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

Susanne Krogh,[1] Mary O'Reilly,[1] Niamh Nolan[1,2] and Kevin M. Devine[1]

Author for correspondence: Kevin M. Devine. Tel: +353 1 6081872. Fax: +353 1 6798558.
e-mail: kdevine@tcd.ie

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

PBSX and *skin* are two unusual genetic elements resident on the *Bacillus subtilis* chromosome. PBSX is a phage-like element located at approximately 100° which is induced by the SOS response and results in cell lysis with the release of phage-like particles. The phage particles contain bacterial chromosomal DNA and kill sensitive bacteria without injecting DNA. The *skin* element is located at approximately 230° on the chromosome and is positioned within the *sigK* open reading frame (ORF). It is excised at a particular stage of sporulation, leading to reconstitution of the complete *sigK* gene. In this paper, we show that there are phage-like operons present in the *skin* element which are highly homologous to the region of PBSX comprising part of the control region and the late operon. These operons are similar in terms of their gene organization, the percentage identity of the products of homologous ORFs and the positioning and strengths of ribosome-binding sites for each ORF. Although this high degree of conservation suggests that the phage-like operons in *skin* can be expressed, expression of the late operon was not detected during exponential growth, during sporulation or after induction of the SOS response. However two non-phage-like operons in the *skin* element are expressed and have distinct expression profiles that are dependent on the growth and developmental status of the cell.

Keywords: PBSX, *skin*, high level of similarity, expression of *skin* operons

INTRODUCTION

Many species of *Bacillus* have a prophage-like element resident on their chromosomes (Steensma *et al.*, 1978). Expression of this element is initiated by agents which induce the SOS response and results in cell lysis with the concomitant release of phage-like particles (Okamoto *et al.*, 1968; Steensma *et al.*, 1978). The particles can be distinguished serologically, morphologically by the number of cross-striations in the tail, and by their range of killing activity. Using these criteria, the phage-like elements PBSW, PBSX, PBSY and PBSZ have been identified in *Bacillus subtilis* var. *vulgatus* and in strains 168, S31 and W23, respectively. *Bacillus licheniformis* harbours an apparently related element PBSV. The phage-like particles can bind to and kill bacterial cells which carry

heterologous phage-like elements, whereas the host bacterium is resistant to their killing activity. For PBSX and PBSZ, the type of teichoic acid resident in the cell wall determines the adsorption spectrum of the phage and hence their killing selectivity (Glaser *et al.*, 1966; Karamata *et al.*, 1987; Young *et al.*, 1989). PBSX binds to erythritol teichoic acids of *B. subtilis* strain W23, but cannot bind to the glycerol teichoic acids present in the cell wall of the host *B. subtilis* strain 168. The reverse holds for the PBSZ phage-like element resident in *B. subtilis* strain W23. The mechanism whereby the bound phage kills the cell is not known, but it does not involve injection of DNA into the cell. An additional unusual feature of PBSX is that the head of the phage-like particle is capable of packaging only 13 kb of DNA, even though the PBSX genome exceeds 33 kb in size. These 13 kb fragments found in the phage-like particle head are predominantly *B. subtilis* chromosomal DNA (Okamoto *et al.*, 1968; Haas & Yoshikawa, 1969). From these

observations it is evident that PBSX induction is suicidal both for the phage and for the bacterial cell. Why then does the element persist in bacterial cells? All these elements encode proteins which can assemble into a phage-like particle, suggesting that it is the particle which holds the key to their function. One model proposes that they are phage-like bacteriocins. This does not exclude the possibility that they also encode genetic functions which are beneficial for cell physiology.

PBSX is the most thoroughly investigated of the phage-like elements resident in strains of B. subtilis. Mutations affecting induction (xin and xhi1479), the ability to form phage-like particles (xhd and xtl), and the killing activity (xki 1479) were isolated. These mutations all map between the metA and metC loci of the B. subtilis chromosome (Garro et al., 1970; Thurm & Garro, 1975a; Buxton, 1976). The proteins synthesized upon PBSX induction, and the structural proteins of the phage-like particle, range in size from 12 to 76 kDa (Thurm & Garro, 1975b; Mauel & Karamata, 1984). The size of the PBSX genome is not known precisely. However it has been established that part of the regulatory region and the entire late operon are approximately 33 kb in size (Wood et al., 1990a). The element is maintained on the chromosome by a repressor (xre) which binds to operator sites located in the vicinity of the promoters for divergent transcription units (Wood et al., 1990b; McDonnell & McConnell, 1994). One of these transcription units contains the gene pcf, which encodes a sigma-factor-like protein necessary for transcription of the late operon (McDonnell et al., 1994; McDonnell & McConnell, 1994). The structural proteins of the phage-like particle are encoded within the late operon. The last four cistrons of the late operon encode proteins which are involved in host cell lysis and include a holin-like protein and an autolysin (N-acetyl-muramoyl-L-alanine amidase; Foster, 1993; Longchamp et al., 1994; S. Krogh & K. Devine, unpublished).

The B. subtilis chromosome encodes a second unusual genetic element called skin (sigK intervening sequence), located at approximately 230° on the B. subtilis chromosome. During sporulation, a genetic rearrangement occurs in the mother cell compartment only, which results in the excision of the skin element from the chromosome. This is catalysed by a site-specific recombinase. Expression of the spoIVCA gene which encodes the recombinase (spoIVCA is located within the skin element) is activated at stage III of sporulation (Stragier et al., 1989; Kroos et al., 1989; Sato et al., 1990, 1994; Kunkel et al., 1990; Popham & Stragier, 1992). Excision of the skin element is necessary to reconstitute the sigK open reading frame (ORF). SigK then directs mother-cell-specific gene expression. The skin element has been completely sequenced (Takemaru et al., 1995). The products of ORFs 5, 7, 8 and 50 show homology to proteins encoded by phage ϕ105, suggesting that this element is a remnant of a cryptic ancestral phage (Takemaru et al., 1995). However, skin may be a composite element since it also contains genes encoding arsenate resistance which are homologous to those resident on the Staphylococcus aureus plasmid pI258 (Takemaru et al., 1995). Two findings suggest that

excision of skin is not critical for developmentally regulated expression of sigK: (i) strains engineered such that the skin element is excised and the sigK ORF is contiguous grow and sporulate normally (Kunkel et al., 1990); (ii) in closely related Bacillus strains, the sigK ORF is not interrupted (Adams et al., 1991). These data indicate that developmental regulation of sigK expression is effected by mechanisms which are independent of skin excision (Kunkel et al., 1988; Cutting et al., 1990; Lu et al., 1990). These observations raise questions about the origin of the skin element and the role it plays (if any) in cell physiology.

In this paper, we show that part of the skin element is highly homologous to the control and late operon regions of PBSX. We also investigate the expression of skin operons and show that two operons are expressed with distinct profiles which are dependent on the growth and developmental status of the cell. The phage-like late operon within skin, however, is not expressed during exponential growth, during sporulation, or during the SOS response.

