The flagellin \(N\)-methylase gene \(fliB\) and an adjacent serovar-specific IS200 element in \(Salmonella\) \(typhimurium\)

André P. Burnens,1 John Stanley,2 Ragna Sack,3 Peter Hunziker,3 Isabelle Brodard1 and Jacques Nicolet1

Author for correspondence: André P. Burnens. Tel: +41 31 631 24 85. Fax: +41 31 631 26 34.

e-mail: BURNENS@VBI.UNIBE.CH

1 Institute for Veterinary Bacteriology, University of Berne, CH-3012 Berne, Switzerland
2 Molecular Biology Unit, Virus Reference Division, Central Public Health Laboratory, London NW9 5HT, UK
3 Institute for Biochemistry, University of Zurich, CH-8057 Zurich, Switzerland

INTRODUCTION

The insertion sequence IS200 was discovered as a 708 bp insertion in the \(hisD\) gene of a strongly polar mutant \(Salmonella\) \(typhimurium\) LT2 \(hisD984\) (Lam & Roth, 1983). Its nucleotide sequence exhibited certain features which were unusual for insertion sequences. It was the smallest insertion element reported to date, lacked inverted repeats and failed to generate duplications in the target sequence upon insertion (Gibert et al., 1991). In a subsequent study of another \(S.\) typhimurium mutant, insertion of IS200 into the \(gpt\) gene was found. When the regions of the \(gpt\) gene flanking the ends of the element were sequenced, the absence of terminal inverted repeats in IS200 was again detected (O'Reilly et al., 1990).

IS200 was initially considered to be \(Salmonella\)-specific, but it was subsequently shown that isofoms of the element were present in certain rare clones of \(Escherichia\) \(coli\) (Bisercic & Ochman, 1993a, b). When the central parts of IS200 elements amplified by PCR from strains of \(E.\) coli and \(Salmonella\) were sequenced, their analysis showed that IS200 must have been present in the common ancestor of the two species, and that it had not been transferred between them since their divergence. This was in contrast to the IS1 element, which was found to be isogenic in both species, reflecting recent horizontal transfer (Bisercic & Ochman, 1993a).

Random mutagenesis with Tn10 followed by genetic mapping of those transposons which were linked to IS200 insertion sites showed that six IS200 copies in \(S.\) typhimurium LT2 were located at evenly spaced sites around the chromosome (Sanderson et al., 1993). Molecular epidemiological studies showed that three of these were specific to all isolates of this serovar from cases of human and animal disease (Stanley et al., 1993). A physical genetic map of the LT2 chromosome constructed by macrorestriction placed one of these elements, termed copy IV, close to the \(fliA\) gene which encodes a sigma factor (Sanderson et al., 1993). This gene belongs to one of three groups of chromosomal
genes (flg, flb and fli) within which the order of genes is highly conserved between S. typhimurium and E. coli, and which encode flagellar biosynthesis and motility (Kawagishi et al., 1992). Unusually, a genetic determinant nml, which mediates post-translational methylation of lysine residues in the flagellin of salmonellae, was mapped to the fli gene cluster but is absent from E. coli. The nml-encoded methylase is the only one in either prokaryotes or eukaryotes acting on e-amino groups of lysine. At the same locus in Salmonella muenchen are found two genes fliUV, essential for motility of E. coli mutants transformed with cloned flagellar genes of salmonellae (Doll & Frankel, 1993). Nothing is known about the methylation status of flagellin or the presence of an nml homologue in S. muenchen.

We wished to study the evolutionary significance and molecular genetic features of the IS200 locus described above, and also to characterize the linked nml region of S. typhimurium. We addressed these aims in the first instance by cloning and sequencing the region between fliC and fliA of S. typhimurium LT2, which contains IS200. In so doing, we sequenced and determined the phenotype associated with the upstream ORF, equivalent to the nml determinant mapped to the region between fliC and fliA by phage transduction (Konno et al., 1976) and renamed fliB accordingly (Iino et al., 1988).

