Selective Amplification of Abequose and Paratose Synthase Genes (rfb) by Polymerase Chain Reaction for Identification of Salmonella Major Serogroups (A, B, C2, and D)

JOHN M. C. LUK,1* URIRAT KONGMUANG,1 PETER R. REEVES,2 AND ALF A. LINDBERG3‡

Department of Clinical Bacteriology, Karolinska Institute, Huddinge Hospital F82, S-141 86 Huddinge, Sweden,1 and Department of Microbiology, University of Sydney, Sydney, New South Wales 2006, Australia2

Received 4 March 1993/Accepted 24 May 1993

Many parts of the Salmonella rfb gene clusters which are responsible for biosynthesis of the oligosaccharide-repeating units of the O-antigen lipopolysaccharide have recently been cloned and sequenced. On the basis of this knowledge, three sets of nucleotide primers were selected to target defined regions of the abequose and paratose synthase genes: rfsB of Salmonella serogroup B, rfbJ of Salmonella serogroup C2, and rfbS of Salmonella serogroup D (also present in serogroup A). For good differentiation among these major serogroups, the primers were designed not only to give precise specificity in priming but also to give DNA products with different sizes in polymerase chain reactions (product sizes, ~720 bp for both serogroups A and D, ~820 bp for serogroup C2, and ~882 bp for serogroup B). In a polymerase chain reaction assay utilizing these rfb-specific primers, all of the 40 salmonellae belonging to serogroups B, C2, and D plus A were accurately identified among a total of 123 clinical isolates tested (including 55 salmonellae from 36 different serotypes and 68 strains from 10 other members of the family Enterobacteriaceae). No false-positive reactions were detected. The selected rfb gene sequences were proved for the first time to be useful DNA-based markers for identification of and differentiation among Salmonella serogroups A, B, C2, and D.

Lipopolysaccharide (LPS [endotoxin]) is a major virulence determinant of most gram-negative bacteria (13). The chemical structure of Salmonella LPS has been elucidated and found to be composed of three distinctive regions: the O-polysaccharide chain, the core oligosaccharide domain, and lipid A (17). Lipid A, which anchors the entire LPS molecule in the outer membrane of the bacterial cell envelope, is highly conserved in many genera, and the core oligosaccharide has only one known variation in all Salmonella species hitherto studied (30). By contrast, the O-polysaccharide chains, which are formed by polymerization of repeating units that contain three to six sugars in Salmonella serogroups A to E, are antigenically extremely variable. Such antigenic variation in the O-chain structure has formed the basis for serological classification of salmonellae (10, 14). Furthermore, immunochromic studies of Salmonella O antigens provided evidence that the presence of a 3,6-dideoxyhexose, e.g., paratose in serogroup A, abequose in serogroups B and C2, and tyvelose in serogroup D, was of paramount importance for determining the O specificities of these organisms (15). In this respect, we have recently generated a comprehensive panel of monoclonal antibodies specific for the O antigens of Salmonella LPS; antibodies which recognized the dideoxyhexosyl saccharides demonstrated precise specificities, making themselves diagnostically useful for serotyping salmonellae (19, 21).