METHODS

Bacterial strains and growth conditions. Escherichia coli strain TP611 (recBC hsdR M cya610 pcn) was used for cloning large chromosomal DNA fragments (Glaser et al., 1993). E. coli strain tg1 [K-12 Δ(lac pro) supE thi hsdR F' traD36 proAB lacI lacZ ΔM15] was used for subcloning and for preparation of sequencing templates. B. subtilis chromosomal DNA for sequencing was isolated from strain 168. Studies on skin and PBSX expression were performed in B. subtilis strain JH642 and in B. subtilis strain SKPOP9, a derivative of strain L8508 (xhi 1479 lyt 2) from which the distal part of the PBSX late operon (from fragment 37 to the distal part of fragment 39, which includes the lysis genes; Wood et al., 1990a) is deleted. B. subtilis and E. coli strains were routinely grown in Luria–Bertani (LB) medium (Miller, 1972). Sporulation in B. subtilis was initiated by resuspension in Sterlini–Mandelstam medium as described by Nicholson & Setlow (1990). The timing of sporulation was monitored by measurement of alkaline phosphatase and dipi-colinic acid accumulation. B. subtilis transformation was carried out according to the method of Anagnostopoulos & Spizizen (1961). E. coli transformation was carried out according to the method of Sambrook et al. (1989).

Plasmid and strain construction. Three oligonucleotide primer pairs were made (Applied Biosystems PCR-Mate DNA Synthesizer) and used to amplify segments of skin DNA by PCR (their location in the skin sequence is in parentheses): (1) 5′ GAAGGGTGCTACCATTAC 3′ (24423 → 24440) and 5′ ATCTCCTCACCTAGATC 3′ (25472 → 25456); (2) 5′ AGCCCTTCTCATTGAGAG 3′ (38023 → 38040) and 5′ GAAGCATCGACTAACACC 3′ (39018 → 39001); and (3) 5′ CTGCGTTGCTGTAGTAGC 3′ (43543 → 43560) and 5′ GTAATATCACTTGGCGG 3′ (44530 → 44514). The DNA amplified using each pair of primers was isolated, the ends were polished and the fragments were subcloned in pUC19 to give plasmids pSkin1F, pSkin1R, pSkin2R, pSkin3F and pSkin3R, where F and R represent the forward (the sequence in GenBank) and reverse orientations relative to sites within the polylinker. The integrity of the cloned fragments was checked by sequencing. Insert DNA with non-compatible ends was prepared from each plasmid and ligated directionally to the integrating plasmid pJM783 (a plasmid conferring resistance to chloram-

phenicol and containing a promoterless *lacZ* gene; a gift from J. A. Hoch) so that transcriptional fusions are generated upon integration of the recombinant plasmid into the chromosome of *B. subtilis*. The point of fusion in integrant strains SK1F and SK1R (the insertion is directed by the chromosomal inserts of pSkin1F and pSkin1R, respectively) is located within *yqbO* (ORF 54; Takemaru *et al.*, 1995); the point of fusion in integrant strain SK2R (directed by the chromosomal insert in pSkin2R) is located immediately 3′ to the *yqcF* (ORF 95; Takemaru *et al.*, 1995) cistron and the point of fusion in integrant strains SK3F and SK3R (directed by the chromosomal inserts in pSkin3F and pSkin3R, respectively) is positioned within the *yqcK* (ORF 2; Takemaru *et al.*, 1995) cistron. The location and structural integrity of DNA after integration at each locus was verified by Southern analysis using pSkin 1F as probe for strains SK1F and SK1R, pSkin 2R as probe for SK2R and pSkin3R as probe for SK3F and SK3R. Fragment 37 from PBSX was also used to generate a strain, SK37, which carried a *lacZ* fusion in the late operon of PBSX orientated in the direction of transcription of the late operon (Wood *et al.*, 1990a). The *lacZ* fusions resident in SK1F, SK2F, SK3F and SK37 were transferred into the *B. subtilis* strain SKPOP9 by chromosomal transformation, generating strains SKPOP1F, SKPOP2R, SKPOP3F and SKPOP37.

Isolation of *B. subtilis* chromosomal DNA and sequencing strategy. DNA was isolated from the chromosome of *B. subtilis* strain 168 by chromosomal walking. A fragment from the vicinity of the PBSX *xre* gene was cloned into the integrating vector pDIA5304 (Glaser *et al.*, 1993), giving plasmid pDIArep. This plasmid was then inserted into the chromosome of *B. subtilis* strain 168 (giving strain SKREP) with selection for resistance to chloramphenicol (3 μg ml⁻¹). Southern analysis was performed on restricted chromosomal DNA from integrant strain SKREP using pDIArep as a probe and it was determined that digestion with *Bam*HI releases a fragment of 12 kb in size adjacent to the insertion site. Chromosomal DNA from SKREP was restricted with *Bam*HI, ligated and transformed into *E. coli* strain TP611. Transformants containing the 12 kb chromosomal insert subcloned in plasmid pDIA5304 (pBam12) were identified. This chromosomal walking procedure was repeated, giving pSal14, which is contiguous with, and overlaps, pBam12. A 5·5 kb chromosomal fragment, pEco5, was isolated from a lambda clone of chromosomal DNA from *B. subtilis* strain SO113 (Wood *et al.*, 1990a). Parts of pEco5 were sequenced directly using oligonucleotides (S. Krogh, unpublished results). This sequence was used to design further oligonucleotides through which the sequence of this region was obtained from *B. subtilis* strain 168. Southern analysis was performed to verify the integrity of pBam12, pSal14 and pEco5. After sequencing, the actual insert sizes were: pBam12, 13 832 bp; pSal12, 11 483 bp and pEco5, 5497 bp. The overlap between pBam12 and pSal12 is 898 bp and that between pSal12 and pEco5 is 468 bp.

The isolated chromosomal DNA was sequenced using a shotgun strategy (pBam12 and pSal14) and by a directed approach using oligonucleotides (pEco5). Plasmid DNA (30 μg in 120 μl TE buffer) (TE buffer is 10 mM Tris/HCl; 1 mM EDTA) was randomly sheared either by sonication (Braun Labsonic: 7 pulses/0·22 cycles/0·25 W) or using DNaseI in the presence of manganese (Sambrook *et al.*, 1989). The sheared DNA was resolved on agarose gels and fragments of 2–5 kb in size were isolated. Fragment ends were polished using Klenow enzyme and T4 DNA polymerase, ligated to pUC18 and cloned into *E. coli* strain tg1. Template DNA for sequencing was prepared by two methods: (1) the miniprep procedure of Holmes & Quigley (1981) and (2) using magnetic beads. In the latter method, inserts were amplified using universal forward and reverse

primers, one of which was biotinylated. An aliquot of amplified DNA was then added to 2 μl of a streptavidin-conjugated magnetic bead suspension (Dynabeads; Dynall AS) and incubated for 10 min at 37 °C. The beads were then washed extensively, resuspended in TE buffer and used as templates for sequencing. Sequencing reactions were carried out using GENPAK fluorescent dye-primer sequencing kits according to the manufacturer's instructions. Reactions were resolved on an ABI Automated Sequencer model 373A. Gaps in the sequence were filled using custom synthesized oligonucleotides (Applied Biosystems PCR-Mate DNA Synthesizer) and the Applied Biosystems dye-terminator sequencing kit.