METHODS

Bacterial strains. The wild-type strain S. typhimurium LT2 was obtained from the American Type Culture Collection (Rockville, MD, USA; ATCC 23564). S. muenchen KOP54, Kauflmann’s original isolate, was from the International Salmonella Centre, Copenhagen, Denmark, and is the ancestor of strain ATCC 8388. The SARB reference collection of Salmonella isolates (Boyd et al., 1993) was obtained from K. E. Sanderson, Salmonella Genetic Stock Centre, Alberta, Canada. The remaining Salmonella strains were clinical isolates from the routine investigative programme of the Swiss Reference Laboratory for Foodborne Diseases. E. coli strain JA11, carrying a kanamycin resistance cassette in the fliC gene was a gift from R. Belas (Belas & Flaherty, 1994). Fresh clinical isolates were kept as stab cultures at room temperature and reference strains were kept frozen at −80 °C in nutrient broth containing 12% sucrose.

Nucleic acid techniques. Total genomic DNA was extracted by a guanidium thiocyanate micromethod (Pitche et al., 1989). Restriction enzymes (Boehringer Mannheim) were used according to the manufacturer’s instructions. Electrophoresis of DNA, ligation and transformation were carried out by standard protocols (Sambrook et al., 1989), using plasmid vector pBluescript KS(−) and E. coli strain XL1-Blue (Stratagene). For nonradioactive hybridization, DNA was blotted onto nylon membranes (Boehringer Mannheim) by alkaline transfer in 0.4 M NaOH using vacuum-blotting (Pharmacia LKB Vacugene 2016). Blots were probed with labelled PCR amplicons as described below. CSPD substrate (Boehringer Mannheim) was used for luminescent detection of the alkaline phosphatase reaction and the blots were exposed to Fuji RX films at room temperature for approximately 45 min. For reprobing, blots were stripped of probes and antibodies by washing the membranes twice in distilled water with 0.1% SDS in a boiling water bath for 5 min. Nucleotide sequences were determined with an Applied Biosystems AB373 DNA Sequenator using the Taq Dye Deoxy Terminator Cycle Sequencing kit (Applied Biosystems/Perkin Elmer), using T3 and T7 primers on appropriate subclones according to the strategy outlined in Fig. 1(a).

PCR amplification of flic, fliA and IS200 probes. Digoxigenin-labelled probes were prepared by PCR as follows. The reaction contained, in volumes of 100 μl: 100 ng genomic DNA of S. typhimurium LT2 as template, 0.25 μM primers, 170 μM dNTP, 50 μM digoxigenin-dUTP (Boehringer Mannheim) and 2.5 U Taq DNA polymerase (Boehringer Mannheim). The reaction buffer contained 10 mM Tris/HCl pH 8.3, 1.5 mM MgCl2, 50 mM KCl, 0.005% Tween 20 (Calbiochem), 0.005% NP-40 (Calbiochem). Primer sequences for flic, 5′-ATG GCA CAA GTC ATT AAT AC-3′ (forward) and 5′-TTA ACG CAG TAA AGA GAG GA-3′ (reverse) and reaction conditions were as described by Kilger & Grimont (1993). Primer sequences for fliA, based on the sequence of fliA of S. typhimurium (EMBL X52624; Ohnishi et al., 1990), were 5′-TGG AGC ATG TGC TAC AAG-3′ (forward) and 5′-TTT GAT GGC CGT ACT ATG-3′ (reverse) and yielded a 569 bp amplicon. Samples were subjected to 35 cycles of amplification with incubation for 1 min each at a denaturation temperature of 94 °C, annealing temperature of 54 °C and elongation temperature of 74 °C in a Perkin-Elmer Cetus Thermal Cycler. Primers for IS200, designed to anneal to nucleotides conserved between the two published sequences of IS200, EMBL L25848 (Bisicic & Ochman, 1993a) and EMBL X56834 (Gilbert et al., 1991), were 5′-CCT AAC AGG CGC ATA CGA TC-3′ (reverse) and 5′-ACA CCT TGC CTT CGT CTA GCA ATG-3′ (forward) and yielded a 557 bp amplicon under the conditions described above for fliA. The identities of the amplicons obtained without adding digoxigenin-dUTP were verified by size and restriction products.