The rfb gene cluster, located at approximately 42 min and linked to the his operon on the Salmonella enterica chromosome (28), encodes glycosyl synthase and transferase enzymes for the biosynthesis of oligosaccharide-repeating units (22, 23). In recent years, the entire rfb loci of Salmonella serovars Typhimurium, Paratyphi A, Typhi, Muenchen, and Anatum, representing serogroups B, A, D1, C2, and E1, respectively, have been cloned and sequenced (2, 9, 11, 16, 33–35). Many of these genes are common, but the genes for the final steps of dideoxyhexose synthesis provide unique opportunities for creating serogroup-specific probes (Fig. 1A). Salmonella species of both serogroups B and C2 make abequose-containing O antigens; to do this, they must convert CDP-4-keto-3,6-dideoxyglucose to CDP-abequose by use of an enzyme which, for brevity, we call abequose synthase. However, the base sequence of the serogroup B gene, rfbJ (B), is very different from that of the serogroup C2 gene, rfbJ (C2) (44% difference at the nucleotide level and 64% difference at the amino acid level). Serogroups A and D make CDP-paratose in a similar step, and the paratose synthase gene (rfsS) is quite distinctive. The further biosynthetic step, conversion of CDP-paratose to CDP-paratose, is carried out by the rfsE gene (Fig. 1B). However, the restriction maps of the rfb clusters of serogroups A and D appear identical, apart from a triplication of one region which probably has no significant effect on phenotypic expression (16, 32). Sequencing of the rfsE gene showed that the serogroup A form had a single base deletion which converted the fourth codon to an amber stop codon (33), thus accounting for serogroup A strains making paratose and not tyvelose. Scanning the homologous sequences along the related Salmonella rfb gene clusters revealed distinctive molecular fingerprints, of which we have taken advantage to differentiate major Salmonella serogroups in correlation with their serological specificities. Here, we report the use of oligonucleotide primers targeting the abequose and paratose synthase genes and show that their use in a polymerase chain reaction (PCR) amplification assay was selective for salmonellae belonging to serogroups A, B, C2, and D.

In the present paper, we treat all Salmonella strains as...
belonging to one species, *Salmonella enterica*, as suggested by Le Minor and Popoff (12), on the basis that the range of variation within *Salmonella* species, particularly as indicated by total chromosomal DNA hybridization studies (1), is within that found within related species such as *Escherichia coli*. The old species names are retained as serovar names. Despite the high level of sequence homology within the species for most genes, those that encode the highly variable O-antigen genes show substantial sequence variation, and it is this variation which can be used for selective PCR amplification.

**MATERIALS AND METHODS**

**Bacterial cultures, DNA extraction, and oligonucleotide primers.** Six 18-mer oligonucleotides of high-performance liquid chromatography-purified grade were prepared for us by Scandinavian Gene Synthesis AB (Köping, Sweden). The nucleotide sequences of three sets of the primers used for specific amplification of *Salmonella rfb* gene clusters (as shown in Table 1) were based on the sequences of *rfbJ* of serogroup B (9), *rfbJ* of serogroup C2 (3), and *rfbS* of serogroup D (33). The specificities of these primers for PCR amplification were first examined with *Salmonella* reference strains (Table 2) and then by clinical isolates from the department strain collection, Diagnostic Microbiology Laboratory of Huddinge University Hospital, Huddinge, Sweden. Bacterial isolates grown on agar plates (blood and nutrient agar) and/or in liquid medium (Luria-Bertani broth) (Oxoid Ltd., Hampshire, England) at 37°C overnight were harvested and resuspended (10^5 to 10^6 CFU) with 0.1 ml of TE (Tris-EDTA) buffer in a microcentrifuge tube. Total genomic DNA from the bacterial cultures was extracted by standard procedures (27).

**Guanidium thiocyanate extraction.** Cells were lysed with 0.5 ml of a solution containing 5 M guanidium thiocyanate (Sigma Co., St. Louis, Mo.), 0.1 M EDTA, and 0.5% (v/v) Sarkosyl. The DNA was partially purified by chloroform-isooamyl alcohol (24:1), precipitated with absolute ethanol, and redissolved in TE buffer.

**Heat treatment.** Bacterial cultures were heated to 96°C for 10 min, the cell debris was removed by microcentrifugation at 20,000 × g for 10 min, and the supernatant fraction was collected. DNA concentrations were measured by calculating the UV A_260 (27). DNA purity was measured by calculating the UV A_260 and dividing by the UV A_280.

**DNA amplification and electrophoresis.** Extracted bacterial supernatants diluted in 50 μl of TE buffer (approximate DNA concentration, 0.02 μg/μl) were first heated to 96°C for 5 min and then cooled on ice, prior to PCR amplification. PCR was performed with 20 to 200 ng of crude genomic DNA and 80

