Microbiology (2002), **148**, 3971–3982 Printed in Great Britain

Proteomics characterization of novel spore proteins of *Bacillus subtilis*

Ritsuko Kuwana,[1] Yasuhiro Kasahara,[2] Machiko Fujibayashi,[1]
Hiromu Takamatsu,[1] Naotake Ogasawara[2] and Kazuhito Watabe[1]

Author for correspondence: Kazuhito Watabe. Tel: +81 72 866 3112 or 3114. Fax: +81 72 866 3112 or 3114.
e-mail: watabe@pharm.setsunan.ac.jp

[1] Faculty of Pharmaceutical Sciences, Setsunan University, Hirakata, Osaka 573-0101, Japan

[2] Nara Institute of Science and Technology, Ikoma, Nara 630-0101, Japan

The spores of *Bacillus subtilis* have characteristic properties and consist of complex structures including various types of proteins. To perform comprehensive analysis of the protein composition of the spores, the proteins extracted from the spore were analysed by a combination of one-dimensional PAGE and liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) using Turboquest SEQUEST software interfaced with the DNA sequence database of *B. subtilis*. A total of 154 proteins were identified, and 69 of them were novel. The remaining 85 proteins have been previously reported as sporulation-specific proteins or as proteins that are synthesized in vegetative cells. The expression pattern of each gene deduced to encode novel spore proteins was analysed using a series of strains carrying a *lacZ* reporter gene. The results revealed that the expression of 26 genes was dependent on sporulation-specific sigma factors, namely σ^F, σ^E, σ^G and σ^K. In this study, it is demonstrated that the combination of the techniques of SDS-PAGE and LC-MS/MS, with the mutant library of *B. subtilis*, is an effective tool for the analysis of complicated cellular structures.

Keyword: mass spectrometry

INTRODUCTION

Bacterial sporulation is a relatively simple model for cell differentiation and its progress is marked by sequential and drastic changes in the physiological state of the cell. When nutrients are exhausted, the Gram-positive soil micro-organism *Bacillus subtilis* initiates sporulation by dividing asymmetrically. After asymmetric septation, the resultant larger and smaller cells are termed the mother cell and the forespore, respectively. As development proceeds, the mother cell engulfs the forespore and eventually lyses, releasing the mature spores into the medium. The mature spores can remain in the dormant stage for long periods of time and are resistant to heat, toxic chemicals, lytic enzymes and other factors that cause cell damage (Aronson & Fitz-James, 1976; Goldman & Tipper, 1978; Gould, 1983; Driks, 1999, 2001; Henriques & Moran, 2000). Dormant spores respond rapidly to nutrients in the environment and return to an active physiological state through the process of germination (Paidhungat & Setlow, 2001). Genes involved in the spore developmental system of *B. subtilis* have been identified and their biological functions have been analysed (Stragier & Losick, 1996). These genes are mostly transcribed during sporulation by an RNA polymerase containing developmentally specific sigma factors, namely σ^F, σ^E, σ^G and σ^K. These factors, forming a sigma cascade, are temporally and spatially activated and regulate gene expression in a compartment-specific fashion (reviewed by Piggot & Losick, 2001). In the sigma cascade, σ^F is the primary sigma factor existing in the prespore and is required for the activation of pro-σ^E in the mother-cell compartment. Mature σ^E is required for the activation of σ^G in the forespore, and σ^G is required for the activation of pro-σ^K in the mother-cell compartment.

B. subtilis spores have three distinct structures, which can be observed using transmission electron microscopy (Aronson & Fitz-James, 1976). The central part of the spore, called the core, includes chromosomal DNA. Some proteins required for protection of nucleotides and for the processes of germination and outgrowth are thought to exist in the core. The cortex is a thin peptidoglycan layer outside of the core that includes some proteins involved in germination. The outermost

---

Abbreviation: LC-MS/MS, liquid chromatography coupled to tandem mass spectrometry.

Table 1. Bacterial strains and plasmids used in this study

We obtained the *B. subtilis* strains, used for the LacZ assay shown in Fig. 2, from the Japanese and European Consortia for Functional Analysis of the *B. subtilis* Genome (respectively, http://bacillus.genome.ad.jp/ and http://genolist.pasteur.fr/).

Strain/plasmid	Genotype/description	Source/reference
Strain		
B. subtilis 168	*trpC2*	1A1 (*Bacillus* Genetic Stock Center)
B. subtilis SGF602C	*trpC2 sigF::cat*	This work
B. subtilis SGE603C	*trpC2 sigE::cat*	This work
B. subtilis SGG604C	*trpC2 sigG::cat*	This work
B. subtilis SGK605C	*trpC2 spoIVCB(sigK)::cat*	This work
E. coli JM109	*relA supF44 endA1 hsdR17 gyrA96 mcrA* $mcrB^+$ Δ*thi* (*lac–proAB*)/F′ (*traD36* $proAB^+$ $lacI^q$ *lacZ*ΔM15)	Sambrook *et al.* (1989)
Plasmid		
pUC18	*bla*	Sambrook *et al.* (1989)
pDH88	*bla cat spac-1* promoter *lacI*	Henner (1990)
pCAT5	*cat*	This work
pSIGF62C	*cat sigF′*	This work
pSIGE63C	*cat sigE′*	This work
pSIGG64C	*cat sigG′*	This work
pSIGK65C	*cat spoIVCB(sigK)′*	This work

layer is called the spore coat and is composed of two layers, the inner coat and the outer coat. The spore coat is mainly composed of variously sized proteins, some of which function in cell development and morphogenesis; however, a limited number of them have been characterized because of the difficulty in their isolation or purification by conventional procedures (Driks, 2001). Therefore, the complete composition of spore-associated proteins and their characteristics are still unknown.

The *B. subtilis* genome-sequencing project revealed about 4100 protein-encoding genes, half of which have unknown functions (Kunst *et al.*, 1997). Systematic disruption of the remaining genes has already been carried out by the Japanese and European Consortia for Functional Analysis of the *B. subtilis* Genome (respectively, http://bacillus.genome.ad.jp/ and http://genolist.pasteur.fr/). The strategy for construction of the mutants was insertional inactivation of the target genes with pMutin carrying a *lacZ* reporter gene (Vagner *et al.*, 1998). Recently, almost all of the disruptant strains of *B. subtilis* have been made available for analysis.

The recent abundance of genome sequence data has brought an urgent need for the use of systematic proteomics to gain an understanding of the protein network that directs cellular functions, especially the sporulation, germination and spore dormancy of bacilli. Liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) is a powerful technique for the analysis of peptides and proteins and is now widely applied to proteome analysis. Multiprotein complexes function in the processes of bacterial sporulation and germination as well as most other cellular processes in many organisms. Here, we applied LC-MS/MS to a comprehensive analysis of proteins extractable from whole spores of *B. subtilis*, and 154 proteins were found. We also analysed the regulation of genes encoding each spore protein using cells with mutations in genes for sporulation-specific sigma factors.