Purification of flagellin and amino acid analysis. Flagellin was initially purified from S. typhimurium and from motile transformants of E. coli JA11 carrying cloned S. typhimurium flagellar genes by dissociation at low pH, reassociation and ultracentrifugation (Ilim et al., 1985). Additional final purification was done by ion exchange chromatography on DEAE-Sepharose (Pharmacia) as follows. Flagellin loaded onto the column and extensively washed with 30 mM phosphate buffer pH 5.9 was eluted in the same buffer containing 100 mM NaCl. It was then extensively dialysed against distilled water and lyophilized. Amino acid analysis was performed on a model 420 A/H amino acid analyser (Applied Biosystems) according to the manufacturer’s instructions.

RESULTS

Cloning of the fliA-flic intergenic region

Examination of the sequence of fliA of S. typhimurium (Ohnishi et al., 1990) showed that a 323 bp sequence homologous to IS200 was to be found upstream of the gene. A cloning strategy for the fliA-flic intergenic region was deduced from analysis of the nucleotide sequences of fliA, flic and fliD (Homma et al., 1990; Joys, 1985; Ohnishi et al., 1990). Unique restriction sites for the enzymes BglII and SacI were identified in the restriction maps of fliD and fliA, respectively (Fig. 1a). A fragment of approximately 5 kbp reacted with both flic and fliA probes in genomic Southern blots made by BglII
Fig. 1. Recombinant plasmids, phenotypes and nucleotide sequences. (a) Diagrammatic representation of the flagellar genes and associated IS200 insertion site in S. typhimurium LT2. The fragment of Salmonella DNA cloned in plasmid pIVB20 is represented by the black bar. The lower part shows an enlargement of the 195 kbp KpnI–ClaI subfragment of pIVB20 cloned in pIVB23 and the strategy used for sequencing it. Asterisks denote sequences obtained from the overlapping EcoRI–EcoRI fragment. B, BssHII; Bc, BclI; Bg, BglII; Bl, BglII; C, ClaI; E, EcoRI; Ev, EcoRV; H, HindIII; K, KpnI; Ks, KspI; Sc, Sacl; Sl, SalI. (b) Diagrammatic representation of DNAs subcloned in pBluescript KS(−) which conferred motility when transformed into E. coli JA11. The methylation status of the resulting mature flagellin is indicated on the left. (c) Nucleotide and deduced amino acid sequences of the gene ORF1 (flIB) upstream of the IS200 insertion site in S. typhimurium LT2. The 3' end of the flIC gene and the beginning of the IS200 sequence are shown in boldface and underlined.
Table 1. Amino acid composition and content of methylated lysine residues (Meth-Lys) of flagellin isolated from S. typhimurium ATCC 23564 and E. coli JA11 carrying flagellar genes of S. typhimurium

