Analysis of virulence plasmid gene expression defines three classes of effectors in the type III secretion system of Shigella flexneri

Tony Le Gall,1,2 Maria Mavris,3 Maria Celeste Martino,4 Maria Lina Bernardini,4 Erick Denamur1 and Claude Parsot3

Correspondence
Claude Parsot
cparsot@pasteur.fr

1INSERM E0339, Faculté de Médecine Xavier Bichat, 16 rue Henri Huchard, F-75018 Paris, France
2Laboratoire de Bactériologie, Faculté de Médecine de Brest, F-29285 Brest Cedex, France
3Unité de Pathogénie Microbienne Moléculaire, INSERM U389, Institut Pasteur, 25 rue du Dr Roux, F-75724 Paris Cedex 15, France
4University Roma La Sapienza, Sez Sci Microbiol, Dipartimento Biol Cellulare & Sviluppo, Via Sardi 70, Rome, I-00185 Italy

Proteins directly involved in entry and dissemination of Shigella flexneri into epithelial cells are encoded by a virulence plasmid of 200 kb. A 30-kb region (designated the entry region) of this plasmid encodes components of a type III secretion (TTS) apparatus, substrates of this apparatus and their dedicated chaperones. During growth of bacteria in broth, expression of these genes is induced at 37 °C and the TTS apparatus is assembled in the bacterial envelope but is not active. Secretion is activated upon contact of bacteria with host cells and is deregulated in an ipaB mutant. The plasmid encodes four transcriptional regulators, VirF, VirB, MxiE and Orf81. VirF controls transcription of virB, whose product is required for transcription of entry region genes. MxiE, with the chaperone IpgC acting as a co-activator, controls expression of several effectors that are induced under conditions of secretion. Genes under the control of Orf81 are not known. The aim of this study was to define further the repertoires of virulence plasmid genes that are under the control of (i) the growth temperature, (ii) each of the known virulence plasmid-encoded transcriptional regulators (VirF, VirB, MxiE and Orf81) and (iii) the activity of the TTS apparatus. Using a macroarray analysis, the expression profiles of 71 plasmid genes were compared in the wild-type strain grown at 37 and 30 °C and in virF, virB, mxiE, ipaB, ipaB mxiE and orf81 mutants grown at 37 °C. Many genes were found to be under the control of VirB and indirectly of VirF. No alteration of expression of any gene was detected in the orf81 mutant. Expression of 13 genes was increased in the secretion-deregulated ipaB mutant in an MxiE-dependent manner. On the basis of their expression profile, substrates of the TTS apparatus can be classified into three categories: (i) those that are controlled by VirB, (ii) those that are controlled by MxiE and (iii) those that are controlled by both VirB and MxiE. The differential regulation of expression of TTS effectors in response to the TTS apparatus activity suggests that different effectors might be required at different times following contact of bacteria with host cells.

INTRODUCTION

Bacteria of Shigella species are responsible for shigellosis, a human disease characterized by the destruction of the colonic epithelium. Shigellae and enteroinvasive Escherichia coli both contain a virulence plasmid (VP) of approximately 200 kb that encodes most proteins directly involved in entry and dissemination of bacteria into epithelial cells. Sequence analysis of the VP pWR100 (Buchrieser et al., 2000) and its derivative pWR501 (Venkatesan et al., 2001) from a Shigella flexneri strain of serotype 5 and the VP pCP301 (Jin et al., 2002) from an S. flexneri strain of serotype 2a indicated that the VP is composed of a mosaic of approximately 100 genes and numerous insertion sequences. The VP genes exhibit a G+C content from 30 to 60 mol%, suggesting that they were acquired from different sources.

Abbreviations: TTS, type III secretion; VP, virulence plasmid.
Genes required for entry of bacteria into epithelial cells and the induction of apoptosis in infected macrophages are clustered on a 30 kb region (designated the entry region) of the VP. This region encodes components of a type III secretion (TTS) apparatus (the Mxi-Spa TTS apparatus), substrates of this secretion apparatus (the translocators IpaB and IpaC and the effectors IpaD, IpgB1, IpgD and IcsB) and their dedicated chaperones (IpgA, IpgC, IpgE and Spa15) and two transcriptional activators (VirB and MxiE).

The current model of the TTS pathway proposes that, upon contact of bacteria with host cells, translocators insert into the membrane of the host cell to form a pore through which effectors transit to reach the cell cytoplasm (Hueck, 1998). Other substrates of the TTS apparatus are encoded by genes scattered throughout the VP, such as virA, ospB, C, D, E, F and G and ipaH genes (Buchrieser et al., 2000).

Several of these putative effectors are encoded by multigene families, with five ipaH, four ospC, three ospD and two ospE genes carried by the VP. Genes encoding components of the TTS apparatus and its substrates exhibit a similar low G+C content (approx. 34 mol%), suggesting that the entire TTS system was acquired once by lateral transfer. In addition, the VP encodes at least five other proteins that are involved in virulence, including IcsA, IcsP, VirK, MsbB2 and SepA. IcsA (VirG) is an outer-membrane protein directly involved in promoting actin polymerization at one pole of intracellular bacteria (Bernardini et al., 1989; Goldberg & Theriot, 1995; Lett et al., 1989). IcsP (SopA) is an outer-membrane protease involved in the release of a certain proportion of surface-exposed IcsA (Egile et al., 1997; Shere et al., 1997). VirK is required for production of IcsA by an unknown mechanism (Nakata et al., 1992). MsbB2 is an acyl transferase that, in conjunction with the product of the chromosomal msbB1 gene, acts to produce full acyl-oxy-acylation of the myristate at the 3’ position of the lipid A glucosamine disaccharide (d’Hauteville et al., 2002). SepA is a secreted serine protease of the IgA1 protease family whose substrate is not known (Benjelloun-Touimi et al., 1995, 1998).