---

**TABLE 1. Properties and nucleotide sequences of PCR primers for amplification of *Salmonella rfb* genes**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Target gene</th>
<th>Nucleotide sequence</th>
<th>Position*</th>
<th>T&lt;sub&gt;mel&lt;/sub&gt; (°C)</th>
<th>Size of PCR products (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><em>rfbJ</em> (B)</td>
<td>5'-AGA ATA TGT AAT TGT GAG-3'</td>
<td>15-32</td>
<td>48</td>
<td>~882</td>
</tr>
<tr>
<td>2</td>
<td><em>rfbJ</em> (B)</td>
<td>5'-TAA CGG TTT CAG TAG TGG-3'</td>
<td>897-880</td>
<td>50</td>
<td>~820</td>
</tr>
<tr>
<td>3</td>
<td><em>rfbJ</em> (C2)</td>
<td>5'-ATG CTG GAT GTG AAT AAG-3'</td>
<td>1-18</td>
<td>48</td>
<td>~820</td>
</tr>
<tr>
<td>4</td>
<td><em>rfbJ</em> (C2)</td>
<td>5'-CTA ATG GAG TCA AGA AGA-3'</td>
<td>820-803</td>
<td>50</td>
<td>~820</td>
</tr>
<tr>
<td>5</td>
<td><em>rfbS</em> (D)</td>
<td>5'-TCA CGA CTT ACA TGC TAC-3'</td>
<td>40-57</td>
<td>52</td>
<td>~720</td>
</tr>
<tr>
<td>6</td>
<td><em>rfbS</em> (D)</td>
<td>5'-CTG CTA TAT CAG CAG AAC-3'</td>
<td>760-743</td>
<td>52</td>
<td>~720</td>
</tr>
</tbody>
</table>

* Positions given are from the first base of the coding sequence.  
  b T<sub>mel</sub>, melting temperature.
### RESULTS

**Salmonella serotyping and monoclonal antibodies.** The serotyping of salmonellae was performed by bacterial agglutination testing on glass slides as previously described (19, 20) with rabbit polyclonal *Salmonella* serogroup-specific antisera provided by the National Bacteriology Laboratory (SBL, Stockholm, Sweden). The *Salmonella* reference strains used in this study were also tested by agglutination against monoclonal antibodies specific for the O2; O3; O4; O6,7; O8; and O9 antigens of *Salmonella* LPS. Generation and characterization of these mouse monoclonal antibodies have been reported in previous studies (19, 20).

**TABLE 2. Salmonella reference strains used in this study**

<table>
<thead>
<tr>
<th>Salmonella serovar</th>
<th>Serogroup and O antigens</th>
<th>Dideoxyhexose and relevant rfb genes</th>
<th>Reaction with serogroup-specific probea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paratyphii A IS2</td>
<td>AO; 2 and 12</td>
<td>Paratose; 7.6, 10.4, F, G, S, E&lt;sup&gt;b&lt;/sup&gt;</td>
<td>++ – – –</td>
</tr>
<tr>
<td>Typhimurium LT2</td>
<td>BO; 4, 5, and 12</td>
<td>Abequose; 7.6, 10.4, F, G, J</td>
<td>– – ++ –</td>
</tr>
<tr>
<td>Paratyphii B IS248</td>
<td>BO; 4, 5, and 12</td>
<td>– – ++ +</td>
<td></td>
</tr>
<tr>
<td>Paratyphii C IS32</td>
<td>CI0; 6 and 7 (Vi)</td>
<td>– – ++ –</td>
<td></td>
</tr>
<tr>
<td>Thompson IS40</td>
<td>CI0; 6 and 7</td>
<td>– – ++ –</td>
<td></td>
</tr>
<tr>
<td>Newport IS50</td>
<td>C20; 6 and 8</td>
<td>Abequose; 7.6, 10.4, F, G, J&lt;sup&gt;c&lt;/sup&gt;</td>
<td>++ – – ++</td>
</tr>
<tr>
<td>Typhi T2</td>
<td>DO; 9 and 12</td>
<td>Tyvelose; 7.6, 10.4, F, G, S, E</td>
<td>++ – – –</td>
</tr>
<tr>
<td>Typhi IS9</td>
<td>DO; 9 and 12</td>
<td>++ – – –</td>
<td></td>
</tr>
<tr>
<td>Typhi SL1058</td>
<td>DO; 9 and 12</td>
<td>++ – – –</td>
<td></td>
</tr>
<tr>
<td>Typhi 541 Ty</td>
<td>DO; 9 and 12 (Vi)</td>
<td>+++ – – –</td>
<td></td>
</tr>
<tr>
<td>Typhi 543 Ty</td>
<td>DO; 9 and 12</td>
<td>+++ – – –</td>
<td></td>
</tr>
<tr>
<td>Enteritidis SH1262</td>
<td>DO; 9 and 12</td>
<td>+++ – – –</td>
<td></td>
</tr>
<tr>
<td>Anatum IS78</td>
<td>E10; 3 and 10</td>
<td>– – – – –</td>
<td></td>
</tr>
<tr>
<td>Typhimurium TV119</td>
<td>Rough Ra</td>
<td>Abequose; 7.6, 10.4, F, G, J, Δ(rfaL)</td>
<td>– ++ – –</td>
</tr>
<tr>
<td>Typhimurium SL 805</td>
<td>Rough Rc</td>
<td>Abequose; 7.6, 10.4, F, G, J, Δ(galE)</td>
<td>– ++ – –</td>
</tr>
</tbody>
</table>