METHODS

Bacterial strains, plasmids, media and general techniques. The *B. subtilis* and *Escherichia coli* strains used in this study are listed in Table 1. *B. subtilis* 168 (*trpC2*) (BGSC 1A1), obtained from the *Bacillus* Genetic Stock Center (Ohio State University, OH, USA), was used in this study as a wild-type strain. We also used a series of strains that have been constructed by the Japanese and European Consortia for Functional Analysis of the *B. subtilis* Genome (respectively, http://bacillus.genome.ad.jp/ and http://genolist.pasteur.fr/). The unified method for the construction of these strains has been described previously; in each strain, pMutin carrying a *lacZ* reporter gene is inserted into an appropriate position within the genome to inactivate a target gene (Vagner *et al.*, 1998). Oligonucleotide primers CATM172 (5′-TCTA-GTACTAAAGCACCCATTAGTTC-3′), CAT686R (5′-TT-AGATATCCGACTGTAAAAAGTACAGT-3′), PUC1697 (5′-AAAGATATCTAGGTGAAGATCCTTTTT-3′) and PUC723R (5′-TTAAGTACTCATGAGCGGATACATATT-3′) were used to amplify a 515 bp segment of a chloramphenicol-resistance gene (*cat*) of pDH88 (Henner, 1990) and a 974 bp segment of pUC18 (Sambrook *et al.*, 1989) including the origin of replication and the multiple cloning site. The PCR products were digested at the *Eco*RV and *Sca*I sites, and the resultant DNA fragments were ligated to create pCAT5. Oligonucleotide primers SIGF2428 (5′-TAAAAGC-TTCGGCAAAAACGCTCAGCT-3′) and SIGF207R (5′-TT-AGGATCCCGATGCAGCCGATCT-3′) were used to amplify a 183 bp segment internal to *sigF* on the *B. subtilis* 168

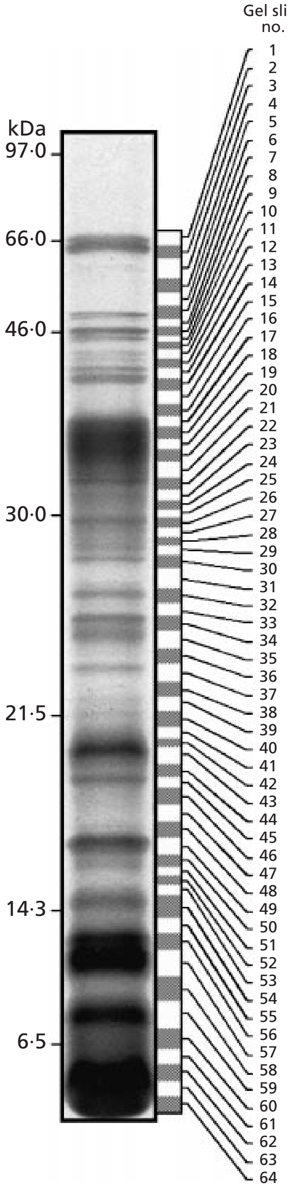

Fig. 1. SDS-PAGE was used to resolve the protein extracts from *B. subtilis* spores. Proteins were extracted from mature spores of *B. subtilis* in the presence of SDS and mercaptoethanol, as described in Methods. The protein preparation (approx. 20 μg) was resolved by SDS-PAGE (14 %, w/v, acrylamide gel) and the proteins were stained with Coomassie brilliant blue. The gel was then divided into 64 slices and analysed by LC-MS/MS.

chromosome. The PCR product was digested at the *Bam*HI and *Hin*dIII sites introduced by the primers and then inserted into *Bam*HI- and *Hin*dIII-restricted pCAT5 to create pSIGF62C. Oligonucleotide primers SIGE16 (5′-ACTAAGCTTACGGTTGACGCACCTC-3′) and SIGE208R (5′-TATGGATCCAGACGCAAAATGCGTTC-3′) were used to amplify a 192 bp segment internal to *sigE* on the *B. subtilis* 168 chromosome. The PCR product was digested at the *Bam*HI and *Hin*dIII sites introduced by the primers and then inserted into *Bam*HI- and *Hin*dIII-restricted pCAT5 to create pSIGE63C. Oligonucleotide primers SIGG121 (5′-GAAAAGCTTGTAAACGGGAACTT-3′) and SIGG315R (5′-GTAGGATCCTGCGGATCTCTCCGA-3′) were used to amplify a 194 bp segment internal to *sigG* on the *B. subtilis* 168 chromosome. The PCR product was digested at the *Bam*HI and *Hin*dIII sites introduced by the primers and then inserted into *Bam*HI- and *Hin*dIII-restricted pCAT5 to create pSIGG64C. Oligonucleotide primers SIGK16 (5′-AGGAAGCTTCGCAGCGCTCGGCTT-3′) and SIGK228R (5′-TCGCTCGAGAAGTTACCATTGTGGATTC-3′) were used to amplify a 212 bp segment internal to *spoIVCB* (*sigK*) from the *B. subtilis* 168 chromosome. The PCR product was digested at the *Bam*HI and *Hin*dIII sites introduced by the primers and then inserted into *Bam*HI- and *Hin*dIII-restricted pCAT5 to create pSIGK65C. Plasmids pSIGF62C, pSIGE63C, pSIGG64C and pSIGK65C were introduced into *B. subtilis* 168 under selection for chloramphenicol resistance (5 μg chloramphenicol ml^{-1}) to produce the transformants SGF602C, SGE603C, SGG604C and SGK605C, respectively. The plasmids were integrated into the chromosomal DNA by homologous recombination and the resultant transformants were tested by PCR. The chromosomal DNA of pMutin strains was introduced into SGF602C, SGE603C, SGG604C or SGK605C to construct double mutants of a sigma factor gene and each gene tested in this study. The resultant cells were resistant to both erythromycin and chloramphenicol. *B. subtilis* strains were grown in Difco Sporulation (DS) medium; *E. coli* was grown in Luria–Bertani (LB) medium (Takamatsu *et al.*, 2000). The conditions for sporulation of *B. subtilis* have been described previously (Takamatsu *et al.*, 2000). Recombinant DNA techniques were carried out following standard protocols (Sambrook *et al.*, 1989). Methods for preparing competent cells, for transformation and for the preparation of chromosomal DNA of *B. subtilis* have been described previously (Cutting & Vander Horn, 1990).

Preparation of spores. The wild-type strain of *B. subtilis* (168) was grown in DS medium at 37 °C as described previously, and mature spores were harvested 18 h after the cessation of exponential growth (T_{18}) and washed once with 10 mM sodium phosphate buffer (pH 7·2) (Takamatsu *et al.*, 2000). To remove the cell debris and vegetative cells, the pellets were suspended in 0·1 ml lysozyme buffer [10 mM sodium phosphate (pH 7·2), 1 % (w/v) lysozyme, Complete protease inhibitor cocktail (Roche)] and incubated at room temperature for 10 min. They were then washed repeatedly with buffer [10 mM sodium phosphate (pH 7·2), 0·5 M NaCl] at room temperature (Takamatsu *et al.*, 2000). After these treatments, more than 99·9 % of the spores were refractile and few dark spores were visible under the phase-contrast microscope. Neither vegetative cells nor cell debris were included in the sample (data not shown).

Solubilization of proteins from mature spores for SDS-PAGE. Spore proteins were solubilized in 0·1 ml loading buffer [62·5 mM Tris/HCl (pH 6·8), 4 % (w/v) SDS, 10 % (v/v) 2-mercaptoethanol, 10 % (v/v) glycerol, 0·05 % (w/v) bromophenol blue] and boiled for 5 min as described previously (Takamatsu *et al.*, 2000). The proteins were separated by 14 % SDS-PAGE and visualized by Coomassie brilliant blue R-250 staining (Takamatsu *et al.*, 2000).