The regulation of expression of VP genes has been the subject of extensive studies [see Dorman & Porter (1998) for a review]. Expression of genes of the entry region is regulated by the growth temperature, these genes being expressed at 37 °C but not at 30 °C (Maurelli et al., 1984), and by a cascade involving two VP gene products, VirF and VirB. VirF, a member of the AraC family of transcriptional activators, is required for transcription of virB (Sakai et al., 1986; Tobe et al., 1993) and VirB, a member of the ParB family of partitioning proteins, is required for transcription of genes of the entry region (Adler et al., 1989; Tobe et al., 1991). Control by temperature is dependent upon the chromosomally encoded protein H-NS, as inactivation of hns leads to expression of genes of the entry region at both 30 and 37 °C (Maurelli & Sansonetti, 1988; Prosseda et al., 1998). The current model proposes that binding of H-NS to the virF and virB promoters, as well as to promoters controlled by VirB, prevents transcription of these genes at 30 °C (Beloin & Dorman, 2003). At 37 °C, changes in DNA conformation at virF and virB promoters lead to an increased transcription of virF and to the activation of the virB promoter by VirF (Dorman et al., 2001). Production of VirB, in turn, leads to activation of promoters of the entry region by a mechanism that has not yet been elucidated but that might involve displacement of H-NS (Adler et al., 1989; Beloin et al., 2002; Beloin & Dorman, 2003; McKenna et al., 2003). During growth of bacteria in broth at 37 °C, the TTS apparatus is assembled in the bacterial envelope but is only weakly active (Menard et al., 1994a) and substrates of the TTS apparatus are stored in the cytoplasm in association with specific chaperones (Menard et al., 1994b; Niebuhr et al., 2000; Ogawa et al., 2003; Page et al., 2001, 2002). The TTS apparatus is activated upon both contact of bacteria with epithelial cells (Menard et al., 1994a) and exposure of bacteria to the dye Congo red (Bahrani et al., 1997; Parsot et al., 1995). The TTS apparatus is deregulated, i.e. constitutively active in bacteria growing in broth, by inactivation of ipaB or ipaD (Menard et al., 1994a). Conditions of active or deregulated secretion lead to increased transcription of members of the ipaH family, ospF, virA and ospC1, but not genes of the entry region (Demers et al., 1998; Mavris et al., 2002a). Transcription of genes regulated by the TTS apparatus activity involves both MxiE, a transcriptional activator of the AraC family encoded by the entry region, and IpgC, the chaperone for the IpaB and IpaC translocators, acting as a co-activator of MxiE (Mavris et al., 2002a). Under conditions of non-secretion, IpaB and IpaC are associated with and titrate IpgC and MxiE is not active. Under conditions of secretion, IpaB and IpaC are secreted and IpgC is liberated in the cytoplasm, where it activates MxiE by an unknown mechanism likely to involve an interaction between MxiE and IpgC. Much less is known about the regulation of expression of VP genes that are not part of the TTS system. Production of IcsA (VirG) was not affected in a virB mutant (Adler et al., 1989) and transcription of an icsA-lacZ fusion carried by a reporter plasmid was decreased 2.5-fold in a virF mutant compared with the wild-type strain (Sakai et al., 1988), suggesting that expression of icsA was directly under the control of VirF. Northern blot analysis indicated that expression of icsA was moderately (2-fold) modulated by the temperature of growth (Porter & Dorman, 1997). Expression of sepA was not modulated by temperature (Benjelloun-Touimi et al., 1995) and expression of icsP was shown to be controlled by VirB (Berlotti et al., 1998; Wing et al., 2004). In addition to VirF and MxiE, pWR100 encodes a third protein of the AraC family, OrfB1, whose function has not yet been investigated (Buchrieser et al., 2000).

The numerous studies that have been performed on the regulation of expression of VP genes have provided a wealth of information on the nature of some genes that are regulated and the mechanism(s) of their control. However, these studies used a variety of techniques and different strains and growth conditions and, as a consequence, their
results are difficult to compare. To obtain a global view on expression of VP genes, we performed a macroarray analysis using membranes carrying PCR fragments corresponding to most VP genes. To define further the repertoires of genes that are under the control of (i) the growth temperature, (ii) each of the known transcriptional regulators encoded by the VP, i.e. VirF, VirB, MxiE and OrfB1, and (iii) the activity of the TTS apparatus, i.e. activated in an ipaB mutant, we analysed the expression profiles in the wild-type strain grown at 37 and 30 °C and in virF, virB, mxiE, ipaB, ipaB mxiE and orfB1 mutants grown at 37 °C. For the sake of clarity, results are presented in the reverse order of the regulatory cascade presented above, i.e. from the analysis of an ipaB mutant at 37 °C, in which most genes are expressed, to that of the wild-type strain at 30 °C, in which most genes are not expressed.