<sup>a</sup> Probes for rfbS (D), rfbJ (B), and rfbJ (C2) were cloned fragments covering, respectively, bases 1 to 609, 1 to 896, and 1 to 641. –, no reaction; ++, strong reaction; +++, very strong reaction.

<sup>b</sup> rfbE in mutant form.

<sup>c</sup> Heterologous type of rfbJ gene in serogroup C2.

pmol (or 0.2 µg) of each of the PCR primers in a total volume of 100 µl of reaction buffer that contained 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl₂, 200 µM each dideoxynucleotide triphosphate (dATP, dTTP, dCTP, and dGTP), 0.002% (wt/vol) gelatin, and 2 units of Taq DNA polymerase (Perkin-Elmer Cetus, Norwalk, Conn.). The reaction was taken through 25 or 30 cycles in a DNA thermal cycler (Perkin-Elmer Cetus). Each cycle consisted of the following: 94°C for 1 min (denaturation), 45°C for 1 min (annealing), and 72°C for 2 min (extension). After PCR, the profiles of amplification products were detected by gel electrophoresis. Ten microliters of total reaction mixture was loaded on a 1.5% agarose gel and electrophoresed at 100V at 25 mA for 45 to 60 min. The amplified DNA fragments were visualized by UV illumination after agarose gel electrophoresis and ethidium bromide staining by standard procedures (27).

The type strains of serogroup B and serogroup C2 not only differ from each other but are also distinctive from both serogroups A and D at several genes (3). Of these genes, only the *rflJ* gene (abequose synthase) has been characterized so far, and it is sufficiently different in the two strains for specific probes to be designed. Probes for PCR were synthesized by using sequences chosen for good differentiation among serogroups to give specificity in priming and also on the basis of spacing to give products of different sizes (Table 1).

The *rbfJ* and *rbfS* genes contained serogroup-specific DNA sequences for *Salmonella* abequose and paratose synthesis. Three pairs of *rbf* primers for *rbfJ* (B), *rbfJ* (C2), and *rbfS* (D), which are the *Salmonella* abequose and paratose synthase genes, were used as probes for organisms belonging to serogroups B, C2, and A plus D, respectively. All 12 *Salmonella* reference strains containing a 3,6-dideoxyhexose (abequose, paratose or tyvelose) in their O antigens showed specific reactions by PCR with separate *rbf* primers (Table 2), i.e., serogroup B reacted with the *rbfJ* (B) only, serogroup C2 reacted with *rbfJ* (C2), and serogroups A and D reacted with *rbfS* (D). Clearly, both *Salmonella* serogroup C1 and serogroup E bacteria lacking a dideoxyhexosyl residue in their O-repeating units were negative, whereas two *Salmonella typhimurium* rough strains, TV119 and SL805, in which no O chain is phenotypically expressed, were also reactive with the *rbfJ* (B) probe. This is not unexpected, because the two rough mutants tested are genetic constructs of a serogroup B parental strain.