In-gel enzymic digestion. Coomassie-blue-stained bands were serially divided into 64 slices from the top band (66 kDa) to the bromophenol blue line at the bottom of the gel at intervals of about 1 mm. The gel slices were washed, reduced by DTT, alkylated by iodoacetamide and digested with modified trypsin or chymotrypsin (Roche) at 37 °C for 16 h as described previously (Shevchenko *et al.*, 1996). The resultant peptides

Table 2. Proteins identified in extracts from *B. subtilis* spores

The 64 gel slices described in Fig. 1 were analysed by LC-MS/MS; the proteins detected in each slice are listed alphabetically.

Gel slice no.	Molecular mass (kDa)	Detected protein
1	66	CotA CotB Epr
2	63	CotA CotB DppE
3	59	CotA CotB CotE OppA YcgN YuiE
4	55	CotA CotB CotE YcgN
5	52	CotB CotE CotY YaaH YdbR YopK
6	50	CotB YaaH
7	48	CotB CotE CotY MelA YaaH YpeB YvdP
8	46	CotB PhoA TufA YgaK YtxN YvdP
9	45	AhpF CotE PhoA YdhO YgaK YtxN YvdP
10	44	CotB CotY YdhD YhcY YtcC YtxN YvrK
11	43·5	CotB Eno YdhD YkvP
12	43	CotB CotE DacF YhfE YjbX YopG YtaA
13	42	AprX CotB CotE CotH CotS DacF YheC YoaN YtaA
14	41	CotB CotS Gap YoaN
15	39	CotB CotE DacB YoaN
16	38	CotB CotE CotG CotY YufD
17	37	CotB CotE CotG CitH CotY YerC YhxC
18	36	CotE CotG
19	35	CotG CotY YtdI
20	34	CotB CotE CotG CotY
21	33	CotB CotE CotG PrsA YdjK YqfQ YuxL YvdO
22	32	CotB CotG CtaC CysK Gdh PrsA YfmC YvdO
23	31	CotB CotE CotG YjqC YusA YvdO
24	30·5	CotB CotE CotG YhcH YjqC YusA
25	30	CotE CotG SleB YisY YvdO
26	29·5	CotE CotG SleB YcdA YckK YisY YitS YtxO YvdO
27	29	CotE LipA
28	28·5	CotD CotE DnaE RplA RpsC SpsI YfkD YqiX YrbA YvgV YwrJ
29	28	CotE Tgl YhjI YvdD YybI
30	27·5	CotE YbaN YdcC YerQ
31	27	CotE Mpr YdcC YvaB
32	26	CotE Mpr SodF YcsK YdcC
33	25·5	CotE SodA SqhC YcsK YlaJ YrbB
34	25	CotE CotJC CotY YhcN YlaJ YrbB
35	24·5	CotE CotJC GltA YhcN YrbB
36	23·5	CotE YwdL YxeE
37	23	CotE CotY YwdL YxeE
38	22·5	CotE CotD CotY
39	22	CotE CotY YhcM YxkC
40	21·5	CotE CotY YdcN
41	20·5	CotE CotY YqaN YwqH YxeE
42	20	CotC CotE CotY YhcM YnzH
43	19·5	CotC CotE CotY GerD SpoIVA YnzH YxeE
44	19	CotE CotY QcrA RpsG SpoIVA YnzH YxeE
45	18·5	CgeA CotC CotD CotE CotY CwlJ FtsH RpsG RpsL SpoIVA
46	18	CotC CotD CotE CotY CotZ CwlJ RpsL YkuD YopY YtfJ YxeE
47	17·5	CotC CotD CotE CotY SspE YdjK YjfA YvdP
48	17	CotC CotE CotD SspE YjcN YrkC
49	16·5	CotC CotD CotE SspE YrkC
50	16	CotC CotD CotE SspE YhcV YhjR
51	15·5	CotC CotD CotE CotS PksM RplR RplX SspE YhjR
52	15	CotC CotD CotE CotS RbfA RplQ RplR RplX YhjR
53	14·5	CotD CotE CotW CotY
54	14	CotD CotW YfhD YqjR
55	13·5	CotB CotD CotE CotW DnaE RpmA Tlp YqaN YrdD
56	12	CotB CotC CotE RpmA
57	11	CotB CotC CotD CotE YabG YfhD YhfS
58	9·5	CotB CotC CotD RpmI YxeE
59	8	CotC CotD RpsP YdbO YfjU YxeE
60	6·5	CotD Hbs RpsP YjdH YkzE YxeE
61	6	CotB CotD CotE CotJA RpsO RpsT YjdH YkzE YuzC YxeE
62	5·5	CotC CotD CotE CotF YsnD
63	5	CotD CotF CotJB CotK CotL SspC YclI YodI YkzE YmaG
64	<4·5	CotD CotJB CotK RpmC SspB SspC YmaG YodI

were eluted from the gel with several changes of extraction buffer [70 μl of 70 % (v/v) acetonitrile, 5 % (v/v) formic acid] and concentrated by evaporation. The samples were then diluted with 5 % formic acid and 50 % acetonitrile for further analysis.

Peptide analysis and protein identification by LC-MS/MS. Nanoliquid chromatography combined with nanoelectrospray ion-trapping tandem mass spectrometry (nanoLC-nanoESI-MS/MS) analysis of peptides was carried out with an LC-Q Deca mass spectrometer (Thermoquest) coupled with a microcapillary nano-flow LC Magic 2002 (Michrom Bioresources). Protein identification was performed using the SEQUEST search program (Thermoquest) (Yates *et al.*, 1995). The searched database was the *Bacillus subtilis* Protein Database.

LacZ assay for evaluation of gene expression. The Japanese and European Consortia for Functional Analysis of the *B. subtilis* Genome constructed the pMutin strains that we used in this study. Each strain contains a *lacZ* fusion protein to monitor gene expression (Vagner *et al.*, 1998). Chromosomal DNA from the strains was extracted and introduced into the sigma-defective mutants SGF602C, SGE603C, SGG604C and SGK605C by competent cell transformation. Resultant cells were grown on DS sporulation agar medium including X-Gal for 48 h at 37 °C and the colony colour was monitored to detect expression of each gene. Five independent colonies of each transformant were tested.

RESULTS

Analysis of spore proteins by SDS-PAGE and LC-MS/MS

The proteins solubilized from dormant spores in this study were first resolved by SDS-PAGE and the gel was divided into 64 slices from the position of the 66 kDa band to the bottom line (4·5 kDa) as shown in Fig. 1. Each gel slice in-gel-digested was then analysed by LC-MS/MS. The MS/MS raw data were used to search with the SEQUEST search program. To identify protein, generally, a correlation factor (*Xcorr*) greater than 2·0 indicates a highly significant match, and a delta cross-correlation factor (*DelCn*) higher than 0·1 indicates a significant distinction between the best match and the second-best match. From search results, the proteins detected in each gel slice are summarized in Table 2. A total of 154 proteins were identified. The predicted characteristics of 85 proteins are described in the *B. subtilis* Genome Database (http://bacillus.genome.ad.jp/). Of these proteins, 41 have been previously reported as sporulation-specific proteins, and another 44 proteins have been previously reported to be synthesized in vegetatively growing cells or sporangia (Table 3). The remaining 69 proteins identified here have not been reported previously (Table 4). The molecular masses of the proteins identified from gels mostly coincided with the deduced molecular masses in the database. However, some proteins, such as the spore-coat proteins CotB, CotC, CotD, CotE, CotS and CotY, were dispersed among several gel slices and their sizes were calculated based on the range of molecular masses estimated for each protein from the SDS-PAGE analysis.