METHODS

Bacterial strains. The S. flexneri wild-type strain M90T-Sm and the ipaB4, mxiE and ipaB4 mxiE mutants were described previously (Allaoui et al., 1992; Mavris et al., 2002a). The orfB1 mutant was constructed by allelic exchange, inserting an aphA3 cassette at codon 25 of the orfB1 gene carried by the VP. The virF mutant was constructed by integration of a suicide plasmid carrying a DNA fragment corresponding to codons 27–118 of virF into the virF gene of pWR100. The virB mutant was constructed by allelic exchange, replacing codons 135–158 of virB by a cassette conferring resistance to chloramphenicol.

Macroarray design. We developed a set of membranes each carrying PCR fragments corresponding to most pWR100 genes (except those encoded by insertion sequences) and four chromosomal genes, adk, cysK, udp and tufB. For each gene, primer pairs were designed to amplify internal DNA fragments of approximately 600 bp specific to each gene, with the following exceptions. In the case of the small mxiH and mxiI genes, the PCR fragment included both genes and was designated mxiHI. The sequences of ipaH1 and ipaH2, ospE1 and ospE2, and ospC2, ospC3 and ospC4 are identical or too similar to be distinguished by hybridization and the corresponding PCR fragments were designated ipaH1/2, ospE1/2 and ospC2/3/4, respectively. The ipaH1.4, ipaH2.5, ipaH4.5, ipaH7.8 and ipaH9.8 genes are designated here ipaH1, ipaH2, ipaH4, ipaH7 and ipaH9, respectively. PCR products were spotted in duplicate copies onto nitrocellulose membranes at the Ge´ nopole at the Pasteur Institute (Paris). As negative controls, S. flexneri genomic DNA at the same concentration as that used for PCR amplifications of plasmid genes and mouse genomic DNA were also spotted onto the membranes.

DNA and RNA extraction and labelling. Bacteria were grown aerobically at 37 °C in 869 medium with constant shaking. Total genomic DNA from the wild-type strain was isolated from 1 ml culture in stationary phase using the standard protocol of the Wizard Genomic DNA purification kit (Promega). Five hundred nanograms of total genomic DNA was used as a template for incorporation of [α-32P]dCTP (Perkin Elmer) in a polymerization reaction using random hexamer primers and the Klenow fragment of DNA polymerase I (Roche) according to the manufacturers’ recommendations. For RNA extraction, an overnight culture of bacteria was grown at 37 and 30 °C, respec-

For transcription profiles, standardization was performed by expressing the intensity value of individual spots relative to the mean calculated over the intensities obtained from the two tufB spots such that the tufB intensities in each experiment were close to 1. The accuracy and reproducibility of the results were analysed by comparing the distribution of intensity in each spot between duplicate dots and between hybridization repeats. In every possible comparison, the results showed a high degree of correlation (r>0.90). The expression profiles of mutants were compared to those of the wild-type strain by calculating the consistency of differential expression across replicate hybridization values using the Wilcoxon two-sample test (Z-test). This non-parametric statistical method contained in the StatView 5.0.1 package (SAS Institute Inc.) tests the hypothesis that one of the paired variables is larger or smaller than the other variable regardless of the magnitude of the difference and is appropriate for the analysis of small samples. This allowed us to identify genes that were...
differentially expressed between two strains: only genes with the lowest possible Z-test P value were considered for analysis. All raw and processed data obtained in this study can be downloaded in tabular format from the websites http://www.pasteur.fr/recherche/unites/Pmm/contents/gb/08_suppldata/suppldata.html or http://www.bichat.inserm.fr/equipes/Emi0339/publications_excel/pWR100transcriptome.xls

RESULTS AND DISCUSSION

Design of the array and overview of results

The VP pWR100 carries approximately 100 genes, not including those of insertion sequences (Buchrieser et al., 2000). Pairs of primers were designed for 75 of these plasmid genes and four chromosomal genes, including tufB, to amplify PCR fragments of approximately the same size (500–600 bp) specific for each gene, except for ipaH1 and ipaH2, ospE1 and ospE2, and ospC2, ospC3 and ospC4, which are identical or too similar to be distinguished by hybridization. PCR fragments for these latter genes were designated ipaH1/2, ospE1/2 and ospC2/3/4. PCR fragments were spotted onto nitrocellulose membranes and each membrane was subjected to hybridizations using the labelled genomic DNA as a probe. Spots corresponding to 71 plasmid genes gave identical signals over all membranes and were considered in the transcriptional analysis presented below. These genes are representative of most, if not all, transcriptional units of the VP. For all these genes, except ipaH1/2 and ospC2/3/4, signals obtained using the labelled genomic DNA as a probe were not significantly different, indicating that differences in the G+C content of the various genes had no influence on the efficiency of hybridization under these experimental conditions. Stronger signals were detected for ipaH1/2 and ospC2/3/4, consistent with the presence of two or three copies of these genes on the VP. For the transcriptional analysis, the intensity of each spot was standardized relative to the mean value calculated over intensities obtained for the tufB spots, since expression of that chromosomal gene was similar in the wild-type strain and in its derivative that had been cured of the VP (T. Le Gall and E. Denamur, unpublished observation). Similar results were obtained using sepA, a plasmid gene whose expression is not controlled by temperature (Benjelloun-Touimi et al., 1995), instead of tufB to standardize the data (data not shown).