Computer sequence analysis. Sequence alignment and editing were performed using the xBAP program of the STADEN package. Conceptual translation of the sequence and other sequence analyses were performed using the NIP program of the STADEN package. The GenBank database was accessed using ACNUC (Gouy *et al.*, 1985) and homology searches of this database were performed using the TBLASTN program (Altschul *et al.*, 1990). Multiple sequence alignments were performed using CLUSTALV (Higgins *et al.*, 1992).

General molecular biological methods. Minipreps of plasmid DNA from *E. coli* were made according to the method of Holmes & Quigley (1981). Restriction analysis, nick translation, Southern analysis and other general molecular biological methods were performed as detailed in Sambrook *et al.* (1989).

Enzyme assays. Alkaline phosphatase and dipicolinic acid accumulation during sporulation were measured as described by Nicholson & Setlow (1990). Expression of *lacZ* was measured as described by Ferrari *et al.* (1986) with the activity being expressed in Miller units (Miller, 1972).

RESULTS AND DISCUSSION

Sequence comparison between PBSX and *skin*

The chromosomal region of *B. subtilis* strain 168 extending from *xre* (100°) to the vicinity of *xpa* (103°), encoding part of the control region and the entire late operon of the phage-like element PBSX, has been sequenced. This region of PBSX is approximately 28 kb in size. Two DNA fragments from within this region of *B. subtilis* strain SO113 were previously sequenced: a 5·5 kb segment comprising part of the PBSX control region (accession number Z34287, McDonnell *et al.*, 1994) and a 2·5 kb segment encoding the PBSX lysis genes (accession number L25924, Longchamp *et al.*, 1994; accession number Z36941, S. Krogh & K. Devine, unpublished). The PBSX sequence shows a very high level of similarity to part of the sequence of the *skin* element located at 230° on the *B. subtilis* chromosome. The similarity between the two elements is evident in their operon organizations and in the gene content of each operon (Fig. 1). There are three operons within this region of PBSX: the first (which for ease of discussion is called the early operon) is comprised of *xre* (the phage repressor) and *xkdA*; the second (called the middle operon) contains the genes *xkdB*, *xkdC*, *xkdD*, *xtrA* and *xpf* (formerly *pcf*, the sigma-factor-like transcription factor); and the third (called the late operon) includes the 25 genes from *xtmA* to *xlyA* which encode the head and tail proteins of the phage particle, the lysis proteins and perhaps other phage functions. There are three phage-like operons within the *skin* element, corresponding to each of the PBSX operons,
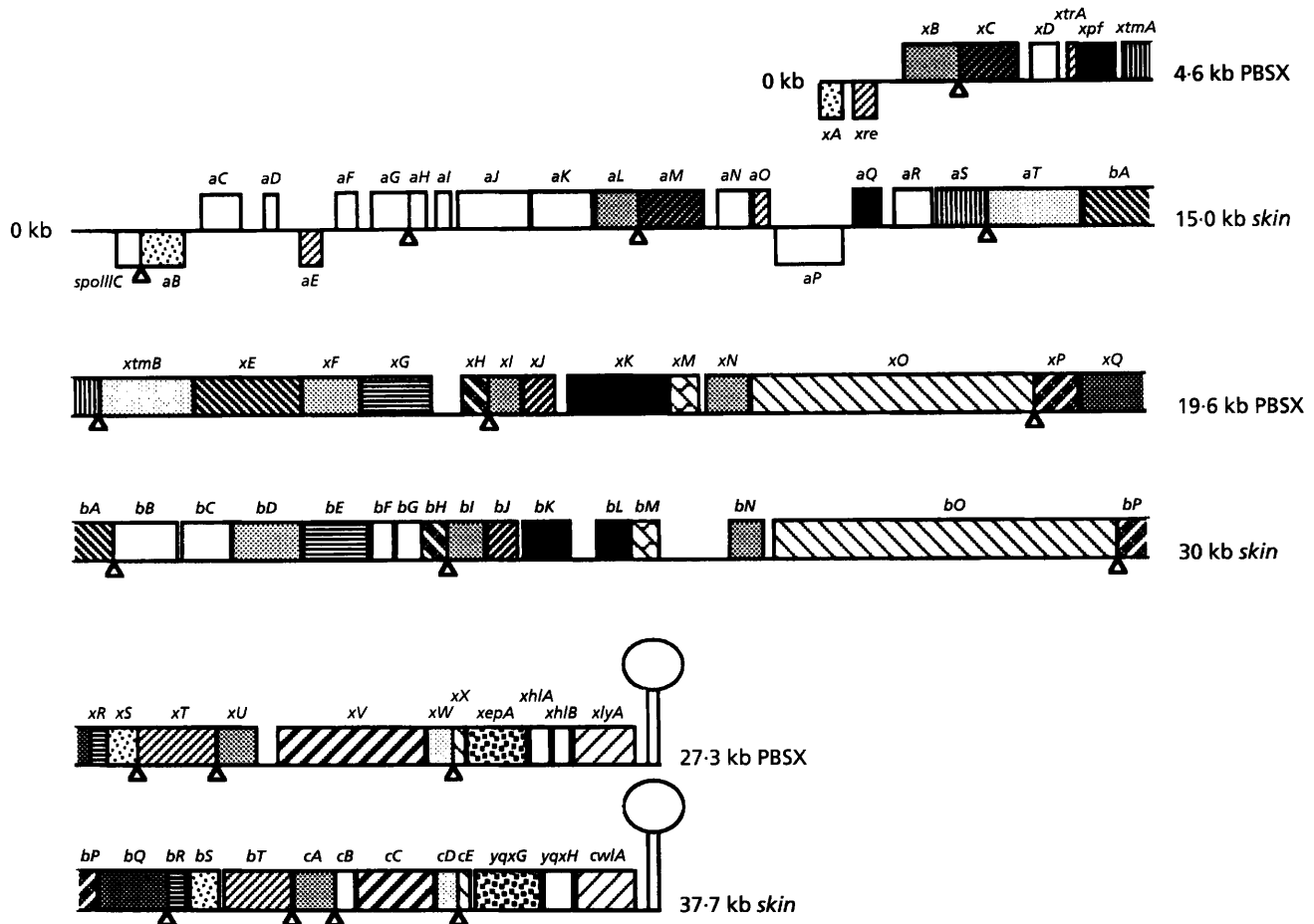
Fig. 1. A comparison of the ORFs identified in PBSX and in the *skin* element. Each ORF is indicated by a box with a distinctive shading. ORFs above the line are transcribed from left to right while those below the line are transcribed from right to left. Homologous ORFs in the two elements have the same shading. The name of each *skin* ORF is that assigned by Médigue *et al.* (1995), with the '*yq*' part of the name omitted for clarity of presentation. The full gene name and the corresponding number assigned to each gene by Takemaru *et al.* (1995) are shown in Table 1. Likewise, the PBSX genes are indicated with the '*kd*' part of the name omitted (the full gene name and the corresponding *skin* ORF are shown in Table 1). ORFs which do not have a homologue in the other element are shown as open boxes. Adjacent ORFs which overlap have an open triangle located at the junction between them. The stem–loop structures signify putative transcriptional terminators positioned after an autolysin gene in both elements.

