	275	32.25						
<i>yotC</i>	ATG/TAG	131 198 > 131 389	32.2	63	7.50						
<i>yotD</i>	ATG/TGA	131 429 > 131 560	31.0	43	5.01						
<i>yotE</i>	ATG/TGA	131 596 > 131 742	35.3	48	5.80						
<i>yotF</i>	ATG/TGA	131 774 > 131 851	26.9	25	3.02						
<i>yotG</i>	ATG/TGA	131 864 > 132 181	31.4	105	12.46						
<i>yotH</i>	CTG/TGA	132 197 > 132 370	32.1	57	6.41						
<i>yotI</i>	ATG/TAA	132 367 > 132 729	29.7	120	13.85						
<i>yotJ</i>	ATG/TAA	132 795 > 133 007	30.5	70	8.48	<i>xis</i>	Excisionase	Bacteriophage T12	60	30	McShan <i>et al.</i> (1997)
<i>yotK</i>	ATT/TAA	133 091 > 133 276	34.4	61	7.29						
<i>yotL</i>	GTG/TAA	133 278 < 133 520	32.9	80	9.29						
<i>yotM</i>	TTG/TAG	133 595 > 134 182	35.2	195	22.81						
<i>yotN</i>	TTG/TAA	134 185 > 134 361	32.2	58	7.07						

* The arrowheads indicate the direction of transcription.

† Incomplete sequence.

‡ Might also start with GTG at position 120 475.

3-ORF core (Fig. 4). Other peculiarities of this putative tail-fibre operon are (i) the absence of a transcriptional terminator downstream of its most distal ORF, *yomM*, which encodes a putative recombinase and (ii) the close proximity of the two convergently oriented ORFs *yomM* and *yozP*, whose stop codons are separated by only two nucleotides. Incidentally, the transcriptional terminator downstream of *yomL* is followed by a long imperfect palindrome (Fig. 4). Such palindromes were found in phages of Gram-negative bacteria, where they correspond to crossover sites for site-specific recombinases (DNA invertases), encoded by genes located next to the tail-fibre genes. Inversion of DNA segments of the tail-fibre gene cluster leads to phage host-range variation (Sandmeier, 1994). However, so far there is no experimental evidence supporting this putative DNA-inversion system in SPβ. The middle part of the palindrome downstream of the *yomL* terminator matches the palindrome-like sequence located at the very extremity of the 1.3 kb *dal*-linked host region that is 97% identical to the corresponding SPβc2 segment (Fig. 4; Table 2). The presence of these palindromic sequences and a possible involvement of a site-specific recombinase may account for chromosomal rearrangements characteristic of certain *B. subtilis* merodiploid strains (Anagnostopoulos, 1990; A. Regamey and others, unpublished).

The previously sequenced gene *d*, identified in an attempt to clone the SPβ repressor gene (McLaughlin *et al.*, 1986), corresponds to *yomJ*. The presence of this gene on a multicopy plasmid protected non-lysogens from both superinfection and lysogenization with SPβ

(McLaughlin *et al.*, 1986). The product of *yomN* belongs to the histone-like family of prokaryotic DNA-binding proteins and is highly similar to the essential host histone-like protein HBsu (Micka & Marahiel, 1992). By analogy with the histone-like protein of the African swine fever virus (Borca *et al.*, 1996), YonN may be located in the virion nucleoid.

The third and largest cluster, which is likely to encompass the early phage operon(s), begins with the putative phage repressor gene *yomR*, whose product is 40 and 50% identical to the PBSX repressor Xre (Krogh *et al.*, 1996) and to the hypothetical transcriptional regulator YqaE of the *skin* element (Takemaru *et al.*, 1995), respectively. This cluster encodes a possible antirepressor, YoqD, and the transcriptional regulator YopN. Since over half of the 123 predicted ORFs in this region specify proteins of less than 120 residues, the mean ORF length is only 164 codons. Large numbers of short ORFs have also been identified in the early operon region of the lytic bacteriophage sk1 of *Lactococcus lactis*, another Gram-positive bacterium (Chandry *et al.*, 1997).