We analysed the transcription profiles of 71 plasmid genes in the wild-type strain grown at 37 and 30 °C and in virF, virB, mxiE, ipaB, ipaA mxiE and orf81 mutants grown at 37 °C. Fig. 1 shows the expression profiles of the wild-type strain grown at either 30 °C (open bars) or 37 °C (light shaded bars) and an ipaB mutant grown at 37 °C (dark shaded bars). Since signal intensities obtained upon hybridization of membranes with genomic DNA were similar (except for ipaH1/2 and ospC2/3/4 as indicated above), the comparison of signal intensities obtained upon hybridization of cDNA suggests that genes for which the amounts of mRNA were the most abundant in the wild-type strain grown in broth at 37 °C are those of the entry region.

Fig. 1. Schematic representation of expression levels of plasmid genes in the wild-type strain grown at 30 °C (open bars) and 37 °C (light shaded bars) and in the ipaB mutant grown at 37 °C (dark shaded bars). Genes are indicated according to their position on the VP pWR100 (Buchrieser et al., 2000). Expression levels (in arbitrary units) are indicated relative to the expression threshold (ET). The ET was defined as the mean of signals obtained on non-spotted areas plus two standard deviations.
especially the \textit{ipa} operon, and \textit{virA}. Under conditions of deregulated secretion, such as in the \textit{ipaB} mutant, strong signals were also detected for \textit{ipaH} and some \textit{osp} genes. In most cases, genes either demonstrated or postulated from the genetic organization to belong to the same operon exhibited signals of similar intensities. The few cases where different intensities were observed for genes belonging to the same operon, e.g. \textit{ospB} and \textit{phoN2} or within the \textit{mxi} operon, might be due to different stabilities of parts of the mRNA resulting from mRNA processing. Expression levels of each gene for each experimental condition are shown in Fig. 2 along with the ratios calculated for pairwise comparisons between a mutant and the wild-type strain or between two mutants. Differences in gene expression leading to a ratio >0.5 and <2 were not considered as significant, although this threshold is arbitrary.

\textbf{Genes regulated by the activity of secretion and MxiE}

Previous analyses indicated that transcription of a number of genes encoding substrates of the TTS apparatus is induced when the TTS apparatus is deregulated in \textit{ipaB} and \textit{ipaD} mutants constructed either by insertion of a non-polar cassette in \textit{ipaB} or \textit{ipaD} or by integration of a suicide plasmid in \textit{ipaB} (Demers et al., 1998; Mavris et al., 2002a, b). To determine the repertoire of genes whose expression is increased when the activity of secretion is deregulated, the transcription profile of a strain carrying the \textit{ipaB4} allele (corresponding to the inactivation of \textit{ipaB} by integration of a suicide plasmid) was compared with that of the wild-type strain. As expected, expression of \textit{ipaC}, \textit{ipaD} and \textit{ipaA}, which are located downstream from \textit{ipaB}, was substantially decreased in this mutant compared with the wild-type strain, as a consequence of the polar effect resulting from insertion of the suicide plasmid into \textit{ipaB}. Genes whose expression was increased in the \textit{ipaB} mutant compared with the wild-type strain include \textit{ospB}, \textit{phoN2}, \textit{ospF}, \textit{ipaH7}, \textit{ipaH4}, \textit{ospD3}, \textit{ospC1}, \textit{virA}, \textit{ipaH9}, \textit{ospG}, \textit{ipaH1}/2 and \textit{ospE1}/2 (Fig. 2; ratio \textit{ipaB}:w.t. 37°C).

To determine whether the increased transcription of these genes in the \textit{ipaB} mutant was dependent on MxiE, we analysed the transcription profile of an \textit{ipaB} \textit{mxiE} mutant in which the activity of secretion is also deregulated but that does not produce MxiE. Expression of all genes that were activated in the \textit{ipaB} mutant was decreased in the \textit{ipaB} \textit{mxiE} mutant at a level that was similar to that observed in the wild-type strain (Fig. 2; ratio \textit{ipaB} \textit{mxiE}:w.t. 37°C), confirming that activation of these genes in response to the deregulation of the secretion activity is MxiE dependent.

We also analysed the transcription profile of an \textit{mxiE} mutant that lacks MxiE but in which the activity of secretion is not deregulated, i.e. the TTS apparatus is not active during growth of bacteria in broth. Expression of a few genes was slightly increased in the \textit{mxiE} mutant compared with the wild-type (Fig. 2; ratio \textit{mxiE}:w.t. 37°C); however, the ratios for these genes ranged from 1.2 to 1.5. This indicates that no genes are strongly controlled, either positively or negatively, by MxiE under conditions of non-secretion. This is consistent with the previous observation that the activity of MxiE is dependent on the presence of the co-activator IpgC, which is not available under conditions of non-secretion as it is titrated by IpB and IpC (Mavris et al., 2002a).