and their organization is similar to that found in *skin* (Fig. 1). In both elements, the middle and late operons are orientated in the same direction while the early operon is adjacent to the middle operon but orientated in the opposite direction. There are also a number of differences in the operon organization of the elements: (i) there is a single ORF located between the middle and late operons of the *skin* element (and oriented in the opposite direction to them) which is not present in PBSX and (ii) *xkdA* and *xre* are adjacent in PBSX, whereas in the *skin* element there are two interposed ORFs and these are transcribed in the opposite direction.

In terms of gene content, the similarity between the elements is most evident in the late operon (Fig. 1). The homologous ORFs are similarly shaded and the number assigned to each ORF of the *skin* element by Takemaru *et al.* (1995) is indicated in parentheses after the name of the

gene (Médigue *et al.*, 1995) in Table 1. Twenty-three of the ORFs present in the late operon of PBSX have a homologue in the *skin* late operon. These genes are present in the same order in both operons and the positioning of start and stop codons for many adjacent ORFs is very similar. There are also a number of differences between the late operons: (i) there are no homologous ORFs in the late operon of PBSX for the *skin* genes *yqbB* (66), *yqbC* (84), *yqbF* (77), *yqbG* (76) and *yqcB* (130); (ii) there is a single ORF present in the lysis gene cluster of *skin* (*yqxH*, 121), whereas there are two smaller ORFs (*xhlA* and *xhlB*) present at the corresponding position in PBSX which do not show homology to *yqxH*; and (iii) there are two instances of rearrangements within ORFs in the late operons. The *xkdV* gene from PBSX contains a 963 nucleotide segment which is absent from the homologous ORF *yqcC* (131) in the *skin* element and the *yqbO* (54) from *skin* contains a 765 nucleotide segment

Table 1. Percentage identity and similarity between ORFs from PBSX and from *skin*

Operon	Gene*	*skin* homologue†		Identity (%)	Similarity (%)
Early	*xkdA*	*yqaB*	(45)	Gene is incomplete	
	xre	*yqaE*	(7)	35·3	66·4
Middle	*xkdB*	*yqaL*	(33)	19·5	44·0
	xkdC	*yqaM*	(34)	45·7	64·5
	xtrA	*yqaO*		36·7	73·5
	xpf	*yqaQ*	(41)	11·5	39·5
Late	*xtmA*	*yqaS*	(43)	16·6	46·8
	xtmB	*yqaT*	(50)	74·3	82·2
	xkdE	*yqbA*	(51)	67·2	84·9
	xkdF	*yqbD*	(52)	48·7	67·4
	xkdG	*yqbE*	(53)	79·7	93·9
	xkdH	*yqbH*		54·1	77·5
	xkdI	*yqbI*	(67)	49·7	74·2
	xkdJ	*yqbJ*	(68)	51·7	77·9
	xkdM	*yqbM*	(71)	95·2	98·0
	xkdN	*yqbN*	(55)	82·8	95·7
	xkdO	*yqbO‡*	(54)		
			1–1040 aa	20·7	39·9
			1040–1587 aa	77·3	91·8
	xkdP	*yqbP*	(85)	82·9	92·7
	xkdQ	*yqbQ*	(86)	81·6	94·5
	xkdR	*yqbR*	(87)	72·7	90·9
	xkdS	*yqbS*	(88)	85·8	98·6
	xkdT	*yqbT*	(89)	88·8	95·4
	xkdU	*yqcA*	(123)	86·4	94·2
	xkdV§	*yqcC*	(131)		
			1–97 aa	83·5	90·7
			417–687 aa	53·7	70·4
	xkdW	*yqcD*	(129)	63·9	75·7
	xkdX	*yqcE*	(128)	68·5	87·0
	xepA	*yqxG*	(122)	54·9	75·8
	xlyA	*cwlA*	(*cwlA*)	39·0	61·0

* Constituent genes of the PBSX early, middle and late operons.

† The nomenclature of the *skin* genes is taken from Médigue *et al.* (1995) and from Takemaru *et al.* (1995).

‡ Contains 255 aa which are absent in the homologous PBSX protein.

§ Contains 321 aa which are absent in the homologous *skin* protein.

which is absent from the homologous ORF in PBSX. The differences in gene content of the early and middle operons of the elements are more striking. There are seven contiguous ORFs in *skin* between *yqaE* (7) and *yqaL* (33) (the PBSX homologues are *xre* and *xkdB*, respectively) which are not present in PBSX. There are also three additional ORFs, *yqaN* (39), *yqaP* (40) and *yqaR* (42), in this region of *skin* which do not have homologues in PBSX. One of these, *yqaP* (40), is oriented in the opposite direction to the others within this region. In summary, it is clear that the phage-like operons of *skin* are similar, in both organization and gene content, to those of PBSX. However, significant rearrangements have occurred since the elements diverged, with the result that there are now 17 fewer ORFs within this region of PBSX than are found in the corresponding region of *skin*.

The percentage identity and similarity (at the amino acid level) between products of homologous ORFs from the *skin* and PBSX elements are presented in Table 1. The percentage identities between the gene products of the homologous ORFs of the late operons are very high and range (with two exceptions detailed below) from 48·7 % (67·4 % similarity) to 95·2 % (98 % similarity). The exceptions are the *xtmA* (16·6 % identity/46·8 % similarity) and *xlyA* (39 % identity/61 % similarity) gene products. The *xtmA* gene encodes the small subunit of the terminase (the large subunit being encoded by *xtmB*) and the *xlyA* gene encodes N-acetylmuramoyl-L-alanine amidase, an activity which functions in lysis of the bacterial cell upon PBSX induction. The identities between the products of the genes which show intragenic rearrangement are interesting: (i) in the case of *xkdV* and

Table 2. Positioning of ribosome-binding sites and their strengths for the ORFs located in *skin* and PBSX

Regions of complementarity with the 16S rRNA are underlined. ΔG^0 values are in kcal mol^{-1} (1 kcal = 4·2 kJ).