A number of ORFs from cluster III specify proteins involved in the synthesis of DNA precursors and in phage DNA replication. The 19-gene operon extending from *yorH* to *yorZ* specifies, among others, the putative DNA helicase YorI, the ssDNA-specific exonuclease YorK, the alpha chain of the DNA polymerase YorL, and the previously characterized modification methylase MtbP that specifically methylates the central cytosine in sequences 5' GGCC 3' and 5' GCNGC 3' (Trautner *et*

Fig. 2. Duplication of *att* following integration of circular phage DNA. Left- and right-facing arrows indicate inverted repeats. Phage (*attP*), host (*attB*) and prophage (*attL* and *attR*) attachment sites are indicated. Recombination events were reconstructed from sequences obtained from phage DNA (this work), from an SPβ-cured strain (accession number M81760), as well as from an SPβ lysogenic strain (accession numbers M81762 and AF012906).

Fig. 3. Repeated motifs in YomR. The three motifs are specified by hatched boxes. Residues conserved within a given motif are marked by an asterisk.

al., 1980; Tran-Betcke *et al.*, 1986). Inspection of 12.5 kb *EcoRI* fragments of SPβc2 containing *mtbP*, inserted into pPS232, and cloned in *E. coli* TP610, a strain used for plasmid propagation, provides further confirmation of the role of MtbP. Indeed, out of the four *PstI*-recognition sites present on the insert, two, which overlapped with the 5' GCNGC 3' sequence, were not cleaved by their cognate restriction enzyme, revealing that the *mtbP* gene was expressed and that its product was functional.

Sequence homologies of the products of the *yorE-Z* operon suggest their involvement in the synthesis of DNA precursors, which may be required for the replication of the rather large SPβ genome. Thus, the dUTPase that produces dUMP, the precursor of thymidine nucleotides, could be encoded by *yosS*. The putative large and small ribonucleotide reductase subunits are encoded by *bnrDE* and *bnrDF*, respectively (Lazarevic *et al.*, 1998). The former was shown to contain a group I intron and an intein coding sequence, while the latter contains the group I intron that harbours the putative

Fig. 1. Physical map of the integrated SPβc2 prophage. Solid lines represent regions cloned in plasmid vectors; dotted lines, PCR-amplified regions; ds, *EcoRI* segment sequenced using SPβc2 DNA as template. Restriction sites are also shown: A, *Avall*; B, *BamHI*; E, *EcoRI*; S, *SacI* (not all sites shown). Distances in kb from *attL* are indicated. Hatched bars indicate regions over 90% identical to different host DNA regions (see Table 1). T, putative stem-loops; <, putative P^A promoter; O, SPBRE sequence; ♦, SOS box; □, approximate location of the extremities of the linear phage genome deduced from the experimental restriction fragment map (Fink & Zahler, 1982); —, group I intron.

Fig. 4. Genetic organization of *yomP*–*yomS* genes and their homologues in PBSX and the *skin* element. The three-ORF core is boxed. Dotted horizontal lines correspond to domains of *yqgC* and *xepA*, which are absent from *yomS*, their *SPβ2* homologue. The *SPβ2* 1:3 kb segment, 97% homologous to the host region at 556 kb from the origin of replication (Kunst *et al.*, 1997), is indicated by a hatched bar. Identical nucleotide residues of one of its extremities (*SPβ*) with those of the homologous host region (*Bsu*) are aligned (|). The position of the loss of homology is indicated by a triangle. Converging arrowheads indicate palindromes.

Analysis of the promoter regions

Searches for putative regulatory signals in the non-coding sequences revealed a novel 31 nt motif, designated SPBRE (SPbeta repeated element), that may represent the recognition site for the SP β repressor. Allowing up to six mismatches to the consensus sequence, nine SPBRE copies were identified in intergenic regions of the SP β c2 genome (Fig. 5), while none were

found in the host genome. An additional copy was found within the SP β c2 ORF *yorM*, 4 nt downstream of the putative start codon. The SPBRE sequences are restricted to cluster III and the regulatory region of *yonP*, the proximal gene of cluster II. In five cases (Fig. 5), SPBRE sequences exhibit 9/12 to 11/12 adherence to the consensus of the promoters recognized by the major σ^A factor of *B. subtilis*. In other cases, with the exception of the *yonT* regulatory region, SPBRE is either preceded or followed by the P^A consensus-like sequence. Finally, the putative P^A promoter of *yonX* is inserted between two copies of SPBRE. Over half of the SPBREs are separated by 2–3 residues from the upstream consensus sequence AAAAAC.