The \textit{cis}-acting element required for increased transcription of \textit{ospC1}, \textit{virA} and \textit{ipaH9} promoters was shown to be a 17 bp motif (5'-GTATCGTTTTTTTANAG), designated the MxiE box, which is located between positions −49 and −33 with respect to the transcription start site (Mavris et al., 2002b). The increased expression of \textit{ospF}, \textit{ipaH4}, \textit{ipaH7}, \textit{ospC1}, \textit{virA} and \textit{ipaH9} detected in the present analysis is consistent with previous results obtained using \textit{lacZ} transcriptional fusions (Demers et al., 1998; Mavris et al., 2002b) and with the presence, in the 5′ region of these genes, of a putative MxiE box exhibiting at least 14 matches out of 16 positions with the MxiE box consensus sequence (Fig. 3). The control of \textit{ospD3} transcription by MxiE in the absence of an MxiE box in the \textit{ospC1}–\textit{ospD3} intergenic region suggests that \textit{ospD3} is part of the same operon as \textit{ospC1}, which has an MxiE box in its promoter region. Similarly, control of \textit{ospG} transcription by MxiE suggests that \textit{ospG} is expressed from the promoter of the upstream \textit{ipaH9} gene. There is evidence that \textit{ospE2} is controlled by MxiE (Kane et al., 2002), which is consistent with the presence of an MxiE box upstream from \textit{ospE2}. Since \textit{ipaH2} is separated from \textit{ospE2} by an insertion sequence, it seems unlikely that this gene is expressed, and the signal detected on \textit{ipaH1}/2 spots should correspond only to expression of \textit{ipaH1}. Since there is no MxiE box in the \textit{ospE1}–\textit{ipaH1} intergenic region, regulation of \textit{ipaH1} transcription by the activity of secretion and MxiE suggests that transcription of \textit{ipaH1} is regulated by the MxiE box that is present in the 5′ region of \textit{ospE1}. Expression of \textit{ospB} and \textit{phoN2}, which belong to the same operon (Santapaola et al., 2002), was increased in the \textit{ipaB} mutant and reduced to a wild-type level in the \textit{ipaB} \textit{mxiE} mutant. The DNA region located between positions −106 and −89 with respect to the \textit{ospB} translation start site exhibits 13 matches (out of 16 positions) with the sequence of the MxiE box and is likely to be involved in the MxiE-dependent activation of the \textit{ospB}–\textit{phoN2} operon (Fig. 3). The VP carries two other regions that exhibit 13 matches with the sequence of the MxiE box, at coordinates 125 388 and 144 574; however, these regions are unlikely to represent functional MxiE boxes, since the former is located downstream from and in the opposite orientation to \textit{virA} and the latter is located within \textit{spa47}, whose expression, as that of downstream genes, was not increased in the \textit{ipaB} mutant (Fig. 2).

Using reporter plasmids in which the 5′ region of a number of genes encoding secreted proteins was driving expression of the green fluorescent protein, Kane et al. (2002) showed that \textit{ospB}, \textit{ospF}, \textit{ospC1}, \textit{virA}, \textit{ipaH9} and \textit{ospE2} promoters...
<table>
<thead>
<tr>
<th>Gene</th>
<th>Expression levels</th>
<th>Expression ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>w.t. 37 °C</td>
<td>ipaB</td>
</tr>
<tr>
<td></td>
<td>w.t. 30 °C</td>
<td>w.t. 37 °C</td>
</tr>
<tr>
<td>acpP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cpnB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pheN2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>otnD2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>orfE13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>orfD1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>orf22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pirB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>virF</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Gene expression range colour code (relative to the expression threshold):

- <1
- >1 ; 3>
- >3 ; 5>
- >5 ; 7>
- >7 ; 9>
- >9 ; 11>
- >11 ; 13>
- ≥13
had been activated upon entry of bacteria into epithelial cells, as shown previously for ipaH1, ipaH4, ipaH7, ipaH9 and virA genes (Demers et al., 1998), and that this activation was dependent on the presence of MxiE. No MxiE-dependent expression of ospG was observed from the reporter plasmid (Kane et al., 2002), which is consistent with the hypothesis presented above that ospG is expressed from the ipaH9 promoter.

In addition to the five ipaH genes carried by the VP discussed above, other ipaH genes are carried by the chromosome (Buysse et al., 1995). There are five chromosomal ipaH genes in the S. flexneri strain M90T of serotype 5 and we showed previously that at least some of these genes are controlled by MxiE and IpgC (Mavris et al., 2002b). The recent elucidation of the genomic sequence of two S. flexneri strains of serotype 2a, CP301 and 2457T, revealed the presence of seven chromosomal ipaH genes in these strains, including four complete copies and three copies inactivated by insertion sequences or frameshifting mutations (Jin et al., 2002; Wei et al., 2003). Sequence analysis revealed the presence of an MxiE box upstream from each chromosomal ipaH gene copy of strain CP301 (Fig. 3) and strain 2457T (data not shown). Sequences exhibiting two mismatches with the MxiE box consensus sequence are also present upstream from genes encoding mannose-6-phosphate isomerase and L-aspartate oxidase; however, these sequences are unlikely to represent regulatory sites as they are also present at the same position in the E. coli strain MG1655, which does not contain a TTS system.