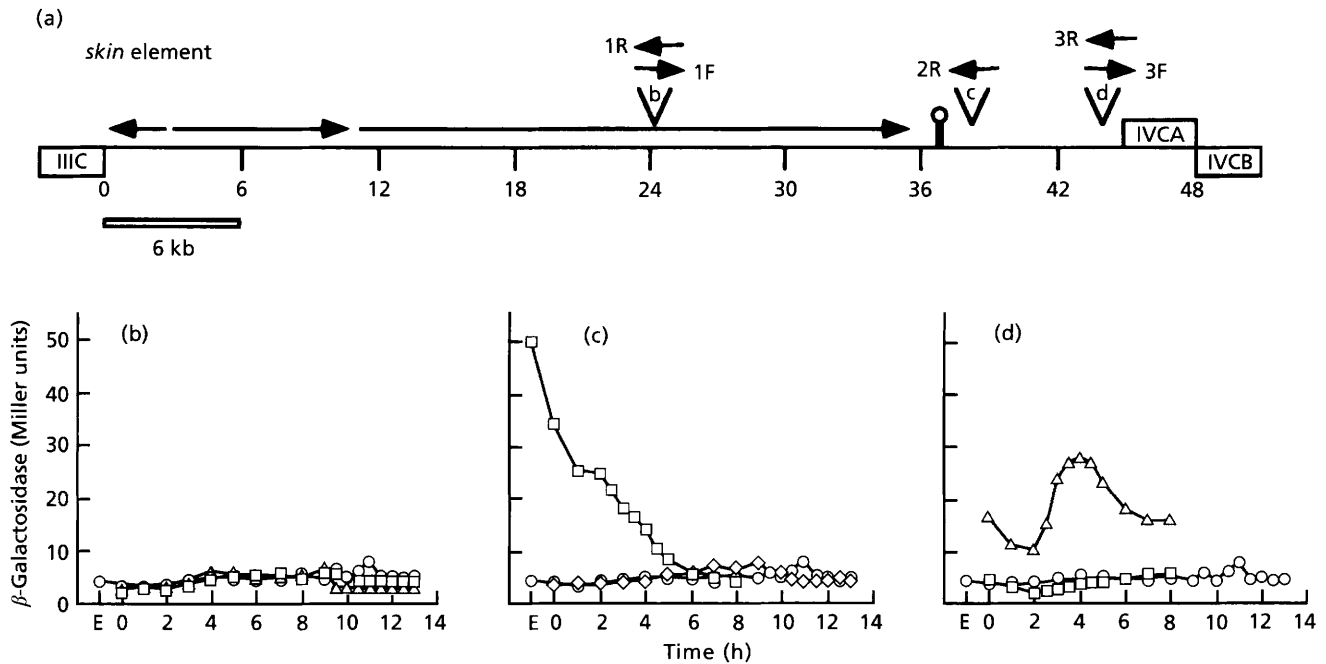| Gene	*skin*	ΔG^0	ΔG^0	PBSX	Gene
yqaB	ATTGTCGTCTTTAAGGAGGTTTTTCATC.TTG	−20·0	−11·8	GAACACACGTTCGAAAGGGAGTATTCAA.TTG	xkdA
yqaE	GAATAAAAACAAAAGTCAGGTGATTAAT.ATG	−15·6	−9·4	AGGGAGTTTAAAAAAGAGAGGGCATAGT.ATG	xre
yqaL	CATAACTGCACTCTGTGAAAGGAAGTAG.GTG	−14·2	−14·0	ATGATGTTAAAAGGAGAAGCGGCGTACA.ATG	xkdB
yqaM	TTCACTCAAAAGAGAAAGGAGAAACAAT.ATG	−18·4	−9·4	GCAATTCAGCATGAGAGGATGAAAAAAC.ATG	xkdC
yqaO	TTCTATCATAACATAGGGGGCGCTTTGA.GTG	−7·2	−14·2	CTGCACCGAAAAGGGGAGGAAATAAAAC.ATGA	xtrA
yqaQ	TGAATTCATTAAAGCATAAAATATCCAG.TTA	−4·6	−16·6	ATTTGCCAAAAATGAGGAGGATCATAGA.ATG	xpf
yqaS	ATCTAAAAGGAGGCGGCAGTGAATGTAG.ATG	−19·0	−18·8	ATGCTGGAGGTGGCGGTGATGCAGTAGC.ATG	xtmA
yqaT	ATCACCATTGTAAACAAAGGTGATGACA.GTG	−15·6	−8·4	TTATGATCAAACGAAAAGAGCGCAAGTC.ATGA	xtmB
yqbA	AAGATCAAGACCAGGAAGGAGGTAAAGC.ATG	−20·0	−23·4	GAACCGAGAAAGGAGGAGGTCATAAATC.ATG	xkdE
yqbD	CTTTTCTTATTTTAAGGAGGTGAATAAC.ATG	−24·4	−21·0	ACGTCCTGAAGGGAGGTGAATCAAGCAG.GTG	xkdF
yqbE	GTTGTTTTAATGGATAAGGAGGATACGA.ATG	−17·8	−18·6	TAAAACCAGTCAAGGAGGAGGAAATCAA.TTG	xkdG
yqbH	TTGAAGCGAAAGTGCGGATGAGATCAAT.ATG	−11·6	−15·2	AATCCTCCAGACTGAAGGTGAGATCTTT.ATG	xkdH
yqbI	TGACGGCAGTACGGGAGGTTGAATACTT.ATGA	−16·6	−9·4	AGGTCGTCGCAGTCAGGGATGAAAGCCT.ATGA	xkdI
yqbJ	TTATGAAGAATTTTAAGGCGGTGCTGCC.ATG	−15·8	−19·0	GCTGAAGAAGCTGTAAAGGAGGAGCAGG.ATG	xkdJ
yqbK	TCAGTTTCTTCAAAAGGAGGTCAAATAG.ATG	−21·2	−20·0	CATTTCTAACCAAGGAGGTCATGTCATC.ATG	xkdK
yqbM	TCTACTTTAATGTTGAGGTGAAATAGGA.ATG	−16·0	−16·6	TTCTACTTTAATGTGGAGGTAAACTAAT.ATG	xkdM
yqbN	AACTAAAAAGTAAAAGGAGTTAATAAAC.ATG	−14·0	−9·4	CAAACAATAGACAAGGGAGTTTTTTTAC.ATG	xkdN
yqbO	AAGAAGCGAAGAAGGGAGGTAATTAACT.ATG	−16·6	−16·6	GAATCAAGGAGAGGAGGGAACTAATATC.GTG	xkdO
yqbP	GTAGCTAAGATTAAGCAGGCCCTTGTCG.ATG	−8·6	−8·6	GCGGCGAAGATCAAGCAGGCGCTTATAG.ATG	xkdP
yqbQ	AAAAATACCGCAGTAAGCAGGTGATGAC.ATG	−17·2	−16·6	GATACCGCAATGAAACAGGTGATGTATG.ATG	xkdQ
yqbR	AAAATTCTATCCCTGATGGTGTTTCCTC.ATGA	−9·4	−12·8	AGAATACGATGCCGGAAGGAGTTTCCTG.ATG	xkdR
yqbS	TCTCGACAAAATATAGGAGGTGTTTGGA.ATG	−21·0	−14·0	TATTTAGCATGTAAAAGGAGTGGGCAAC.ATG	xkdS
yqbT	TTGGAGATCGAGGAGGTGCTTGGCGAGG.ATG	−21·0	−25·0	CTTGAGATTGAGGAGGTGATTGGCGAAG.ATG	xkdT
yqcA	ATCCCTAAATTGGGGCAGGTGAACATCA.TTG	−13·8	−13·8	ATTCCTAAGCTTGGGCAGGTGAAGATCA.TTG	xkdU
yqcC	TTAGGATTACTTTTGAGGTGGTTTAACA.ATG	−13·8	−14·6	TCAGATTTACTTTTGAGGTGGTGTAAAC.ATG	xkdV
yqcD	AAACTATCCAAGATTTTAGGAGGTAACG.ATG	−18·8	−18·8	ATTACCCGAGAGTTTAGGAGGTACCTTT.ATG	xkdW
yqcE	TGCTTACTCTACAAGGAGGGCAAGGTTC.ATGA	−17·8	−16·2	TACTTTCCTTGAAAGGAGATTATGCTGA.ATGA	xkdX
yqxG	TGATCTACAATAGCAAGGAGGATTTATT.ATG	−18·0	−19·0	CCTGTTTTTAGATCAAAGGAGGATGAAG.ATG	xepA
cwlA	ATCTTACTGACAGAAGGAGAGAGGGTAT.ATG	−12·8	−12·8	TAATCTGACGAAATAAGGAGAGATGAAA.ATG	xlyA