The *yorA–yorB* and *yopS–yopT* intergenic regions each contain the GAAC-N₄-GTTC sequence, whereas the *yolC–yolD* spacer contains two such motifs (Fig. 5). This sequence, called the SOS box, was identified within the promoter regions of DNA damage-inducible genes of *B. subtilis* (Cheo *et al.*, 1991) and was shown to serve as an operator site for DinR, the repressor of the SOS regulon (Winterling *et al.*, 1997). Thus, the (auto) digestion of DinR as a result of DNA damage (Winterling *et al.*, 1997) may be at least partly responsible for SP β induction through derepression of three, or possibly more, operons, since the three palindromic SOS box sequences are located in the regulatory regions of divergons (Figs 1 and 5). Some products of the operon(s) activated by DNA damage, including the SP β operons mentioned above, may interfere with the SP β repressor activity, allowing full prophage induction. One of the three SP β c2 regulatory regions containing a SOS box is located within a segment that is 92.5% identical to the host region encoding the *B. subtilis* UmuC homologue YobH (Kunst *et al.*, 1997) (Fig. 1, Tables 1 and 2). Other SP β c2 SOS boxes are found in regulatory regions of cluster-II ORFs of unknown function.

The difference in the G + C content between the SP β c2 prophage and the entire *B. subtilis* genome suggests a relatively recent acquisition of the former. Why would SP β and related phages, occupying more than 3% of the chromosome, be retained in many *B. subtilis* isolates, since prophage-cured strains seem to grow and sporulate normally? The same question has been raised for defective prophage PBSX (Krogh *et al.*, 1996). A possible explanation is that some phage genes, normally expressed in the prophage state, are beneficial for survival in natural habitats. For instance, the previously sequenced and characterized gene *sspC*, encoding a small acid-soluble protein (Connors & Setlow, 1985), was, unexpectedly, localized within the prophage. It was shown to be transcribed from a promoter recognized by the forespore-specific σ^G factor (Nicholson *et al.*, 1989), and integrated in the host sporulation regulatory network. Its product, a DNA-binding protein, confers to spores a high resistance to UV light (Tovar-Rojo & Setlow, 1991). The SP β bacteriocin and bacteriocin-exporter cassette offer an example of genes that render the host cell more competitive. Therefore, the SP β genes are not restricted to the phage life cycle, but some of

them concern host functions. Other candidates for such genes are ORFs whose role could presently not be predicted from sequence homology. A significant proportion of these genes, notably of those in cluster II, may be involved in phage morphogenesis.

About 25% of SP β c2 ORF products exhibit significant homology to protein sequences (EMBL release 54), mainly to host paralogues. Although some of the latter are encoded by PBSX and the *skin* element, those exhibiting over 60% identity to their SP β counterparts are apparently not specified by prophage-like elements. This suggests that many SP β genes might have been picked up from its host(s), either by excision of the prophage-neighbouring regions from the host chromosome during induction of the prophage, which could have been integrated at different positions, or by the host and phage recombinase-mediated events.

ACKNOWLEDGEMENTS

This work was supported by Grant 96.0245 from the Office Fédéral de l'Éducation et de la Science (D. Karamata & C. Mauël) and Grant BIO4-CT96-0655 from the European Commission (A. Düsterhöft).

REFERENCES

- Alonso, J. C., Luder, G., Stiege, A. C., Chai, S., Weise, F. & Trautner, T. A. (1997). The complete nucleotide sequence and functional organization of *Bacillus subtilis* bacteriophage SPP1. *Gene* **204**, 201–212.
- Anagnostopoulos, C. (1990). Genetic rearrangements in *Bacillus subtilis*. In *The Bacterial Chromosome*, pp. 361–371. Edited by M. Riley & C. Drlica. Washington, DC: American Society for Microbiology.
- Borca, M. V., Irusta, P. M., Kutish, G. F., Carillo, C., Afonso, C. L., Burrage, A. T., Neilan, J. G. & Rock, D. L. (1996). A structural DNA-binding protein of African swine fever virus with similarity to bacterial histone-like proteins. *Arch Virol* **141**, 301–313.
- 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.
- Chandry, P. S., Moore, S. C., Boyce, J. D., Davidson, B. E. & Hillier, A. J. (1997). Analysis of the DNA sequence, gene expression, origin of replication and modular structure of the *Lactococcus lactis* lytic bacteriophage sk1. *Mol Microbiol* **26**, 49–64.
- Chen, N.-Y., Zhang, J.-J. & Paulus, H. (1989). Chromosomal location of the *Bacillus subtilis* aspartokinase II gene and nucleotide sequence of the adjacent genes homologous to *uurC* and *trx* of *Escherichia coli*. *J Gen Microbiol* **135**, 2931–2940.
- Cheo, D. L., Bayles, K. W. & Yasbin, R. E. (1991). Cloning and characterization of DNA damage-inducible promoter regions from *Bacillus subtilis*. *J Bacteriol* **173**, 1696–1703.
- Chung, C. T. & Miller, R. H. (1988). A rapid and convenient method for the preparation and storage of competent bacterial cells. *Nucleic Acids Res* **16**, 3580.
- Connors, M. J. & Setlow, P. (1985). Cloning of a small, acid-soluble spore protein gene from *Bacillus subtilis* and determination of its complete nucleotide sequence. *J Bacteriol* **161**, 333–339.