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Theoretical percentage composition</th>
<th>S. typhimurium phase 1, H'</th>
<th>E. coli JA11(pIVB20)</th>
<th>E. coli JA11(pIVB94)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>12.2</td>
<td>13.7</td>
<td>11.2</td>
<td>12.3</td>
</tr>
<tr>
<td>Arg</td>
<td>2.9</td>
<td>2.6</td>
<td>2.0</td>
<td>2.3</td>
</tr>
<tr>
<td>Asx</td>
<td>16.2</td>
<td>14.6</td>
<td>13.8</td>
<td>13.9</td>
</tr>
<tr>
<td>Glx</td>
<td>9.6</td>
<td>10.9</td>
<td>13.0</td>
<td>11.8</td>
</tr>
<tr>
<td>Gly</td>
<td>8.0</td>
<td>10.4</td>
<td>13.7</td>
<td>14.5</td>
</tr>
<tr>
<td>His</td>
<td>0.2</td>
<td>0.5</td>
<td>1.0</td>
<td>0.6</td>
</tr>
<tr>
<td>Ile</td>
<td>5.3</td>
<td>5.0</td>
<td>4.2</td>
<td>4.1</td>
</tr>
<tr>
<td>Leu</td>
<td>8.2</td>
<td>8.2</td>
<td>6.9</td>
<td>6.5</td>
</tr>
<tr>
<td>Met</td>
<td>0.8</td>
<td>0.4</td>
<td>ND</td>
<td>0.8</td>
</tr>
<tr>
<td>Phe</td>
<td>1.2</td>
<td>1.7</td>
<td>1.8</td>
<td>1.6</td>
</tr>
<tr>
<td>Pro</td>
<td>1.0</td>
<td>1.5</td>
<td>2.6</td>
<td>2.6</td>
</tr>
<tr>
<td>Ser</td>
<td>8.1</td>
<td>6.4</td>
<td>10.0</td>
<td>7.5</td>
</tr>
<tr>
<td>Thr</td>
<td>11.4</td>
<td>10.9</td>
<td>8.4</td>
<td>8.6</td>
</tr>
<tr>
<td>Tyr</td>
<td>2.7</td>
<td>0.4</td>
<td>0.5</td>
<td>1.0</td>
</tr>
<tr>
<td>Val</td>
<td>6.3</td>
<td>6.9</td>
<td>6.1</td>
<td>6.1</td>
</tr>
<tr>
<td>Lys</td>
<td>3.4</td>
<td>3.6</td>
<td>3.4</td>
<td>5.8</td>
</tr>
<tr>
<td>Meth-Lys</td>
<td>2.5</td>
<td>2.2</td>
<td>1.4</td>
<td>0.0</td>
</tr>
</tbody>
</table>

ND: Not determined.

and SacI double digestion (data not shown). Genomic DNA fragments (BglII–SacI digest) of this size were extracted from agarose gels, and cloned into the plasmid pBluescript KS(—). Recombinant plasmids were then screened by hybridization with the fliC probe. One of these, designated pIVB20, carried an insert of 4.8 kb. Upon further analysis, this was shown by restriction mapping and hybridization with fliC, fliA and IS200 probes (Fig. 1a) to represent the desired clone. A predicted KpnI site within the coding sequence of fliC (Joys, 1985) and a predicted EcoRI site within the coding sequence of fliA (Ohnishi et al., 1990) were mapped.

Sequence and conserved features of an ORF between fliC and the IS200 site

The sequence (1949 bp) of the KpnI–ClaI fragment and that of the overlapping EcoRI–EcoRI fragment (574 bp) were determined from subcloned fragments, according to the strategy delineated in Fig. 1(a). These sequences were found to contain two ORFs. The first of these was a 1203 bp ORF termed ORF1 coding for a protein of 401 amino acid residues with a predicted molecular mass of 45 kDa, commencing 79 nucleotides downstream of the 3' end of the fliC gene and transcribed in the same direction as fliC. The second, termed ORF2, was due to IS200. No promoter sequences were found upstream of ORF1, the nucleotide sequence, amino acid translation and significant regulatory features of which are depicted in Fig. 1(c). As shown in Fig. 1(a), the start codon of ORF1 coincided with the start of the S. muenchen fliU gene sequence, while its stop codon coincided with the end of S. muenchen fliV (Doll & Frankel, 1993). The only significant nucleotide or protein level homologies found between ORF1 and sequences in the databases were with fliU and fliV of S. muenchen, with a 406 bp fragment in the Salmonella typhi fliC sequence (Frankel et al., 1989) and with a putative fliV gene from Yersinia enterocolitica (Iriarte et al., 1995). A probe consisting of the ORF1 sequence between nucleotides 923 and 1332 reacted in colony blots with 60 Salmonella clinical isolates belonging to 19 serovars frequently isolated from cases of human disease (data not shown), showing this gene to be conserved in the genus.