A mixture of the three pairs of primers was used for PCR for 30 cycles, starting with 0.2 µg of genomic DNA, which was followed by simple gel electrophoresis and ethidium bromide staining (Fig. 2). The profile of target DNA fragments from each serogroup showed the following distinctive patterns: (i) amplified DNA bands at 720 bp for both serogroup A and serogroup D salmonellae, (ii) a specific band at 820 bp for serogroup C2 salmonellae only, and (iii) 882-bp DNA segments in serogroup B salmonellae as well as two rough Ra (TV119) and Rc (SL805) mutants which were derived from the serovar Typhimurium parental strain (see reference 18). In addition, PCR products were tested in a dot
Enterobacteriaceae were targeting prominent and Typhi (C2), rfbJ (D), ribS (B) groups A, enteric correctly identified all genotypes, generating the ples, intestinal blot assay (data others found to isolates including were the nine testing. expected, the from gene-specific restriction of the target ized products from the PCR products, being positive in this study. As expected, use of the rfbJ (B) primers correctly identified all of the 10 serogroup B salmonellae, yielding prominent DNA bands of ~882 bp. Similarly, all of the nine serogroup C2 isolates, but none of the serogroup C1 salmonellae, were positive in a reaction with the rfbJ (C2) primers, generating PCR products of ~820 bp. Also as expected, the rfbS (D) primers in this study showed cross-reaction between serogroup A and serogroup D salmonellae, yielding target DNA products of similar sizes (~720 bp). It was consistent with previous findings that the rfbS gene was found to be present (and common) in the Salmonella serovar Typhi and Paratyphi A strains (16, 33). To examine the possible cross-reactivities of the selected rfb primers among normal intestinal microfloras commonly found in stool samples, 68 isolates from 10 other members of the family Enterobacteriaceae were tested by the PCR procedure, and none showed a positive result. To further evaluate the primer specificities for Salmonella species, many more facultative florals and strict anaerobes (e.g., Bacteroides species) are being included in our screening trials.

FIG. 2. Agarose gel electrophoresis of PCR products after primer-directed amplification of sequences targeted in the Salmonella rfb gene clusters. Products of PCRs were electrophoresed in 1.5% agarose in Tris-borate-EDTA buffer for 45 min at 5 V/cm. The approximate sizes of the target DNA fragments were 882, 820, and 720 bp for Salmonella serogroup B rfbJ (B), serogroup C2 rfb (C2), and serogroup A and D rfbS (D), respectively.

DISCUSSION

In spite of the magnitude of salmonellosis as a serious public health burden, the identification of Salmonella species by conventional culture methods continues to be laborious and time-consuming. This has sparked our interest in searching for new probe-based methods for rapid identification of salmonellae and other important enteric pathogens in clinical and food specimens. Since its invention, the PCR technique (26) has revolutionized approaches to many facets of biomedical science, especially for molecular diagnosis of infectious and genetic diseases (31). In a short time, an array of PCR assays from different laboratories has been introduced for the successful detection and/or identification of Chlamydia trachomatis, Helicobacter pylori, Legionella pneumophila, Mycobacterium tuberculosis, and Mycoplasma pneumoniae, etc. (5, 24).

Herein, we describe a PCR system established for specific detection of Salmonella serogroups A, B, C2, and D, which constitute about 70% of Salmonella infections in humans and animals (25). Our PCR assay for Salmonella identification is based on a defined set of oligonucleotide primers targeting the rfb gene clusters. The rfb genes are required for the biosynthesis of bacterial polysaccharide O antigens (8, 22); putatively, the regions encoding the 3,6-dideoxyhexosyl synthases determine the antigenic specificities of Salmonella serogroups A, B, C2, and D (Fig. 1). Three unique regions of the rfb clusters were subsequently selected as target sites in the PCRs: (i) rfbS, present in serovar Typhi (serogroup D) and in serovar Paratyphi A (serogroup A), which has a mutation in the final step of the tyvelose synthesis pathway; (ii) the rfbJ for abequose synthase in serovar Typhimurium (serogroup B); and (iii) its rfbJ counter part in serovar
TABLE 3. Specificity of DNA amplification by PCR of salmonellae and other enteric bacteria with the rfb and rfbS gene primers