- Deckert, G., Warren, P. V., Gaasterland, T. & 12 other authors (1998). The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature* **392**, 353–358.
- Del Sal, G., Manfioletti, G. & Schneider, C. (1988). A one-tube plasmid DNA mini-preparation suitable for sequencing. *Nucleic Acids Res* **16**, 9878.
- Devereux, J., Haeberli, P. & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* **12**, 387–395.
- Fink, P. S. & Zahler, S. A. (1982). Restriction fragment maps of the genome of *Bacillus subtilis* bacteriophage SP β . *Gene* **19**, 235–238.
- Freymond, P.-P. (1995). *Génétique et biochimie des acides teichoïques secondaires de Bacillus subtilis 168 et W23*. PhD thesis, Lausanne University.
- Gardner, A., Odebralski, J., Zahler, S., Korman, R. Z. & Aronson, A. I. (1982). Glutamine synthetase subunit mixing and regulation in *Bacillus subtilis* partial diploids. *J Bacteriol* **149**, 378–380.
- Garnier, T. & Cole, S. T. (1988). Studies of UV-inducible promoters from *Clostridium perfringens* *in vivo* and *in vitro*. *Mol Microbiol* **2**, 607–614.
- Gaur, N. K., Dubnau, E. & Smith, I. (1986). Characterization of a cloned *Bacillus subtilis* gene that inhibits sporulation in multiple copies. *J Bacteriol* **168**, 860–869.
- 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.
- Graumann, P., Schroder, K., Schmid, R. & Marahiel, M. A. (1996). Cold shock stress-induced proteins in *Bacillus subtilis*. *J Bacteriol* **178**, 4611–4619.
- Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**, 557–580.
- Hatfull, G. F. & Sarkis, G. J. (1993). DNA sequence, structure and gene expression of mycobacteriophage L5: a phage system for mycobacterial genetics. *Mol Microbiol* **7**, 395–405.
- Hemphill, H. E., Gage, I., Zahler, S. A. & Korman, R. Z. (1980). Prophage-mediated production of a bacteriocinlike substance by SP β lysogens of *Bacillus subtilis*. *Can J Microbiol* **26**, 1328–1333.
- Ishihara, T., Tomita, H., Hasegawa, Y., Tsukagoshi, N., Yamagata, H. & Uda, S. (1995). Cloning and characterization of the gene for a protein thiol-disulfide oxidoreductase in *Bacillus brevis*. *J Bacteriol* **177**, 745–749.
- Junier, T. (1996). *Analyse génétique et comparaison des protéines liées au phénotype kill chez les phages défectifs de type PBSX de Bacillus subtilis*. Travail de diplôme, Lausanne University.
- Klenk, H. P., Clayton, R. A., Tomb, J. F. & 48 other authors. (1997). The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature* **390**, 364–370.
- Krogh, S., O'Reilly, M., Nolan, N. & Devine, K. M. (1996). The phage-like element PBSX and part of the *skin* element, which are resident at different locations on the *Bacillus subtilis* chromosome, are highly homologous. *Microbiology* **142**, 2031–2040.
- Kunkel, B., Losick, R. & Stragier, P. (1990). The *Bacillus subtilis* gene for the development transcription factor σ^K is generated by excision of a dispensable DNA element containing a sporulation recombinase gene. *Genes Dev* **4**, 525–535.
- Kunst, F., Ogasawara, N., Moszer, I. & 148 other authors. (1997). The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. *Nature* **390**, 249–256.
- Lazarevic, V., Mauël, C., Soldo, B., Freymond, P.-P., Margot, P. & Karamata, D. (1995). Sequence analysis of the 308° to 311° segment of the *Bacillus subtilis* 168 chromosome, a region devoted to cell wall metabolism, containing non-coding grey holes which reveal chromosomal rearrangements. *Microbiology* **141**, 329–335.