Expression patterns of genes encoding the proteins extracted from mature spores

To characterize *B. subtilis* genes of unknown function, a series of strains have been constructed using derivatives of a plasmid, pMutin, by the Japanese and European Consortia for Functional Analysis of the *B. subtilis* Genome [JAFAN (http://bacillus.genome.ad.jp/) and SubtiList (http://genolist.pasteur.fr/), respectively]. These plasmids have a *lacZ* reporter gene (Vagner *et al.*, 1998). In each pMutin strain, *lacZ* is located downstream of the 5′ portion of a target gene and becomes transcriptionally active when the target gene is activated. The strains provided by the Consortia allow the pattern of target gene expression to be monitored. In this study, we tried to identify sporulation-specific genes encoding the proteins extracted from *B. subtilis* spores. Two different methods are available for monitoring of LacZ activity; one is an enzyme assay on cells grown in liquid medium and the other is a plate colony assay, using X-Gal as a chromogenic substrate. We preferred the plate colony assay, because this method can be advantageous for performing systematic and first-step screening of the genes involved in sporulation. Of the 69 genes known to be involved in sporulation, we could analyse 57; we excluded the remaining 12 genes from our study for the following reasons. Nine genes, *ydcN*, *ykzE*, *ynzH*, *yodI*, *yopG*, *yopK*, *yopY*, *yqaN* and *yrdD*, were not available from the Consortia stocks. These genes either originated from phages or the construction of disruptant strains is difficult. The three genes *yerQ*, *yjcN* and *yufD* are essential or are related to essential genes for growth. Further analysis of these genes was not necessary, because the purpose of this work was to identify genes specifically transcribed during sporulation. We examined 57 strains and found that 48 expressed LacZ activity on sporulation medium (Fig. 2). The remaining nine strains did not show any LacZ activity (data not shown), suggesting that the expression of *yclI*, *ygaK*, *yhcY*, *yitS*, *yjbX*, *yjqC*, *yqjR*, *yvdO* and *yybI* was too weak to be detected under our experimental conditions. The 48 genes that were possibly expressed during sporulation were then further studied as described below.

The expression of the 48 genes possibly expressed during sporulation was analysed by monitoring LacZ activity in sigma-factor-defective cells, such as SigF, SigE, SigG or SigK (Fig. 2, lanes F, E, G and K, respectively). These factors are sporulation-specific and are activated sequentially during sporulation (Driks, 2001; Helmann & Moran, 2001). Each chromosomal DNA of the pMutin strains was transformed into each sigma mutant. The resultant transformants were grown on sporulation medium and their LacZ activity was monitored. The expression of 26 of the 48 genes was controlled by the

Table 3. List of previously identified proteins that were found to be associated with wild-type spores

Among those extracted from mature spores of *B. subtilis*, the proteins whose characteristics have been reported are listed alphabetically. Some of them have been renamed already. The current names and ORF IDs are indicated according to the *Bacillus subtilis* Genome Database JAFAN (http://bacillus.genome.ad.jp). The molecular masses of the proteins were deduced from amino acid sequence data and by the migration positions on the acrylamide gel shown in Fig. 1. The molecular masses of the proteins detected in gel slice no. 64 were estimated to be smaller than 6·4 kDa. Possible signal sequences and the number of transmembrane sequences are based upon information from the JAFAN database.

Protein name	ORF ID	Molecular mass (kDa)		Localization signal	Description
		Deduced from sequence	Calculated by SDS-PAGE		
AhpF	*BG11204*	55	44		Alkyl hydroperoxide reductase
AprX	*BG12567*	48	42		Alkaline serine protease
CgeA	*BG11193*	14	19		Involved in spore-surface property
CitH	*BG11146*	34	37	Signal sequence	Malate dehydrogenase
CotA	*BG10490*	58	66, 63, 59, 55		Spore-coat protein
CotB	*BG10491*	43	66–6·0		Spore-coat protein
CotC	*BG10492*	15	20–5·5		Spore-coat protein
CotD	*BG10493*	8·8	29 to <4·5		Spore-coat protein
CotE	*BG10494*	21	59–5·5		Involved in coat assembly
CotF	*BG10012*	19	5·5, 5·0		Spore-coat protein
CotG	*BG11017*	24	37 to <4·5		Spore-coat protein
CotH	*BG11791*	43	42		Involved in coat assembly
CotJA	*BG11799*	9·7	6·0, 5·0		Spore-coat protein
CotJB	*BG11800*	12	<4·5		Spore-coat protein
CotJC	*BG11801*	22	25		Spore-coat protein
CotK (SspO)	*BG11920*	5·4	5·0, <4·5		Small acid-soluble spore protein
CotL (SspP)	*BG11921*	5·4	5		Small acid-soluble spore protein
CotS	*BG11380*	41	42, 41, 16, 15		Spore-coat protein
CotW	*BG10497*	12	15, 14		Spore-coat protein
CotY	*BG10498*	18	52–15		Spore-coat protein
CotZ	*BG10499*	17	18		Spore-coat protein
CtaC	*BG10215*	40	32	Transmembrane (3)	Cytochrome *c* oxidase I
CwlJ	*BG11172*	16	19, 18		Germination protein
CysK	*BG10136*	33	32	Transmembrane (1)	Cysteine synthase
DacB	*BG10527*	43	39	Signal sequence	Involved in cortex development
DacF	*BG10295*	43	43	Signal sequence	Involved in cortex development
DnaE	*BG12583*	125	29, 14		DNA polymerase III
DppE	*BG10846*	63	63	Signal sequence	Dipeptide ABC transporter
Eno	*BG10899*	47	44		Enolase
Epr	*BG10561*	70	66	Signal sequence	Serine protease
FtsH	*BG10132*	71	19	Transmembrane (2)	Membrane metalloprotease
Gap	*BG10827*	36	41		Glucose dehydrogenase in core
GDH	*BG10545*	28	32		Glyceraldehyde-3-phosphate dehydrogenase
GerD	*BG10644*	21	20	Signal sequence	Germination protein
GltA	*BG10811*	17	25		Glutamate synthase
Hbs	*BG10276*	9·9	6·5		Non-specific DNA-binding protein
LipA	*BG14038*	23	31	Signal sequence	Lipase
MelA	*BG12615*	49	48		α-D-Galactoside galactohydrolase
Mpr	*BG10690*	34	27, 26	Signal sequence	Extracellular metalloprotease
OppA	*BG10771*	61	59	Signal sequence	Oligopeptide ABC transporter
PhoA	*BG10183*	50	46, 45	Signal sequence	Alkaline phosphatase A
PksM	*BG10931*	477	16		Polyketide synthase
PrsA	*BG10464*	32	33, 32	Signal sequence	Protein secretion
QcrA	*BG11325*	19	19	Signal sequence	Menaquinol cytochrome *c* oxidase
RbfA	*BG10270*	13	15		Ribosome-binding factor A
RplA	*BG10164*	25	29		Ribosomal protein L1
RplQ	*BG11041*	14	15		Ribosomal protein L17

Table 3. (*cont.*)