ORF1 downstream of fliC encodes a flagellin methylase

E. coli strain JA11 is non-motile and does not express flagellin due to the insertion of a kanamycin resistance cassette in the fliC gene. JA11 recipients were rendered motile by transformation with pIVB20, which contains the 4.8 kb cloned insert of S. typhimurium DNA described above. Flagellin was found to be expressed and methylated; the amount of methylated lysine residues closely approximated values observed for flagellin purified from S. typhimurium itself (Table 1). The central 2.9 kbp EcoRI–SalI fragment (fliC gene and
part of ORF1) of pIVB20 was subcloned into pBluescript KS(−) to yield pIVB94 (Fig. 1b). Although pIVB94 carried the whole S. typhimurium fliC gene, including the promoter region and 121 bp of the 3′ end of flID, 225 bp at the 5′ end of ORF1 were deleted (i.e. up to the SalI site in Fig. 1b). When pIVB94 was transformed into E. coli JA11, transformants expressed flagellin and were motile, but the methylation of flagellin was completely abolished (Table 1). In a third recombinant plasmid, pIVB95, constructed by deleting a KspI fragment from pIVB94, 402 bp of DNA 5′ to the stop codon of fliC was still present. E. coli JA11 transformed with pIVB95 expressed an unmethylated Salmonella flagellin and was motile (Fig. 1b). Thus the cloned fliC gene of S. typhimurium with 378 bp of upstream and 402 bp of downstream flanking sequences was sufficient to render E. coli JA11 transformants motile, but the whole of ORF1 was needed for methylation of the expressed flagellin (Fig. 1b).

**Comparison of an occupied and unoccupied IS200 insertion site**

An IS200 insertion site starts 38 nucleotides downstream of the stop codon of the methylase gene (Fig. 1c). The central part of the IS200 sequence constituted a single ORF (termed ORF2 above), transcribed in the same direction as the preceding gene. The predicted polypeptide encoded by ORF2 showed significant amino acid sequence similarity to elements from Yersinia pestis (Simonet et al., 1996), Streptococcus pneumoniae (Berry et al., 1994), Clostridium perfringens (Brynestad et al., 1994), Vibrio cholerae (Bik et al., 1996), and even an archaeobacterial phage (Schnabel et al., 1984). The nucleotide sequence of the corresponding ORF1 (fliB)—fliA intergenic region from S. muenchen and two atypical S. typhimurium isolates lacking this copy of IS200 (see below) were determined by direct PCR sequencing. The sequences flanking the unoccupied IS200 insertion site are shown in Fig. 2, together with the data on the IS200 insertion sites described for the hisD984 and gpt81 mutations. The IS200 element in the fliB—fliA intergenic region lacked terminal inverted repeats but the 5′ end of the element had two additional base pairs not previously reported. The original sequence for IS200 was 708 bp in length but since an additional ‘C’ is found at nucleotide 1524, this would bring the length of the element to 711 bp. The additional ‘C’ found at this position has also been reported to be present in the IS200 elements of two Salmonella strains from the SARA reference collection (Biseric & Ochman, 1993a), and therefore probably represents the ancestral sequence of the element. Given the available evidence it cannot be decided whether the occupied or unoccupied insertion site is ancestral. Based on the data presented in Fig. 2, a possible target sequence requirement for IS200 insertion would be a stretch of T(A) residues as confirmed by analysis of five further IS200 insertions in Salmonella enteritidis and S. typhi (A. P. Burnens, unpublished).