In conclusion, comparison of transcription profiles of the wild-type, mxiE, ipaB and ipaB mxiE strains indicates that expression of 13 genes (ipaH1, ipaH4, ipaH7, ipaH9, ospB, phoN2, ospC1, ospD3, ospE1, ospE2, ospF, ospG and virA) was activated under conditions of deregulated secretion in an MxiE-dependent manner. Transcription of operons

**Fig. 2.** Expression profiles of the wild-type strain grown at 30 and 37 °C and virF, virB, ipaB, mxiE, orfB1 and ipaB mxiE mutants grown at 37 °C. Genes are listed in the left column according to their map position on pWR100 (Buchrieser et al., 2000). For each gene, the direction of transcription is indicated by the orientation of the corresponding arrowhead and the G+C content of that gene is indicated by the colour of the arrow head: red, below 40 mol%; blue, 40–50 mol%; green, above 50 mol%. Arrows indicate the direction and extent of potential operons, as suggested by the genetic organization. The position of insertion sequences is indicated by dashed black lines. Expression levels are indicated by colour-coded squares, using the code indicated at the bottom of the figure. The expression threshold was defined as the mean of signals obtained on non-spotted areas plus two standard deviations. Ratios for salient pairwise comparisons were calculated for each gene and only those for which the lowest possible Wilcoxon Z-test P value was obtained are shown. In these comparisons, expression ratios concerning the gene that was inactivated in a given mutant, e.g. the virB gene in the virB mutant, were not considered and are replaced by a solidus (/).

**Fig. 3.** Comparison of MxiE boxes carried by the virulence plasmid pWR100 and the chromosome of strain CP301. Sequences located upstream from the translation start site (ATG) of the various genes were aligned with respect to the region exhibiting similarity with the MxiE box consensus sequence shown at the bottom. Genes ipaHa–e correspond to ORFs SF0722, SF2610, SF1383, SF1880 and SF2022, respectively, of S. flexneri strain CP301 (Jin et al., 2002). For the sake of clarity, SF2202 and SF0887, which are respectively identical to SF2022 and SF2610, are not shown. In each case, nucleotides identical to those present in the MxiE box consensus sequence are shown in upper-case bold characters and nucleotides identical to those in the −10 box consensus sequence are shown in lower-case. The nucleotides corresponding to the transcription start sites determined for ospC1, virA and ipaH9.8 (Mavris et al., 2002b) are underlined. The numbers of nucleotides present between the determined or putative transcription start and the translational start are indicated.

| VP genes | ORF | Sequence | Expression
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ospB</td>
<td>ttttttGTCGCTTTTTTTAAG</td>
<td>Ttctgatatacaagaagc.. TgtcAAtactata - 57 – AGG</td>
<td></td>
</tr>
<tr>
<td>ospC1</td>
<td>agaaaGTACGCTTTTTTTAAG</td>
<td>Tgaacctctcctgctcaaa AaatAtgtatat - 229 – AGT</td>
<td></td>
</tr>
<tr>
<td>ospE1</td>
<td>aaaaaaGTACGCTTTTTTTAAG</td>
<td>Taatccttttttaccgacc.. Ataatgtaa - 38 – AGT</td>
<td></td>
</tr>
<tr>
<td>ospE2</td>
<td>aaaaaaGTACGCTTTTTTTAAG</td>
<td>Taatccttttttaccgacc.. Ataatgtaa - 36 – AGT</td>
<td></td>
</tr>
<tr>
<td>ospF</td>
<td>ataataGTACGCTTTTTTTAAG</td>
<td>Gataaatcaactaatatat aatataaatatca - 20 – AGT</td>
<td></td>
</tr>
<tr>
<td>virA</td>
<td>gatattGTACGCTTTTTTTAAG</td>
<td>Gagaaatctctcctccttt TtatAtgtccga - 56 – AGT</td>
<td></td>
</tr>
<tr>
<td>ipaH4</td>
<td>ctctaGaATGCATTTTTTTAAG</td>
<td>Acctctctccttattgtgc.. Attataaatcaca - 25 – AGT</td>
<td></td>
</tr>
<tr>
<td>ipaH9</td>
<td>aacaccGTACGCTTTTTTTAAG</td>
<td>Ccaataatgtatataatattat AtatAtagatca - 96 – AGT</td>
<td></td>
</tr>
<tr>
<td>ipaH7</td>
<td>aagtGAATGCATTTTTTTAAG</td>
<td>Taatctctccttattgtgc.. Attataaatcaca - 271 – AGT</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chromosomal genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>ipaHa</td>
</tr>
<tr>
<td>ipaHb</td>
</tr>
<tr>
<td>ipaHe</td>
</tr>
<tr>
<td>ipaId</td>
</tr>
<tr>
<td>ipaHe</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MxiE box</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>GTACGCTTTTTTTTAAAG</td>
<td>------ 20 n ------ TtatAtaatatat</td>
</tr>
</tbody>
</table>
carried by the entry region was not affected in mxiE, ipaB and mxiE ipaB mutants, indicating that neither the secretion activity nor MxiE regulates expression of components of the TTS apparatus and translocators. Control of promotor activity by MxiE correlates with the presence of nine MxiE boxes on the VP. MxiE boxes are also present in the 5' region of chromosomal ipaH genes, consistent with the observation that these genes are also controlled by MxiE and IgpC (Mavris et al., 2002b).

Genes controlled by VirB

VirB is necessary for expression of genes of the entry region, as well as those genes that are less expressed in the wild-type strain grown at 37°C. Expression of orf136, phoN2 and virA, which are controlled by VirB, was markedly decreased in this mutant compared with the wild-type strain. In addition, transcription of virB, ipaH, ospD2, ospF, orf13, ospD1, ospC2/3/4, ospC1, orf81, ipaI, orf131b, orf137 and icsA was decreased in this mutant (Fig. 2; ratio virB: w.t. 37°C). Since orf131b is located immediately downstream from spa40, it is probably part of the spa operon (Buchrieser et al., 2000; Sasakawa et al., 1993). The decreased expression of icsA in the virB mutant detected here is not consistent with the initial observation that production of IcsA (VirG) was not affected in a virB mutant, although a 2-fold reduction in the amount of IcsA might not have been detected by immunoblot analysis (Adler et al., 1989). Accordingly, the present analysis suggests that, in addition to genes of the entry region, icsP, phoN2 and virA, the repertoire of VirB-controlled genes includes ospD2, ospF, orf13, ospD1, ospC2/3/4, ospC1, orf81, ipaI, orf137 and icsA.