Protein name	ORF ID	Molecular mass (kDa)		Localization signal	Description
		Deduced from sequence	Calculated by SDS-PAGE		
RplR	*BG11409*	13	16, 15		Ribosomal protein L18
RplX	*BG10759*	11	16, 15		Ribosomal protein L24
RpmA	*BG10335*	10	13, 12		Ribosomal protein L27
RpmC	*BG10756*	7·7	<4·5		Ribosomal protein L29
RpmI	*BG11972*	7·6	9·5		Ribosomal protein L35
RpsC	*BG19005*	24	29		Ribosomal protein S3
RpsG	*BG19006*	18	19		Ribosomal protein S7
RpsL	*BG19009*	15	19, 18		Ribosomal protein S12
RpsO	*BG19010*	11	6		Ribosomal protein S15
RpsP	*BG10831*	10	8·0, 6·5		Ribosomal protein S16
RpsT	*BG11643*	9·5	5·5		Ribosomal protein S20
SleB	*BG11439*	34	30	Signal sequence	Germination protein in cortex
SodA	*BG11676*	25	26		Superoxide dismutase
SodF	*BG12676*	33	26		Superoxide dismutase
SpoIVA	*BG10275*	55	20, 19		Involved in coat and cortex development
SpsI	*BG10617*	28	29		Spore-coat-polysaccharide synthesis
SqhC	*BG12679*	63	44, 26		Squalene/hopene cyclase
SspB	*BG10787*	7	<4·5		Small acid-soluble spore protein
SspC	*BG10882*	7·8	5·0, <4·5		Small acid-soluble spore protein
SspE	*BG10789*	9·3	18, 17, 16		Small acid-soluble spore protein
Tgl	*BG10946*	28	28		Transglutaminase
Tlp	*BG11806*	9·7	14, 10		Thioredoxin-like protein
TufA	*BG11056*	44	46		Elongation factor Tu
YaaH	*BG10080*	49	52, 50, 48	Cell-wall-binding motif	Germination protein
YabG	*BG10106*	33	11		Protease involved in coat assembly
YbaN	*BG11416*	28	28	Signal sequence	Involved in sporulation
YdhD	*BG12181*	49	44	Cell-wall-binding motif	Spore protein involved in germination
YdjK (IolT)	*BG12802*	52	33	Transmembrane (11)	Major inositol-transport protein
YfjU (SspH)	*BG12917*	6·9	8		Small acid-soluble spore protein
YhcN	*BG11592*	21	25	Signal sequence	Involved in outgrowth
YkuD	*BG13288*	18	18	Cell-wall-binding motif	Spore protein
YkvP	*BG13318*	46	43·5	Cell-wall-binding motif	Spore protein
YoaN	*BG13484*	44	42, 41, 39		Oxalate decarboxylase
YpeB	*BG11001*	51	48	Signal sequence	Germination protein
YrbA (SafA)	*BG13781*	43	29	Cell-wall-binding motif	Coat protein involved in coat assembly
YrbB	*BG13782*	20	26, 25	Signal sequence	Cortex protein
YtxN (CotSA)	*BG11381*	43	46, 45, 44		Spore-coat protein
YvrK	*BG14148*	44	44		Oxalate decarboxylase

sporulation-specific sigma factors. From the results shown in Fig. 2 and the known sigma cascade of *B. subtilis*, possible regulators of the novel spore proteins are summarized in Table 4. The transcription of the *yfhD*, *yhcM* and *ytfJ* genes in the prespore is possibly controlled by σ^{F}. The transcription of the *ydcC*, *yheC*, *yhjR*, *yhxC*, *yjdH*, *yjfA*, *ysnD*, *yuzC*, *ywdL* and *yxeE* genes in the mother-cell compartment is possibly controlled by σ^{E}. The transcription of the *yfkD*, *yhcV*, *yisY*, *ylaJ* and *yvdP* genes in the forespore is possibly controlled by σ^{G}. The transcription of *ycsK*, *ymaG*, *yqfQ*, *yrkC*, *ytaA*, *ytcC*, *ytxO* and *ywrJ* in the mother-

Table 4. Description of the novel spore proteins identified in this study

Among those extracted from mature spores of *B. subtilis*, the proteins whose characteristics have not been reported are listed alphabetically. ORF IDs are according to the *Bacillus subtilis* Genome Database JAFAN (http://bacillus.genome.ad.jp). The molecular mass of each protein was deduced from amino acid sequence data and by the migration position on the acrylamide gel shown in Fig. 1. The molecular masses of the proteins detected in gel slice no. 64 were estimated to be smaller than 6·4 kDa. Possible signal sequences and the number of transmembrane sequences are based upon information from the JAFAN database. Predictions of the sigma factors that regulate the synthesis of each protein were based upon the results of the LacZ assay described in Fig. 2. SigA and/or other sigma factors possibly regulate the proteins whose syntheses are independent of SigF, SigE, SigG or SigK.

Protein name	ORF ID	Molecular mass (kDa)		Localization signal	Possible regulator
		Deduced from sequence	Calculated by SDS-PAGE		
YcdA	*BG12757*	39	30	Signal sequence	SigA and/or others
YcgN	*BG12012*	56	59, 55		SigA and/or others
YckK	*BG11187*	29	30	Signal sequence	SigA and/or others
YclI	*BG12030*	52	5·0, <4·5	Transmembrane (4)	ND
YcsK	*BG11232*	24	26		SigK
YdbO	*BG12082*	31	21, 8·0	Transmembrane (4)	SigA and/or others
YdbR	*BG12085*	57	52		SigA and/or others
YdcC	*BG12090*	42	28, 27, 26	Transmembrane (1)	SigE
YdcN	*BG12101*	15	22		NT
YdhO	*BG12192*	48	45	Transmembrane (8)	SigA and/or others
YerC	*BG12831*	12	37		SigA and/or others
YerQ	*BG12843*	33	28		NT
YfhD	*BG12879*	7·3	16, 14, 11		SigF
YfkD	*BG12921*	30	29	Signal sequence	SigG
YfmC	*BG12954*	35	32	Signal sequence	SigA and/or others
YgaK	*BG12231*	41	46, 45		ND
YhcH	*BG11586*	34	31		SigA and/or others
YhcM	*BG11591*	17	22, 20	Signal sequence	SigF
YhcV	*BG11600*	15	16		SigG
YhcY	*BG13005*	42	44		ND
YheC	*BG13035*	42	42		SigE
YhfE	*BG13051*	39	43		SigA and/or others
YhfS	*BG13063*	38	11	Transmembrane (1)	SigA and/or others
YhjI	*BG13075*	45	28	Transmembrane (12)	SigA and/or others
YhjR	*BG13084*	17	16,15		SigE
YhxC	*BG11058*	31	37		SigE
YisY	*BG13104*	31	30		SigG
YitS	*BG12244*	31	30		ND
YjbX	*BG13153*	26	43		ND
YjcN	*BG13167*	12	17	Signal sequence	NT
YjdH	*BG13180*	15	6·5, 6·0, 5·5	Signal sequence	SigE
YjfA	*BG13185*	17	18	Signal sequence	SigE
YjqC	*BG13220*	31	31		ND
YkzE	*BG13333*	6·9	6		NT
YlaJ	*BG13347*	23	25	Signal sequence	SigG
YmaG	*BG13415*	14	5·0, <4·5		SigK
YnzH	*BG13471*	12	20, 19		NT
YodI	*BG13537*	9·1	5·0, <4·5	Transmembrane (1)	NT
YopG	*BG13640*	5·3	43		NT
YopK	*BG13644*	45	52		NT
YopY	*BG13658*	11	18		NT
YqaN	*BG11265*	16	21, 14		NT
YqfQ	*BG11663*	26	33		SigK
YqiX	*BG11727*	28	29		SigA and/or others