The IS200 insertion site under study was located on a 2.4-kbp PstI fragment which hybridized with IS200, ORF1 (Fig. 3) and fliA (data not shown) in all strains of S. typhimurium studied except two. In these rare strains of S. typhimurium (environmental isolates from Malaysia and Sri Lanka) no other fragment hybridized to IS200, fliB and fliA, and the IS200 insertion in the ORF1—fliA intergenic region was absent. These strains also lacked another (1.9-kbp) conserved fragment hybridizing with IS200 (Fig. 3). To further substantiate that the insertion of IS200 between ORF1 and fliA is generally specific to strains of serovar S. typhimurium, PCR amplification of the intergenic region was carried out for a collection of strains belonging to diverse Salmonella serovars. The amplified fragment was 424 bp long for salmonellae lacking this IS200 insertion site, and 1135 bp long for strains carrying IS200 between ORF1 and fliA. In a set of 108 salmonellae belonging to 38 important serovars, 21 of 23 isolates of S. typhimurium yielded the 1135 bp amplicon, the two unusual strains lacking the IS200 copy in the 2.4-kbp PstI fragment (see above) giving the 424 bp amplicon instead. Of the remaining 85 salmonellae belonging to other serovars [S. choleraesuis (6 strains), S. derby (3), S. dublin (3), S. enteritidis (4), S. gallinarum (3), S. heidelberg (8), S. java (3), S. muenchen (6), S. newport (8), S. paratyphi C (4), two strains each of S. infantis, S. javiana, S. miami, S. montevideo, S. paratyphi B, S. saintpaul, S. senftenberg, S. stanley, S. typhi, and S. wien, one strain each of S. agora, S. anatum, S. brandenburg, S. duisburg, S. emek, S. haifa, S. indiana, S. limete, S. manhattan, S. panama, S. paratyphi A, S. reading, S. rubislaw, S. schwazergrund, S. sendai, S. stanleyville, and S. thompson], 84 yielded the 424 bp
amplicon. One isolate of S. newport (strain SARB 37 of Boyd et al., 1993) failed to give a PCR product. Therefore, with the exception of the two strains described above, all isolates of S. typhimurium carried the IS200 locus under study, and none of 85 isolates of 37 other Salmonella serovars contained IS200 inserted between ORF1 and flIA.

DISCUSSION

Of the four genes mapped to the flagellin gene cluster, fliD is the filament capping protein, flIC encodes flagellin itself, fliB an enzyme which post-translationally modifies flagellin and fliA encodes a sigma factor (Homma et al., 1990). Although fliB was recently cloned (Kawagishi et al., 1992), no further analysis or sequence information was provided. We have here presented two lines of evidence that the ORF downstream of flIC is the gene fliB. Firstly, introduction of the recombinant plasmid pIVB20 into E. coli led to motile transformants which expressed a methylated flagellin. This methylation was found to be very similar to that observed in S. typhimurium. Small observed differences may be explained by the presence of some phase 2 flagellar antigen of S. typhimurium in the flagellin preparation. Secondly, deletion of 225 bp from the 5' end of the gene completely abolished methylation of Salmonella flagellin expressed from a recombinant plasmid in E. coli. Since no potential promoter sequences were found 3' of the fliB gene, the gene appears to be co-transcribed with flIC. Whereas the evidence presented above suggests that in S. typhimurium the fliB gene encodes a single polypeptide, two genes (flIU and flIV) have been postulated to occupy the same position within the flagellar gene cluster in S. muenchen (Doll & Frankel, 1993). Comparison of the nucleotide sequences of flIU and flIV of S. muenchen with the sequence determined for fliB of S. typhimurium in this paper suggests that a sequencing error in the flIU sequence (missing 'C' at position 499 in EMBL L06320 in a stretch of eight 'A' residues) led to the erroneous conclusion that the flIU ORF ends at residue number 569. It is difficult to reconcile the findings of Doll & Frankel (1993), showing that sequences downstream of the S. muenchen fliC gene are essential for expression of the cloned flIC gene in E. coli, with the findings reported in the present paper. However, we have successfully expressed second phase as well as rough phase Salmonella flagellar genes, completely lacking any downstream fliB-like sequences, in E. coli with the approach outlined above (A. P. Burnens, unpublished).