Expression of phoN2, ospF, ospC1 and virA was decreased in the virB mutant compared with the wild-type strain and was increased in the ipaB mutant in an MxiE-dependent manner (see above). The decreased expression of these genes in the virB mutant is not due to the decreased production of MxiE (whose expression was indeed decreased in the virB mutant), since no differences in expression of these genes were observed between the mxiE mutant and the wild-type strain (Fig. 2; ratio mxiE: w.t. 37°C). This indicates that these genes are expressed in a VirB-dependent, MxiE-independent manner under conditions of non-secretion (i.e. in the wild-type and mxiE strains) and that their expression is increased in an MxiE-dependent manner under conditions of secretion (i.e. in the ipaB mutant).

Genes controlled by VirF

To investigate which genes are controlled by VirF, we compared the transcription profile of a virF mutant with those of the wild-type and virB mutants, suggesting that Orf81 was produced by the wild-type strain and was potentially active during growth in broth. To determine which genes might be under the control of Orf81, we compared the expression profiles of the wild-type strain and an orf81 mutant; however, the differences were approximately 1-5-fold. This suggests that inactivation
of orf81 did not affect strongly, either positively or negatively, the expression of any VP gene. Since the activity of transcriptional activators of the AraC family is often modulated in response to external signals, it is possible that, in the wild-type strain, Orf81 was not active under the growth conditions used in the present study. Alternatively, orf81 might be a remnant of a regulatory circuit that is no longer used by S. flexneri. This hypothesis is consistent with the observation that orf81 is inactivated by an insertion sequence in an S. flexneri strain of serotype 2a (Jin et al., 2002) and could not be amplified by PCR from a strain of serotype 6 (Lan et al., 2003).

Concluding remarks

The comparison of transcription profiles of most VP genes in the wild-type strain and in mutants for each of the known regulatory genes carried by the VP presented here provides a general view of the repertoires of genes that are under the control of each regulator. Data are remarkably consistent with previous observations, obtained using a wide variety of methods, concerning particular genes that had been shown to be controlled by VirF, VirB, MxiE and the activity of secretion. These results also extend the list of genes that are controlled by VirB and MxiE and give information on genes that are not affected in the various mutants and, hence, on the specificity of each regulatory circuit. Differences in gene expression leading to a ratio >0.5 and <2 were considered cautiously, although this threshold is arbitrary. Within these limits of interpretation, it seems that only virB is under the direct control of VirF and that no genes are controlled by Orf81 and MxiE in the wild-type strain growing in broth. In addition, it appears that a number of genes known to be involved in virulence, including sepA, virK and msbB2, are not controlled by any of the VP-encoded regulators. Among the other genes that are not controlled by VP-encoded regulators or by temperature are genes whose products are involved in plasmid replication, maintenance or transfer, such as parB, stbA, ccdAB, mvpA, trbH, finO, repA, orf159b and orf163.

Numerous genes located outside the entry region appear to be under the control of VirB (and indirectly of VirF). Many of these genes exhibit a low G+C content that is similar to genes of the entry region, suggesting that they might have the same origin as and possibly be part of the TTS system. Such genes include ospC1, ospC2/3/4, ospD1, ospD2, ospF and virA, the products of which are indeed substrates of the TTS apparatus, since they were identified by N-terminal sequencing of proteins secreted by a secretion-deregulated strain (or exhibited sequence similarity with a secreted protein) (Buchrieser et al., 2000). Other VirB-controlled genes that exhibit a low G+C content include orf13, orf137, orf81 and ipaH, the latter being located within the entry region, as well as icsP. The two VirB-controlled genes that exhibit a G+C content higher than genes of the TTS system are phoN2 and icsA. Control of VirB over expression of these genes might be a consequence of their localization on the VP, as phoN2 is located downstream from and transcribed from the same promoter as ospB (encoding a TTS substrate) and icsA is located upstream from and is transcribed in the opposite orientation to virA (also encoding a TTS substrate).

The mechanism by which VirB activates transcription is still not completely understood. In vivo binding and oligomerization of VirB on target DNA has been reported; as well as DNA binding of the purified protein (Beloin et al., 2002; McKenna et al., 2003). A short DNA sequence (5′-ATTTCAT-3′) located upstream from the icsB promoter has been shown to be required for activity of this promoter in vivo and for binding of a GST–VirB protein in vitro (Taniya et al., 2003). However, there are 56 occurrences of this sequence on pWR100, including within the numerous insertion sequences that are scattered along the plasmid, and a similar motif with one mismatch is present more than 1000 times on the plasmid, suggesting that this sequence alone does not define the VirB target. Although not all genes with a low G+C content are controlled by VirB, promoter regions of genes that were less expressed in the virB mutant exhibit a low G+C content. Accordingly, the possible constraint that this G+C content might impose on the local conformation of the DNA (e.g. bending) might be part of the VirB recognition motif. It has been proposed that activation of VirB-controlled promoters might involve displacement of H-NS by VirB at these promoters (Beloin & Dorman, 2003). The present analysis, by further defining the repertoire of genes that are (or are not) under the control of VirB, might help future studies aimed at determining its mode of action.