Table 4. (*cont.*)

Protein name	ORF ID	Molecular mass (kDa)		Localization signal	Possible regulator
		Deduced from sequence	Calculated by SDS-PAGE		
YqjR	*BG11747*	50	14	Signal sequence	ND
YrdD	*BG12284*	6	14		NT
YrkC	*BG11769*	21	17		SigK
YsnD	*BG12335*	12	5·5	Signal sequence	SigE
YtaA	*BG13821*	41	43, 42		SigK
YtcC	*BG13830*	46	44		SigK
YtdI	*BG13837*	30	35		SigA and/or others
YtfJ	*BG13849*	16	18		SigF
YtxO	*BG11382*	17	30		SigK
YufD	*BG12345*	52	38	Transmembrane (14)	NT
YuiE	*BG13970*	54	59	Transmembrane (1)	SigA and/or others
YusA	*BG14013*	30	31	Signal sequence	SigA and/or others
YuxL	*BG10463*	74	33		SigA and/or others
YuzC	*BG14052*	14	6		SigE
YvaB	*BG14058*	23	27		SigA and/or others
YvdD	*BG12413*	21	28		SigA and/or others
YvdO	*BG12423*	35	33, 32, 31		ND
YvdP	*BG12424*	50	48, 46, 45, 18		SigG
YvgV	*BG14104*	25	29	Signal sequence	SigA and/or others
YwdL	*BG10608*	20	24, 23	Signal sequence	SigE
YwqH	*BG12512*	16	21, 14		SigA and/or others
YwrJ	*BG12526*	26	29	Transmembrane (1)	SigK
YxeE	*BG11881*	15	24–6·0		SigE
YxkC	*BG12541*	23	22	Transmembrane (1)	SigA and/or others
YybI	*BG10022*	30	28		ND

ND, No sigma factors can be suggested that are involved in the protein synthesis, because the strains carrying pMutin T3 did not show LacZ activity on sporulation medium; NT, not analysed, because either the strains have not yet been constructed or the genes were essential for vegetative growth.

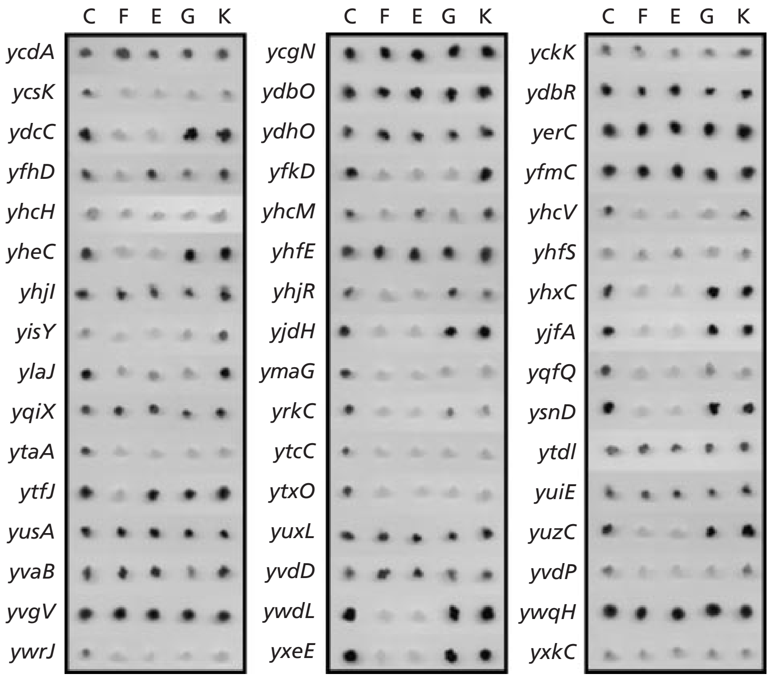

Fig. 2. Expression pattern of genes encoding the proteins extracted from mature spores. LacZ activity of pMutin strains having a *lacZ* fusion was monitored by growth of the cells on DS medium agar containing X-Gal at 37 °C for 48 h. Expression of each gene was monitored by observing the colour of the colonies. Null-mutants of each sigma factor, $SigF^-$, $SigE^-$, $SigG^-$ and $SigK^-$, were also used. The chromosomal DNA of pMutin strains carrying *lacZ* gene fusions was introduced into the null-mutant of each sporulation-specific sigma factor. Lanes: C, strains carrying wild-type sigma factors; F, $SigF^-$ strains; E, $SigE^-$ strains; G, $SigG^-$ strains; K, $SpoIVCB^-(SigK^-)$ strains.

cell compartment is possibly controlled by σ^K. The expression of the remaining 22 genes is independent of these sigma factors. Fawcett *et al.* (2000) have found that the transcription of *yhfS*, *ytfJ* and *ywdL* is reduced in mutant cells of *spoOA* or *sigF* using DNA arrays.

DISCUSSION

Spore proteins of *B. subtilis*

Bacterial spores contain unique high- and low-molecular-mass proteins. Some of these proteins contribute directly or indirectly to the unique characteristics of spores, such as dormancy, a high degree of resistance and particularly a unique cell morphology. Therefore, a comprehensive analysis of protein composition will provide useful basic information and will facilitate our understanding of the cell differentiation of *B. subtilis* at the molecular level. Previous studies have reported that at least 45 proteins are associated with spores of *B. subtilis* (Driks, 1999; Henriques & Moran, 2000). These studies have all been performed by resolution and detection of the spore proteins using one-dimensional gel electrophoresis. Identification of spore proteins has technical problems, because both the extraction of total proteins from mature spores and the purification of each protein are quite difficult. We applied a recently developed LC-MS/MS system to this work and succeeded in identifying over 200 proteins and polypeptides from *B. subtilis* spores (Fig. 1; Table 2). It is notable that this system can distinguish each protein in a protein mixture and identify unknown polypeptides; it can also detect many membrane-bound proteins. These observations indicate that this system is usable for proteome analyses of bacterial spores.

The majority of known spore proteins, small acid-soluble proteins (SASPs) and coat proteins of *B. subtilis* were identified in this work. Some of these proteins are presumed to function in the morphogenesis and germination of spores, as described previously (Piggot & Losick, 2001; Paidhungat & Setlow, 2001). A total of 16 polypeptides originally characterized as spore-coat proteins were identified under our experimental conditions. In this work, CotJB, the coat protein encoded by the second gene of the *cotJ* operon (Seyler *et al.*, 1997; Henriques *et al.*, 1995), was detected in gel slice no. 64, while this protein had not been isolated from spores in previous studies. Four coat proteins, CotM (Henriques *et al.*, 1997), CotT (Bourne *et al.*, 1991), CotV (Zhang *et al.*, 1993) and CotX (Zhang *et al.*, 1993), were not detected in this work. TasA has been reported as the protein associated with spores at least under certain conditions (Serrano *et al.*, 1999). It was also not found in this study. Other coat proteins, CotB, CotC, CotD, CotE, CotS and CotY, were dispersed among several gel slices. These results suggest that some spore proteins are subjected to processing or cross-linking during sporulation. Indeed, we have previously identified one of the sporulation-specific proteases, YabG, which is involved in processing of some coat proteins during morphogenesis of the coat layer (Takamatsu *et al.*, 2000), and this protein was found in gel slice no. 57. SodA and Tgl were found in gel slices nos 33 and 29, respectively; these are possibly involved in protein modification (Henriques *et al.*, 1998; Kobayashi *et al.*, 1998). Other processing enzymes, AprX (Valbuzzi *et al.*, 1999), FtsH (Deuerling *et al.*, 1995) and Mpr (Rufo *et al.*, 1990; Sloma *et al.*, 1990), were also identified from spores in this work. We speculate that these and some unidentified modification enzymes are involved in sporulation or germination processes, either directly or indirectly. Seventeen genes deduced to encode SASPs were found in the genome sequence project (Kunst *et al.*, 1997) and many of the corresponding polypeptides have been identified from acid extracts of spores (Francesconi *et al.*, 1988; Bagyan *et al.*, 1998). In this work, seven of them were identified from spores extracted in the absence of acidic reagents. We prepared spore samples very carefully, as described Methods. In this study, we did not use any spore preparations that contained a protein, Hag (flagellin), which would suggest contamination by vegetative cells. Hag is a major component of flagella, which is associated with vegetative cells and is involved in the motility of *B. subtilis* (LaVallie & Stahl, 1989; Hirose *et al.*, 2000).