- Lazarevic, V., Soldo, B., Düsterhöft, A., Hilbert, H., Mauël, C. & Karamata, D. (1998). Introns and intein coding sequence in the ribonucleotide reductase genes of *Bacillus subtilis* temperate bacteriophage SP β . *Proc Natl Acad Sci USA* **95**, 1692–1697.
- McLaughlin, J. R., Wong, H. C., Ting, Y. E., Van Arsdell, J. N. & Chang, S. (1986). Control of lysogeny and immunity of *Bacillus subtilis* temperate bacteriophage SP β by its *d* gene. *J Bacteriol* **167**, 952–959.
- McShan, W. M., Tang, Y.-F. & Ferretti, J. J. (1997). Bacteriophage T12 of *Streptococcus pyogenes* integrates into the gene for a serine tRNA. *Mol Microbiol* **23**, 719–728.
- Micka, B. & Marahiel, M. A. (1992). The DNA-binding protein HBSu is essential for normal growth and development in *Bacillus subtilis*. *Biochimie* **74**, 641–650.
- Micka, B., Groch, N., Heinemann, U. & Marahiel, M. A. (1991). Molecular cloning, nucleotide sequence, and characterization of the *Bacillus subtilis* gene encoding the DNA-binding protein HBSu. *J Bacteriol* **173**, 3191–3198.
- Nakamura, T., Enomoto, H. & Unemoto, T. (1996). Cloning and sequencing of *nhaB* gene encoding an Na⁺/H⁺ antiporter from *Vibrio alginolyticus*. *Biochim Biophys Acta* **1275**, 157–160.
- Neely, M. N. & Friedman, D. I. (1998). Functional and genetic analysis of regulatory regions of coliphage H-19B: location of shiga-like toxin and lysis genes suggest a role for phage functions in toxin release. *Mol Microbiol* **28**, 1255–1267.
- Nicholson, W. L., Sun, D. X., Setlow, B. & Setlow, P. (1989). Promoter specificity of σ^G -containing RNA polymerase from sporulating cells of *Bacillus subtilis*: identification of a group of forespore-specific promoters. *J Bacteriol* **171**, 2708–2718.
- Ogasawara, N., Nakai, S. & Yoshikawa, H. (1994). Systematic sequencing of the 180 kilobase region of the *Bacillus subtilis* chromosome containing the replication origin. *DNA Res* **1**, 1–14.
- Okamoto, K., Mudd, J. A. & Marmur, J. (1968). Conversion of *Bacillus subtilis* DNA to phage DNA following mitomycin C induction. *J Mol Biol* **34**, 429–437.
- Ozawa, Y., Tanimoto, K., Fujimoto, S., Tomita, H. & Ike, Y. (1997). Cloning and genetic analysis of the UV resistance determinant (*uvr*) encoded on the *Enterococcus faecalis* pheromone-responsive conjugative plasmid pAD1. *J Bacteriol* **179**, 7468–7475.
- Regamey, A. & Karamata, D. (1998). The N-acetylmuramoyl-L-alanine amidase encoded by the *Bacillus subtilis* 168 prophage SP β . *Microbiology* **144**, 885–893.
- Rosenthal, R., Toye, P. A., Korman, R. Z. & Zahler, S. A. (1979). The prophage of SP β c2dcitK₁, a defective specialized transducing phage of *Bacillus subtilis*. *Genetics* **92**, 721–739.
- Rowe, D. B., Iismaa, T. P. & Wake, R. G. (1986). Nonrandom cosmid cloning and prophage SP β homology near the replication terminus of the *Bacillus subtilis* chromosome. *J Bacteriol* **167**, 379–382.
- Salauze, D., Perez-Gonzalez, J. A., Piepersberg, W. & Davies, J. (1991). Characterization of aminoglycoside acetyltransferase-encoding genes of neomycin-producing *Micromonospora chalicea* and *Streptomyces fradiae*. *Gene* **101**, 143–148.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Sandmeier, H. (1994). Acquisition and rearrangement of sequence motifs in the evolution of bacteriophage tail fibres. *Mol Microbiol* **12**, 343–350.