A plasmid containing the entire flIC gene and only 402 bp of DNA downstream of flIC conferred motility on an E. coli strain (JA11) with a flIC gene inactivated by insertion of a kanamycin cassette. This suggests that in E. coli the S. typhimurium flagellin gene functions normally without lysine methylation. This is in contrast to S. muenchen, in which fliB gene function was found essential for motility (Doll & Frankel, 1993). This finding is consistent with early phage transduction experiments which indicated that the methylation status of flagellin did not influence its functional status in salmonellae (Joys & Kim, 1979). All salmonellae tested in the present study by colony blot showed hybridization to a probe for the fliB gene (data not shown). Little is known about the distribution of methylated flagellins among salmonellae, except for a limited study of the distribution of the methylase gene (formerly designated nml for lysine N-methylase) by phage transduction experiments (Stocker et al., 1961). The available evidence suggests that all lysine residues which are surface-exposed in the native flagellum are methylated (Kanto et al., 1991). With respect to recent efforts to express foreign antigens as part of the mature flagellum with the aim of producing live recombinant vaccines (Newton et al., 1989, 1995; Verma et al., 1995), we note that methylation could affect the antigenicity of these foreign epitopes, modifying or abolishing immune responses directed against them. Results presented above would suggest the use of Salmonella strains devoid of the

**Fig. 3.** Southern blots of S. typhimurium DNA restricted with *PstI* and hybridized with IS200 (a) and fliB (b) probes. Lanes: 1, S. typhimurium strain LT2; 2–12, clinical isolates 230-94 (lane 2), 622-94 (lane 3), 3470-94 (lane 4), 84-94 (lane 5), 231-94 (lane 6), 225-94 (lane 7), 179-94 (lane 8), 182-94 (lane 9), 4268-90 (lane 10), 5483-94 (lane 11) and 168-94 (lane 12). (a) Conserved *PstI* fragments of 1.9 kbp, 2.4 kbp and 4.8 kbp hybridizing with an IS200 probe in the majority of isolates are marked with arrows. (b) Blot shown in (a) rehybridized with a fliB probe. Only the 2.4 kbp fragment carrying IS200 also hybridized with this probe, indicating insertion of IS200 near fliB. The 2.4 kbp fragment in atypical strains 168-94 and 5483-94 did not hybridize.
methylase gene fliB or deleting the fliB gene in the construction of appropriate vaccine strains.

Post-translational methylation of surface appendages of pathogenic bacteria may be a key step in the macro-molecular assembly and stabilization of these structures. For example, the methylation of fimbriae in *Pseudomonas aeruginosa* (Strom et al., 1993) is due to the pilD-encoded enzyme which N-methylates the N-terminal phenylalanine residue in mature pilin, and is the only bacterial N-methylase to have been sequenced so far. pilD does not show any sequence homologies with fliB (ORF1). fliB is the only prokaryotic N-methylase to act on ε-aminogroups of lysine to have been described to date.

IS200 is the smallest bacterial insertion sequence known and exhibits unusual features such as a lack of terminal inverted repeats, an apparently low rate of transposition and mutagenic activity insertions (Casadesus & Roth, 1989), and unusual stability which allows it to be used as a marker in molecular epidemiological studies (Baquar et al., 1993, 1994b). A question which follows is whether IS200 shows any target specificity, i.e. whether its insertion sites have any sequence features in common. Since the only two IS200 loci to have been characterized were from insertion mutants, it was of considerable interest to clone and analyse a wild-type IS200 insertion locus from a non-mutated strain. The locus cloned and sequenced in the present report is inserted into a stretch of noncoding DNA between the flagellar genes fliB and fliA. IS200 insertion into noncoding DNA may explain the observed rarity of mutant phenotypes due to IS200 (Casadesus & Roth, 1989), despite the fact that some salmonellae carry a large number of IS200 insertions (Threlfall et al., 1993; Stanley et al., 1994). In ongoing work on *S. typhi*, the insertion of IS200 in intergenic regions has been found to be the rule.