Here, we identified not only spore-specific proteins but also proteins involved in DNA maintenance, protein synthesis, metabolism, transport, secretion, proteolysis, etc. (Table 3). These proteins are probably required for resistance, germination or outgrowth of spores. It is, of course, difficult to solubilize total proteins from dormant spores even with harsh treatments. The above results indicate that at least some proteins organized in the spore coat, cortex and core are obtained under our experimental conditions. Many functionally unknown genes remain in the *B. subtilis* database and some of them could encode spore proteins.

Similarity of the spore proteins

Analysis of amino acid sequences showed the similarity of primary sequence between novel proteins and previously known spore proteins. The entire sequence of YnzH was similar to the corresponding region of CotC (71 %). YtaA is a paralogue of CotS, and YtcC is a paralogue of CotSA (YtxN). The *cotS* operon is located near both *ytaA* and *ytcC*. *cotS*, *cotSA* (*ytxN*), *ytaA* and *ytcC* are transcribed under the regulation of σ^K. Therefore, we hypothesize that the functions of CotS and YtaA and those of CotSA (YtxN) and YtcC are similar. YjqC has primary sequence similarity (31 %) to the corresponding region of CotJC. Both YhxC and YxbG, whose functions are still unknown, are similar to glucose dehydrogenase (GDH). The C-terminal half of YrvJ has similarity (28 %) to the corresponding region of CwlD. Although we obtained important information from comparisons of the amino acid sequences, the function of some novel spore proteins is still unknown. Our preliminary results showed that the phenotypes of some of the pMutin strains, such as the heat resistance of spores, are the same as the wild-type, but others are defective in germination (H. Takamatsu, R. Kuwana, M. Fujibayashi & K. Watabe, unpublished data).

Deduced localization of novel spore proteins

By analysis of gene expression using *lacZ* fusions, we found that the transcription of 26 genes encoding novel spore proteins was controlled by sporulation-specific sigma factors, namely σ^F, σ^E, σ^G and σ^K. These sigma factors are temporally and spatially activated and regulate gene expression in a compartment-specific fashion (Igo & Losick, 1986; Stragier & Losick, 1996).

In general, spores of *B. subtilis* have three distinct structures, the core, the cortex and the coat. The proteins synthesized in the mother-cell compartment under the regulation of σ^E and σ^K will be localized in the cortex or coat. In contrast, proteins synthesized in the forespore under the regulation of σ^F and σ^G will be localized in the cortex and core (Driks, 1999). Additionally, TasA has been reported as a protein that is synthesized under the regulation of σ^H, associated with the spore coat and involved in spore morphogenesis (Serrano *et al.*, 1999). A core protein synthesized in the mother-cell compartment or a coat protein synthesized in the forespore has never been reported. Possible signal sequences or transmembrane regions were found in many proteins, and these may facilitate the translocation of the proteins to outer- and inner-spore membranes. From these observations, we have predicted the site of localization of newly identified proteins. YcsK, YheC, YhjR, YhxC, YmaG, YqfQ, YrkC, YtaA, YtcC, YtxO, YuzC and YxeE are possibly synthesized in the mother-cell compartment and assemble on the spore coat. YdcC, YjdH, YjfA, YsnD, YwdL and YwrJ are also possibly synthesized in the mother-cell compartment. We speculate that they localize to the outer-spore membrane or cortex because they have a deduced signal sequence or transmembrane sequence(s). This speculation is supported by the following facts. Both DacB and DacF are cortex proteins and they have a typical signal sequence (Buchanan & Ling, 1992; Wu *et al.*, 1992). DacB is controlled by σ^E (Simpson *et al.*, 1994) and DacF is controlled by σ^F (Schuch & Piggot, 1994). SleB and YrbB also have a signal sequence and their localization in the cortex has been shown by immunological studies (Moriyama *et al.*, 1999; Takamatsu *et al.*, 1998). DacB, DacF, SleB and YrbB were all detectable in our analysis (Table 2). YfhD, YhcV, YisY, YtfJ and YvdP are possibly synthesized in the forespore and exist in the core. YfkD, YhcM and YlaJ have a deduced signal sequence. These proteins are possibly synthesized in the forespore and localize to the inner-spore membrane or cortex. To understand the precise localization site of the newly identified spore proteins and their functions in the cell, we are currently performing a comprehensive analysis of these proteins by using a combination of SDS-PAGE and LC-MS/MS.

ACKNOWLEDGEMENTS

We thank the Japanese and European Consortia for Functional Analysis of the *B. subtilis* Genome for providing the pMutin strains. This work was supported by Grant-in-aids for Scientific Research on Priority Areas (C) 'Genome Biology' from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

REFERENCES

Aronson, A. I. & Fitz-James, P. (1976). Structure and morphogenesis of the bacterial spore coat. *Bacteriol Rev* **40**, 360–402.

Bagyan, I., Setlow, B. & Setlow, P. (1998). New small, acid-soluble proteins unique to spores of *Bacillus subtilis*: identification of the coding genes and regulation and function of two of these genes. *J Bacteriol* **180**, 6704–6712.

Bourne, N., Fitz-James, P. C. & Aronson, A. I. (1991). Structural and germination defects of *Bacillus subtilis* spores with altered contents of a spore coat protein. *J Bacteriol* **173**, 6618–6625.

Buchanan, C. E. & Ling, M.-L. (1992). Isolation and sequence analysis of *dacB*, which encodes a sporulation-specific penicillin-binding protein in *Bacillus subtilis*. *J Bacteriol* **174**, 1717–1725.

Cutting, S. M. & Vander Horn, P. B. (1990). Genetic analysis. In *Molecular Biological Methods for Bacillus*, pp. 27–64. Edited by C. R. Harwood & S. M. Cutting. Chichester: Wiley.

Deuerling, E., Paeslack, B. & Schumann, W. (1995). The *ftsH* gene of *Bacillus subtilis* is transiently induced after osmotic and temperature upshift. *J Bacteriol* **177**, 4105–4112.

Driks, A. (1999). *Bacillus subtilis* spore coat. *Microbiol Mol Biol Rev* **63**, 1–20.

Driks, A. (2001). Proteins of the spore core and coat. In *Bacillus subtilis and its Closest Relatives from Genes to Cells*, pp. 527–535. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology.