- Sato T., Samori Y. & Kobayashi, Y. (1990). The *cisA* cistron of *Bacillus subtilis* sporulation gene *spoIVC* encodes a protein homologous to a site-specific recombinase. *J Bacteriol* **172**, 1092–1098.
- Scotti, C., Valbuzzi, A., Perego, M., Galizzi, A. & Albertini, A. M. (1996). The *Bacillus subtilis* genes for ribonucleotide reductase are similar to the genes for the second class I NrdE/NrdF enzymes of *Enterobacteriaceae*. *Microbiology* **142**, 2995–3004.
- Seaman, E., Tarmy, E. & Marmur, J. (1964). Inducible phages of *Bacillus subtilis*. *Biochemistry* **3**, 607–613.
- Shortle, D. (1983). A genetic system for analysis of staphylococcal nuclease. *Gene* **22**, 181–189.
- Spancake, G. A., Daignault, S. D. & Hemphill, H. E. (1987). Genome homology and divergence in the SP β -related bacteriophages of *Bacillus subtilis*. *Can J Microbiol* **33**, 249–255.
- Stanley, E., Fitzgerald, G. F., Le Marrec, C., Fayard, B. & van Sinderen, D. (1997). Sequence analysis and characterization of ϕ O1205, a temperate bacteriophage infecting *Streptococcus thermophilus* CNRZ1205. *Microbiology* **143**, 3417–3429.
- Stephens, R. S., Kalman, S., Lammel, C. J. & 9 other authors. (1998). Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. *Science* **282**, 754–759.
- Stragier, P., Kunkel, B., Kroos, L. & Losick, R. (1989). Chromosomal rearrangement generating a composite gene for a developmental transcription factor. *Science* **243**, 507–512.
- Takemaru, K., Mizuno, M., Sato, T., Takeuchi, M. & Kobayashi, Y. (1995). Complete nucleotide sequence of a *skin* element excised by DNA rearrangement during sporulation in *Bacillus subtilis*. *Microbiology* **141**, 323–327.
- Tovar-Rojo, F. & Setlow, P. (1991). Effects of mutant small, acid-soluble spore proteins from *Bacillus subtilis* on DNA *in vivo* and *in vitro*. *J Bacteriol* **173**, 4827–4835.
- Tran-Betcke, A., Behrens, B., Noyer-Weidner, M. & Trautner, T. A. (1986). DNA methyltransferase genes of *Bacillus subtilis* phages: comparison of their nucleotide sequences. *Gene* **42**, 89–96.
- Trautner, T. A., Pawlek, B., Gunthert, U., Canosi, U., Jentsch, S. & Freund, M. (1980). Restriction and modification in *Bacillus subtilis*: identification of a gene in the temperate phage SP β coding for a BsuR specific modification methyltransferase. *Mol Gen Genet* **180**, 361–367.
- Walker, S. A., Dombroski, C. S. & Klaenhammer, T. R. (1998). Common elements regulating gene expression in temperate and lytic bacteriophages of *Lactococcus* species. *Appl Environ Microbiol* **64**, 1147–1152.
- Warner, F. D., Kitos, G. A., Romano, M. P. & Hemphill, H. E. (1977). Characterization of SP β : a temperate bacteriophage from *Bacillus subtilis* 168M. *Can J Microbiol* **23**, 45–51.
- Winterling, K. W., Levine, A. S., Yasbin, R. E. & Woodgate, R. (1997). Characterization of DinR, the *Bacillus subtilis* SOS repressor. *J Bacteriol* **179**, 1698–1703.
- Zahler, S. A. (1982). Specialized transduction in *Bacillus subtilis*. In *Molecular Biology of the Bacilli*, vol. 1, pp. 269–305. Edited by D. A. Dubnau. New York: Academic Press.
- Zahler, S. A. (1993). Temperate bacteriophages. In *Bacillus subtilis and Other Gram-positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, pp. 831–842. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology.
- Zahler, S. A., Korman, R. Z., Rosenthal, R. & Hemphill, H. E. (1977). *Bacillus subtilis* bacteriophage SP β : localization of the prophage attachment site, and specialized transduction. *J Bacteriol* **129**, 556–558.

Received 9 November 1998; revised 4 January 1999; accepted 6 January 1999.