The present study confirms the lack of terminal inverted repeats in IS200. However, evidence from the sequencing data is consistent with the presence of two additional thymidine residues at the 5' end of the element. Our data show that the central part of IS200 has a single ORF, consistent with the data of Bisericic & Ochman (1993a) rather than earlier work (Gibert et al., 1991). Significant similarities with this ORF were found at the protein level in divergent groups of bacteria of medical importance like *Y. pestis* (Simonet et al., 1996), *C. perfringens* (Brennstad et al., 1994), and *Strep. pneumoniae* (Berty et al., 1994).

Profiles of IS200 insertions in salmonellae have been shown to be effective molecular markers for the epidemiological analysis of clinical isolates associated with gastroenteritis. Several studies (Stanley et al., 1993, 1994, 1995; Baquar et al., 1994a; Ezzquerra et al., 1993) have shown that the distribution of IS200 bands (genomic restriction fragments hybridizing with an internal probe for the element) may be a key marker in elucidating the short-term evolution of salmonellae. The analysis of a serovar-specific IS200 insertion locus in the present study provides further evidence for this model. The locus is conserved among almost all isolates of *S. typhimurium* from various sources, and is absent from other *Salmonella* serovars. It is noteworthy that the two *S. typhimurium* strains which do not carry this IS200 insertion also lack a further conserved IS200 band (Fig. 3) previously found to be specific for the serovar (Stanley et al., 1993). This is consistent with them being distinct and rare clonal lines in a polyphyletic serovar background. Beltran et al. (1991) used multilocus enzyme electrophoresis, MLEE, to demonstrate the occurrence of different chromosomal genotypes within the same *Salmonella* serovar, a phenomenon explicable by horizontal transfer of, and recombination at, the flagellar loci (Li et al., 1994; Smith et al., 1990). The two atypical *S. typhimurium* strains in the present study were both from Asia, while the SARA reference collection of *S. typhimurium* isolates (Beltran et al., 1991) contains two strains from Asia, one of which (Tm10) has an unique MLEE genotype. However, none of the *S. typhimurium* strains in the SARB collection, representing all the principal MLEE genotypes of the serovar, failed to carry the IS200 locus between fliA and fliB. Finally, none of the MLEE genotypes of *S. heidelberg* in the SARB collection, which closely resemble *S. typhimurium*, carried this IS200 site, substantiating its exclusive occurrence in the main evolutionary lines of *S. typhimurium*. Conserved fragments hybridizing with IS200 on Southern blots therefore correspond to conserved IS200 insertions, a finding which underscores the phylogenetic importance of IS200 distribution in salmonellae. Furthermore, the exclusive occurrence in *S. typhimurium* of the insertion described here confirms the evolutionary stability of such markers.

**ACKNOWLEDGEMENTS**

The study was supported by a grant from the Swiss National Science Foundation (32-39451.93) to A.P.B. and a British-Swiss Joint Research Programme grant (38BC-032852) to A.P.B. and J.S. The generous gift of *E. coli* JA11 by R. Belas (Belas & Flaherty, 1994) is gratefully acknowledged. The National Center for Biotechnology Information, Washington, DC, is acknowledged for use of its BLAST program.

**REFERENCES**


Stanley, J., Baquar, N. & Threlfall, E. J. (1993). Genotypes and phylogenetic relationships of Salmonella typhimurium are defined by molecular fingerprinting of IS200 and 16S rRNA loci. J Gen Microbiol 139, 1133–1140.


Received 22 July 1996; revised 9 December 1996; accepted 11 December 1996.