Fawcett, P., Eichenberger, P., Losick, R. & Youngman, P. (2000). The transcriptional profile of early to middle sporulation in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* **97**, 8063–8068.

Francesconi, S. C., MacAlister, T. J., Setlow, B. & Setlow, P. (1988). Immunoelectron microscopic localization of small, acid-soluble spore proteins in sporulating cells of *Bacillus subtilis*. *J Bacteriol* **170**, 5963–5967.

Goldman, R. C. & Tipper, D. J. (1978). *Bacillus subtilis* spore coats: complexity and purification of a unique polypeptide component. *J Bacteriol* **135**, 1091–1106.

Gould, G. W. (1983). Mechanisms of resistance and dormancy. In *The Bacterial Spores*, vol. 2, pp. 173–210. Edited by A. Hurst & G. W. Gould. London: Academic Press.

Helmann, J. D. & Moran, C. P., Jr (2001). RNA polymerase and sigma factors. In *Bacillus subtilis and its Closest Relatives from Genes to Cells*, pp. 527–535. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology.

Henner, D. J. (1990). Inducible expression of regulatory genes in *Bacillus subtilis*. *Methods Enzymol* **185**, 223–228.

Henriques, A. O & Moran, C. P., Jr (2000). Structure and assembly of the bacterial endospore coat. *Methods* **20**, 95–110.

Henriques, A. O., Beall, B. W., Roland, K. & Moran, C. P., Jr (1995). Characterization of *cotJ*, a σ^E-controlled operon affecting the polypeptide composition of the coat of *Bacillus subtilis* spores. *J Bacteriol* **177**, 3394–3406.

Henriques, A. O., Beall, B. W. & Moran, C. P., Jr (1997). CotM of

Bacillus subtilis, a member of the α-crystallin family of stress proteins, is induced during development and participates in spore outer coat formation. *J Bacteriol* **179**, 1887–1897.

Henriques, A. O., Melsen, L. R. & Moran, C. P., Jr (1998). Involvement of superoxide dismutase in spore coat assembly in *Bacillus subtilis*. *J Bacteriol* **180**, 2285–2291.

Hirose, I., Sano, K., Shioda, I., Kumano, M., Nakamura, K. & Yamane, K. (2000). Proteome analysis of *Bacillus subtilis* extracellular proteins: a two-dimensional protein electrophoretic study. *Microbiology* **146**, 65–75.

Igo, M. M. & Losick, R. (1986). Regulation of a promoter that is utilized by minor forms of RNA polymerase holoenzyme in *Bacillus subtilis*. *J Mol Biol* **191**, 615–624.

Kobayashi, K., Hashiguchi, K., Yokozeki, K. & Yamanaka, S. (1998). Molecular cloning of the transglutaminase gene from *Bacillus subtilis* and its expression in *Escherichia coli*. *Biosci Biotechnol Biochem* **62**, 1109–1114.

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.

LaVallie, E. R. & Stahl, M. L. (1989). Cloning of the flagellin gene from *Bacillus subtilis* and complementation studies of an *in vitro*-derived deletion mutation. *J Bacteriol* **171**, 3085–3094.

Moriyama, R., Fukuoka, H., Miyata, S., Kudoh, S., Hattori, A., Kozuka, S., Yasuda, Y., Tochikubo, K. & Makino, S. (1999). Expression of a germination-specific amidase, SleB, of bacilli in the forespore compartment of sporulating cells and its localization on the exterior side of the cortex in dormant spores. *J Bacteriol* **181**, 2373–2378.

Paidhungat, M. & Setlow, P. (2001). Spore germination and outgrowth. In *Bacillus subtilis and its Closest Relatives from Genes To Cells*, pp. 537–548. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology.

Piggot, P. J. & Losick, R. (2001). Sporulation genes and intercompartmental regulation. In *Bacillus subtilis and its Closest Relatives from Genes To Cells*, pp. 483–517. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology.

Rufo, G. A., Jr, Sullivan, B. J., Sloma, A. & Pero, J. (1990). Isolation and characterization of a novel extracellular metalloprotease from *Bacillus subtilis*. *J Bacteriol* **172**, 1019–1023.

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

Schuch, R. & Piggot, P. J. (1994). The *dacF–spoIIA* operon of *Bacillus subtilis*, encoding σ^F, is autoregulated. *J Bacteriol* **176**, 4104–4110.

Serrano, M., Zilhao, R., Ricca, E., Ozin, A. J., Moran, C. P., Jr & Henriqous, A. O. (1999). A *Bacillus subtilis* secreted protein with a role in endospore coat assembly and function. *J Bacteriol* **181**, 3632–3643.

Seyler, R. W., Jr, Henriques, A. O., Ozin, A. J. & Moran, C. P., Jr (1997). Assembly and interactions of *cotJ*-encoded proteins, constituents of the inner layers of the *Bacillus subtilis* spore coat. *Mol Microbiol* **25**, 955–966.

Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem* **68**, 850–858.

Simpson, E. B., Hancock, T. W. & Buchanan, C. E. (1994). Transcriptional control of *dacB*, which encodes a major sporulation-specific penicillin-binding protein. *J Bacteriol* **176**, 7767–7769.

Sloma, A., Rudolph, C. F., Rufo, G. A., Jr, Sullivan, B. J., Theriault, K. A., Ally, D. & Pero, J. (1990). Gene encoding a novel extracellular metalloprotease in *Bacillus subtilis*. *J Bacteriol* **172**, 1024–1029.

Stragier, P. & Losick, R. (1996). Molecular genetics of sporulation in *Bacillus subtilis*. *Annu Rev Genet* **30**, 297–341.

Takamatsu, H., Hiraoka, T., Kodama, T., Koide, H., Kozuka, S., Tochikubo, K. & Watabe, K. (1998). Cloning of a novel gene *yrbB*, encoding a protein located in the spore integument of *Bacillus subtilis*. *FEMS Microbiol Lett* **166**, 361–367.

Takamatsu, H., Kodama, T., Imamura, A., Asai, K., Kobayashi, K., Nakayama, T., Ogasawara, N. & Watabe, K. (2000). The *Bacillus subtilis yabG* gene is transcribed by SigK RNA polymerase during sporulation, and *yabG* mutant spores have altered coat protein composition. *J Bacteriol* **182**, 1883–1888.

Vagner, V., Dervyn, E. & Ehrlich, S. D. (1998). A vector for systematic gene inactivation in *Bacillus subtilis*. *Microbiology* **144**, 3097–3104.

Valbuzzi, A., Ferrari, E. & Albertini, A. M. (1999). A novel member of the subtilisin-like protease family from *Bacillus subtilis*. *Microbiology* **145**, 3121–3127.

Wu, J.-J., Schuch, R. & Piggot, P. J. (1992). Characterization of a *Bacillus subtilis* sporulation operon that includes genes for an RNA polymerase σ factor and for a putative DD-carboxypeptidase. *J Bacteriol* **174**, 4885–4892.

Yates, J. R., 3rd, Eng, J. K., McCormack, A. L. & Schieltz, D. (1995). Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. *Anal Chem* **67**, 1426–1436.

Zhang, J., Fitz-James, P. C. & Aronson, A. I. (1993). Cloning and characterization of a cluster of genes encoding polypeptides present in the insoluble fraction of the spore coat of *Bacillus subtilis*. *J Bacteriol* **175**, 3757–3766.

Received 30 May 2002; revised 7 August 2002; accepted 12 September 2002.