The repertoire of genes that were induced under conditions of deregulated secretion includes 13 genes. Activation of all these genes in the ipaB mutant was dependent on MxiE and their regulated expression correlates with the presence of an MxiE box in the 5′ region of each regulated gene (or operon). Some of these genes, such as ospF, ospC1, ospB, phoN2 and virA, were expressed in a VirB-dependent manner under conditions of non-secretion and induced in an MxiE-dependent manner under conditions of secretion. Other genes, such as ipaH genes, ospG and ospE1/2, were not expressed (or expressed at a low level that was independent of VirB) under conditions of non-secretion and were activated in an MxiE-dependent manner under conditions of secretion. Accordingly, substrates of the TTS apparatus might be classified into three categories on the basis of their expression profiles (Fig. 4): (i) those that are controlled by VirB (i.e. their transcription was decreased in the virB mutant and was not increased in the ipaB mutant), (ii) those that are controlled by MxiE (i.e. their expression was barely detectable in the wild-type strain, was not decreased in the virB mutant and was increased in the ipaB mutant) and (iii) those that are controlled by both VirB and MxiE (i.e. their expression was decreased in the virB mutant and was increased in the ipaB mutant). In addition, it appears that homologous effectors do not
all belong to the same expression class. For example, ospD1 and ospD2 are controlled by VirB, whereas ospD3 is controlled by MxiE. Likewise, expression of ospC2/3/4 is controlled by VirB whereas that of ospC1 is controlled by both VirB and MxiE. Since effectors encoded by genes controlled by VirB are expressed whether or not the TTS apparatus is active, these proteins are produced and presumably stored in the cytoplasm prior to activation of the TTS apparatus and, therefore, can transit through the TTS apparatus immediately upon contact of bacteria with eukaryotic cells. In contrast, effectors encoded by genes that are regulated by MxiE are produced only after activation of the TTS apparatus and are likely to correspond to a second wave of effectors that are involved at a later stage or, possibly, in different cells. Effectors that are controlled by both VirB and MxiE are produced both before and after the TTS is activated and would be involved in both situations. Accordingly, the differential regulation of genes encoding TTS effectors is evidence that different effectors might be required at different times following contact of bacteria with host cells. The function of each of these effectors, initially identified as substrates of the TTS apparatus, within eukaryotic cells and the cell type(s) in which these effectors have an important role remain to be identified.

Regulation of the TTS apparatus activity by external signals and control of the activity of some promoters by the TTS apparatus activity is not specific to the S. flexneri TTS system. In pathogenic Salmonella enterica, the pathogenicity island 1 (SPI-1) encodes a TTS system very similar to the one encoded by the entry region of the S. flexneri VP. SPI-1 encodes several transcriptional activators, including HigA (a member of the OmpR family) and InvF (a member of the AraC family). HigA regulates transcription of genes encoding components of the TTS apparatus, whereas InvF controls transcription of genes encoding effectors located both within and outside SPI-1 (Eichelberg & Galan, 1999). The activity of InvF is dependent on the co-activator SicA, which is homologous to IpgC (Darwin & Miller, 2000, 2001). Modulation of the activity of InvF by the activity of secretion has not yet been documented, since there are no conditions known to modulate the activity of the TTS apparatus in Salmonella enterica. Nevertheless, the Salmonella enterica InvF/SicA and S. flexneri MxiE/IpgC systems appear as clear variations on the same theme, despite the fact that InvF, but not MxiE, controls transcription of the operon encoding its co-activator and translocators. In pathogenic Yersinia species, a feedback regulation on transcription of yop genes by the TTS apparatus activity has been extensively studied (Michiels et al., 1991; Pettersson et al., 1996). In contrast to the Shigella and Salmonella systems, the VP-encoded TTS system of Yersinia species contains only one transcriptional activator, VirF (also a member of the AraC family) (Cornelis et al., 1998). Furthermore, in Yersinia species, expression of genes encoding both components of the TTS apparatus and effectors is increased under conditions of active secretion (Lambert de Derouvoit et al., 1992; Rimpilainen et al., 1992). The mechanism by which the TTS apparatus activity is transmitted to VirF has not yet been elucidated, but also involves TTS chaperones (Francis et al., 2001; Wulff-Strobel et al., 2002). Control of gene transcription by the TTS apparatus activity is likely to be a general feature of TTS systems and is strong evidence that a temporal regulation of expression of effectors is a key element in the function of TTS systems (Francis et al., 2002; Miller, 2002). Differences between regulatory circuits operating in each system might reflect the specific constraints imposed by the strategy used by each pathogen to infect its host(s).

ACKNOWLEDGEMENTS

We are pleased to acknowledge Giulia Nigro for help in strain construction, Elsabeth Couve and Jean-Yves Coppée for spotting PCR fragments at the Génopole of the Pasteur Institute, Lionel Frangeul for sequence analysis and John Rohde for critical reading of the manuscript. This work was supported in part by grants from the Italian Ministero dell’Istruzione, Università e Ricerca (COFIN 2002).

REFERENCES


Shigella flexneri: two-domain structure and proteolytic activity. Microbiology 144, 1815–1822.