yqcC, the identities of the amino- and carboxy-terminal regions of their products are very high, but a deletion has occurred within the *skin* gene; (ii) in the case of *xkdO* and *yqbO*, the level of identity between the carboxyl halves of the proteins is high (77·3 % identity/91·8 % similarity), whereas the amino-terminal regions are quite dissimilar. The percentage identities between the products of homologous ORFs of the early and middle operons of the two elements are significantly lower than those observed for the products of genes located within the late operons, ranging between 11·5 % identity (39·9 % similarity) for the *xpf/yqaO* products to 45·7 % identity (64·5 % similarity) for the *xkdC/yqaM* products.

The observed similarities between these regions of PBSX and the phage-like operons of the *skin* element clearly show that they have shared a common ancestor. That this ancestor was probably a phage is suggested by the similarities of the two elements to phages in terms of their gene organization and by the presence within the elements of genes homologous to phage repressors, terminases and lysis genes (Wood et al., 1990a; McDonnell et al., 1994; Médigue et al., 1995; Takemaru et al., 1995). The highest level of divergence is seen in the early and middle operons, suggesting that the regulation of expression of these two elements may now differ. There are differences in the overall structure of the elements which may yield insight into their origins and possible contributions to cell physiology. The PBSX element terminates at the late operon terminator. The genes located downstream of the

PBSX terminator are different to those positioned downstream of the terminator in the *skin* element and they are not homologous to phage genes (S. Krogh, M. O'Reilly & K. Devine, unpublished). The genes located downstream of the terminator in *skin* include a homologue of *B. subtilis gsiA*, an arsenate-resistance operon and the recombinase *spoIVCA* (Takemaru et al., 1995) and do not appear to be of phage origin. These observations suggest that *skin* is a composite genetic element composed both of phage- and non-phage-derived DNA.

Investigation of the expression signals for the ORFs of PBSX and *skin*

The high level of identity between the homologous proteins from the PBSX and *skin* elements suggests that the phage-like operons within *skin* may be expressed and be functional. The late operon of PBSX is a single transcription unit (Wood et al., 1990a; McDonnell et al., 1994) and the similar organization of the *skin* late operon suggests that it may also be similarly expressed. Therefore the ribosome-binding site strength of each ORF will be an important determinant of the levels of head, tail and lysis protein production. The positioning and strength of the ribosome-binding sites for each of the ORFs of PBSX and *skin* are presented in Table 2. The strengths of the ribosome-binding sites were calculated based on complementarity with the 3'-end of the 16S ribosomal RNA (not allowing bulges) according to the rules of Tinoco et al. (1973). It is evident that the ORFs located within the late

(a)

skin element

1R 1F

3R 3F

2R c 3F

b

IIIC IVCA
IVCB

0 6 12 18 24 30 36 42 48

6 kb

(b)

β-Galactosidase (Miller units)

50
40
30
20
10
0

E 0 2 4 6 8 10 12 14

(c)

50
40
30
20
10
0

E 0 2 4 6 8 10 12 14

Time (h)

(d)

E 0 2 4 6 8 10 12 14

Fig. 2. Expression of *lacZ* fusions located within the *skin* element during vegetative growth and sporulation. (a) Three phage-like operons within the *skin* element which are homologous to the control and late operon regions of PBSX are indicated by arrows immediately above the scale bar. The length and direction of each arrow signify the size and transcriptional direction of the operon, respectively. The position and orientation of the *lacZ* fusions are shown by small arrows followed by the fusion-identifying code: insertions 1F and 1R are positioned within the *yqbO* (54) gene; insertion 2R is positioned immediately 3′ to the *yqcF* (95) gene; and insertions 3F and 3R are positioned within the *yqcK* (2) gene. The letters b, c and d at each insertion locus refer to the panel of this figure in which expression of *lacZ* at that locus is shown. The stem–loop structure indicates the putative transcriptional terminator present at the end of the phage-like late operon. The recombinase gene *spoIVCA* (IVCA) is transcribed from left to right and the two halves of the *sigK* gene (IVCB and IIIC) are transcribed from right to left (not drawn to scale). (b) Accumulation of β-galactosidase in strains SK1F (□—□), SK1R (△—△) and JH642 (○—○); (c) accumulation of β-galactosidase in strains SK2R (□—□), SK37 (fusion in PBSX, ◇—◇) and JH642 (○—○); (d) accumulation of β-galactosidase in strains SK3F (△—△), SK3R (□—□) and JH642 (○—○). E signifies exponential growth and t_0 represents the time at which sporulation was initiated by resuspension as described. In all the time courses, accumulation of alkaline phosphatase initiated at t_2 and accumulation of dipicolinic acid initiated between t_4 and t_5.

operons of the PBSX and *skin* elements have ribosome-binding sites which are appropriately positioned for translation. They show extensive complementarity with the 16S ribosomal RNA and this is reflected in ribosome-binding site strengths ranging from $\Delta G^0 = -8.6$ to -25 kcal mol^{-1} (-36.12 to -105 kJ mol^{-1}). There is a striking similarity in the ribosome-binding site strengths of homologous ORFs from the late operons of PBSX and *skin*. The ribosome-binding site strengths for 13 of the 23 pairs of homologous ORFs in the late operons only differ by approximately 1 kcal mol^{-1} (4.2 kJ mol^{-1}), while the ribosome-binding site strengths of five additional ORFs only differ by approximately 3 kcal mol^{-1} (12.6 kJ mol^{-1}). Correspondence analysis between the ribosome-binding site strengths of each ORF in the two elements gives a correspondence coefficient of 0.69, indicating a significant correlation between the two sets of data. This similarity contrasts with the early and middle operons, where the ribosome-binding site strengths are not significantly similar.

The positioning of the stop codon for the ORF preceding each ribosome-binding site is also indicated in Table 2

(e.g. the stop codon designated in bold characters within the ribosome-binding site of *xkdJ* is that of the preceding ORF, *xkdI*). For genes *xkdI*, *xkdJ*, *xkdO*, *xkdR*, *xkdW* and *xkdX*, the stop codon used and its positioning are conserved between the homologous ORFs of the two elements. In total, the stop codons for 11 PBSX ORFs and 13 *skin* ORFs are located within, or in close proximity to, the ribosome-binding site of the downstream ORF. These analyses support the hypothesis that the phage-like operons located within the *skin* element are expressed.

Are the phage-like operons of *skin* expressed?

To investigate the patterns of expression of operons within the *skin* element, three *lacZ* fusions were generated at the positions outlined in Fig. 2(a): (1) a fusion within the phage-like late operon at position b [the point of fusion is positioned within the amino-terminal region of *yqbO* (54)]; (2) a fusion at position c [the point of fusion is positioned immediately 3′ to the *yqcF* (95) gene]; and (3) a fusion at position d [the point of fusion is located within the *yqcK* (2) gene]. The orientation of each insertion is indicated in Fig. 2(a). The progression of sporulation was

established by measuring alkaline phosphatase accumulation (which begins to accumulate at stage II of sporulation, t_2 in our experiments) and dipicolinic acid (which begins to accumulate at stage V of sporulation, t_{4-5} in our experiments). The pattern and level of β-galactosidase accumulation for each fusion are shown: in Fig. 2(b), fusions 1F and 1R at position b; in Fig. 2(c), fusion 2R; and in Fig. 2(d), fusions 3F and 3R. The expression profile for the two fusions 1F and 1R positioned at b within the *skin* late operon is the same as that seen in the control strain JH642, which has no *lacZ* fusion, indicating that it is not expressed during vegetative growth nor during sporulation (Fig. 2b). The fusion 2R positioned at c shows a very high level of β-galactosidase during vegetative growth, followed by a steady decline in levels up to t_6 of sporulation (Fig. 2c). Expression of a *lacZ* fusion positioned within the late operon of PBSX was also examined during vegetative growth and sporulation and the profile observed was the same as the background (Fig. 2c). The insertion at position d in the 3F orientation was moderately expressed during vegetative growth, decreased slightly between t_0 and t_2 of sporulation and then increased to a peak at t_4 (Fig. 2d). The insertion in the reverse orientation at this position (3R) shows only background levels of β-galactosidase activity. These data show that there are two operons within the *skin* element which are transcriptionally active, and each has a distinctive expression profile. However, it is clear that neither the phage-like late operon of *skin* nor the late operon of PBSX are expressed during either vegetative growth or sporulation under these conditions. We then tested the hypothesis that expression of the phage-like *skin* operons may be under the control of the SOS response in a manner similar to PBSX. The strains harbouring the *lacZ* fusions 1F, 2R and 3F within the *skin* element were exposed to mitomycin C and the levels of accumulated β-galactosidase were determined. These *lacZ* fusions were also transferred to a strain in which the distal part of the PBSX late operon, including the lysis genes, is deleted. PBSX is induced normally in this strain but the cells do not lyse, which facilitates the measurement of the intracellularly located β-galactosidase. None of the *skin* fusions examined were expressed after exposure of the cells to mitomycin C (results not shown). The same results were obtained with the fusions in the lysis-negative genetic background. In order to verify that the SOS response was induced by treatment with mitomycin C, we examined the profile of β-galactosidase accumulation in strains with a *lacZ* fusion positioned within the PBSX late operon. In this case, intracellular β-galactosidase activity increased after exposure to the inducing agent and the cells subsequently lysed, confirming that the SOS response was induced (results not shown). These data demonstrate that expression of the three phage-like transcriptional units within the *skin* element is not under the control of the SOS response.

It is evident that the expression patterns for each of the *skin* operons are distinct and that only one of the three operons showed a pattern of expression which was related to sporulation. Thus the location of the *skin* element

within a sporulation gene may be fortuitous. There does not appear to be any obvious hierarchy or interdependency of expression among the operons, indicating that the phage and non-phage-like operons of *skin* are unrelated. The fact that the late operon is not expressed during exponential growth, sporulation or during the SOS response is consistent with the observations of Foster (1993). In his study, expression of the *cwlA* gene (the location of *cwlA* was unknown at that time) was investigated. The *lacZ* fusions in the Foster study were located in the *cwlA* gene, 12 kb downstream of the *lacZ* fusion point b in this study. However expression of neither insertion was observed under any of the conditions used in these studies. It is interesting to note, however, that Foster (1993) observed amidase activity when the structural gene for *cwlA* was inserted into an expression vector and cloned in *E. coli*. This demonstrates that at least one gene within the *skin* late operon is functional. The *cwlA/xlyA* homologous gene pair is only 39 % identical, the second lowest level of identity observed for the 23 homologous gene pairs of the late operon, suggesting that the remaining 22 genes may also be functional.

Is there a condition under which the phage-like operons within *skin* are expressed? The level of conservation which exists between the homologous ORFs in PBSX and *skin*, and the demonstration that *cwlA* encodes a functional enzyme (Foster, 1993), strongly suggest that the *skin* ORFs are functional and can be expressed. The sequence of genes which are not expressed over long periods of time will degenerate. This is the source of pseudo-genes in eukaryotes and there is evidence that a similar process has occurred in *B. subtilis* (Lazarevic *et al.*, 1995). Thus it is certain from our analysis that the constituent ORFs in the *skin* late operon must have been functional until very recently. It is possible that they are still functional and that a stimulus other than those tested in this study and in Foster's study triggers their expression. However, it is also possible that mutations, generated only very recently, have inactivated the induction mechanism (which may have responded to SOS- or sporulation-inducing signals) so that expression of the phage-like operons can no longer be triggered. These mutations could have occurred either during prolonged growth under laboratory conditions or during the process (mutagenesis by X-rays) whereby *B. subtilis* strain 168 was generated from the parental Marburg strain (Anagnostopoulos, 1990). If the ability to induce the *skin* phage-like operons does not provide a survival advantage during growth under laboratory conditions, then such mutations would not have been detected. To test these ideas, we are undertaking (i) to place the *skin* phage-like late operon under the control of the PBSX induction system to see if a functional phage particle can be formed and (ii) to examine expression of the *skin* late operon in the parental Marburg strain.

ACKNOWLEDGEMENTS

College, Dublin. We would like to thank Andrew Lloyd at the Irish National Centre for Bioinformatics for assistance with the use of their facilities and Des Higgins for ribosome-binding site strength calculations. We thank Maria Bradfield for technical assistance and David McConnell and Ken Wolfe for stimulating discussion and for critical reviews of the manuscript.

REFERENCES

Adams, L. F., Brown, K. L. & Whiteley, H. R. (1991). Molecular cloning and characterisation of two genes encoding sigma factors that direct transcription from a *Bacillus thuringiensis* crystal protein gene promoter. *J Bacteriol* **173**, 3846–3854.

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. (1990). Genetic rearrangements in *Bacillus subtilis*. In *The Bacterial Chromosome*, pp. 361–371. Edited by M. Riley & C. Drilica. Washington, DC: American Society for Microbiology.

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

Buxton, R. S. (1976). Prophage mutation causing heat inducibility of defective *Bacillus subtilis* bacteriophage PBSX. *J Virol* **20**, 22–28.

Cutting, S., Oke, V., Driks, A., Losick, R., Lu, S. & Kroos, L. (1990). A forespore checkpoint for mother cell gene expression during development in *Bacillus subtilis*. *Cell* **62**, 239–250.

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

Foster, S. J. (1993). Analysis of *Bacillus subtilis* 168 prophage-associated lytic enzymes; identification and characterization of CWLA-related prophage proteins. *J Gen Microbiol* **139**, 3177–3184.

Garro, A. J., Leffert, H. & Marmur, J. (1970). Genetic mapping of a defective bacteriophage on the chromosome of *Bacillus subtilis* 168. *J Virol* **6**, 340–343.

Glaser, L., Ionesco, H. & Schaeffer, P. (1966). Teichoic acid as components of a specific phage receptor in *Bacillus subtilis*. *Biochim Biophys Acta* **124**, 415–417.

Glaser, P., Kunst, F., Arnaud, M., Coudart, M.-P., Gonzales, W., Hullo, M.-F., Ionescu, M., Lubochinsky, B., Marcelino, L., Moszer, I., Presecan, E., Santana, M., Schneider, E., Schweizer, J., Vertes, A., Rapoport, G. & Danchin, A. (1993). *Bacillus subtilis* genome project: cloning and sequencing of the 97 kilobase region from 325° to 333°. *Mol Microbiol* **10**, 371–384.

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.

Haas, M. & Yoshikawa, H. (1969). Defective bacteriophage PBSH in *Bacillus subtilis*. I. Induction, purification and physical properties of the bacteriophage and its deoxyribonucleic acid. *J Virol* **3**, 233–247.

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

Holmes, D. S. & Quigley, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. *Anal Biochem* **114**, 193–197.

Karamata, D., Pooley, H. M. & Monod, M. (1987). Expression of heterologous genes for teichoic acid in *Bacillus subtilis* 168. *Mol Gen Genet* **207**, 73–81.

Kroos, L., Kunkel, B. & Losick, R. (1989). Switch protein alters

Kunkel, B., Sandman, K., Panzer, S., Youngman, P. & Losick, R. (1988). The promoter for a sporulation gene in the *spoIVC* locus of *Bacillus subtilis* and its use in studies of temporal and spatial control of gene expression. *J Bacteriol* **170**, 3513–3522.

Kunkel, B., Losick, R. & Stragier, P. (1990). The *Bacillus subtilis* gene for the developmental transcription factor σ^K is generated by excision of a dispensable DNA element containing a sporulation recombinase gene. *Genes Dev* **4**, 525–535.

Lazarevic, V., Mauel, 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.

Longchamp, P. F., Mauel, C. & Karamata, D. (1994). Lytic enzymes associated with defective prophages of *Bacillus subtilis*: sequencing and characterization of the region comprising the N-acetyl-muramoyl-L-alanine amidase gene of prophage PBSX. *Microbiology* **140**, 1855–1867.

Lu, S., Halberg, R. & Kroos, L. (1990). Processing of the mother-cell σ factor, σ^K, may depend on events occurring in the forespore during *Bacillus subtilis* development. *Proc Natl Acad Sci USA* **87**, 9722–9726.

Mauel, C. & Karamata, D. (1984). Characterisation of proteins induced by mitomycin C treatment of *Bacillus subtilis*. *J Virol* **49**, 806–812.

McDonnell, G. & McConnell, D. J. (1994). Overproduction, isolation and DNA-binding characteristics of Xre, the repressor protein from the *Bacillus subtilis* defective prophage PBSX. *J Bacteriol* **176**, 5831–5834.

McDonnell, G., Wood, H., Devine, K. M. & McConnell, D. J. (1994). Genetic control of bacterial suicide: the PBSX system of *Bacillus subtilis*. *J Bacteriol* **176**, 5820–5830.

Médigue, C., Moszer, I., Viari, A. & Danchin, A. (1995). Analysis of a *Bacillus subtilis* genome fragment using a co-operative computer system prototype. *Gene* **165**, GC37–GC51.

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

Nicholson, W. L. & Setlow, P. (1990). Sporulation, germination and outgrowth. In *Molecular Biological Methods for Bacillus*, pp. 391–450. Edited by C. R. Harwood & S. M. Cutting. Chichester: J. Wiley and Sons.

Okamoto, K., Mudd, J. A. & Marmur, J. (1968). Conversion of *Bacillus subtilis* DNA to phage DNA following mitomycin C. *J Mol Biol* **34**, 429–437.

Popham, D. L. & Stragier, P. (1992). Binding of the *Bacillus subtilis* *spoIVCA* product to the recombination sites of the element interrupting the σ^K-encoding gene. *Proc Natl Acad Sci USA* **89**, 5991–5995.

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

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.

Sato, T., Harada, K., Ohta, Y. & Kobayashi, Y. (1994). Expression of the *Bacillus subtilis spoIVCA* gene, which encodes a site-specific recombinase, depends on the *spoIIBG* product. *J Bacteriol* **176**, 935–937.

Steensma, H. Y., Robertson, L. A. & van Elsas, J. D. (1978). The

occurrence and taxonomic value of PBSX-like defective phages in the genus *Bacillus*. *Antonie Leeuwenhoek* **44**, 353–366.

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.

Thurm, P. & Garro, A. J. (1975a). Isolation and characterisation of prophage mutants of the defective *Bacillus subtilis* bacteriophage PBSX. *J Virol* **16**, 184–191.

Thurm, P. & Garro, A. J. (1975b). Bacteriophage-specific protein synthesis during induction of the defective *Bacillus subtilis* bacteriophage PBSX. *J Virol* **16**, 179–183.

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.

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

Wood, H. E., Devine, K. M. & McConnell, D. J. (1990b). Characterisation of a repressor gene (*xre*) and a temperature sensitive allele from the *Bacillus subtilis* prophage PBSX. *Gene* **96**, 83–88.

Young, M., Mauel, C., Margot, P. & Karamata, D. (1989). Pseudo-allelic relationship between non-homologous genes concerned with biosynthesis of polyglycerol phosphate and polyribitol phosphate teichoic acids in *Bacillus subtilis* strains 168 and W23. *Mol Microbiol* **3**, 1805–1812.