<table>
<thead>
<tr>
<th>Clinical isolate</th>
<th>No. of strains tested*</th>
<th>No. of positive reactions with the following PCR products:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>720 bp 882 bp 820 bp</td>
</tr>
<tr>
<td>Salmonella serogroup (serotype)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A (1)</td>
<td>10</td>
<td>10 0 0</td>
</tr>
<tr>
<td>B (6)</td>
<td>10</td>
<td>0 10 0</td>
</tr>
<tr>
<td>C1 (5)</td>
<td>5</td>
<td>0 0 0</td>
</tr>
<tr>
<td>C2 (8)</td>
<td>9</td>
<td>0 0 0</td>
</tr>
<tr>
<td>D (7)</td>
<td>11</td>
<td>11 0 0</td>
</tr>
<tr>
<td>E (9)</td>
<td>10</td>
<td>0 0 0</td>
</tr>
</tbody>
</table>

Other members of the family:
- Enterobacteriaceae
  - Arizona species
    - Citrobacter diversus
    - Citrobacter freundii
  - Enterobacter species
    - Escherichia coli
    - Klebsiella pneumoniae
    - Morganella species
    - Proteus mirabilis
    - Shigella species
    - Yersinia enterocolitica

* A total of 123 clinical isolates were tested (55 salmonellae from 36 different serotypes and 68 isolates from 10 other members of the Enterobacteriaceae family).

Muenchen (serogroup C2). The primers were designed to give PCR products with different sizes; since all three genes are homologous, they were from regions of very low similarity. The accuracy of the PCR technique to detect salmonellae is based on the specificities of designed oligonucleotide primers used under defined thermal cycle conditions. This was verified by correct identification of all 40 salmonellae (10, 9, 11, and 10, belonging to serogroup B, serogroup C2, serogroup D, and serogroup A isolates, respectively) among the 123 clinical isolates of Salmonella species as well as other bacterial species (Table 3). On the basis of the exquisite specificity of PCR reactions, we suggest that the selected nucleotide sequences within the rfb regions for abequous and parasute synthases are specific DNA markers for detection and/or identification of Salmonella serogroups A, B, C2 and D. Furthermore, preliminary evaluation of the potential usefulness of this assay in ongoing trials against over 100 clinical isolates from diarrheal patients in our laboratory at Huddinge Hospital demonstrated accurate detection of salmonellae (sensitivity, \( \sim 10^5 \) to \( 10^6 \) bacteria) in stool specimens (10a).

Since the sequences of the rfb genes of serogroups C1 and E are now available (11, 34), it has also been our interest to extend this method to other common Salmonella serogroups. In a recent study (36), we found several single genes with the serogroup C1 probes which did not hybridize with strains of the 44 other serogroups tested and a single gene with the serogroup E probe which hybridized with strains only of serogroups E and D2. Serogroups E and D2 were readily distinguished by several other single-gene probes. On the basis of the immunochromy of Salmonella LPS, to the effect that any two serogroups should be distinguished by at least one sugar composition or linkage, it would be possible to identify any serogroup by PCR of the appropriate rfb genes; however, in some cases it is a specific combination of genes which defines a serogroup, rather than one gene as in the cases described in the present article. The one known exception is serogroup A, which, as discussed above, differs from serogroup D by mutation but not loss of a gene.

Earlier studies have indicated the potential of PCR technology to detect in feces the Shigella invasion plasmid antigen gene (ila) (4), E. coli heat-stable enterotoxin genes (6), and the Vibrio cholerae enterotoxin operon (29), without the necessity of cultivating organisms in laboratory media. Our next attempt will be to apply the described PCR system in routine laboratory diagnoses for direct detection of salmonellae in clinical and food specimens. To alleviate the interference of unknown inhibitors present in the samples, we have been utilizing immunomagnetic beads coupled with specific antibodies against the bacterial somatic O antigens to separate target organisms from interference materials in clinical specimens (20). A similar approach employed to detect Shigella in stool samples proved to be successful with a sensitivity of 10 to 100 bacteria (7). Optimization of the procedures combining both immunomagnetic separation of salmonellae and PCR amplification of the corresponding rfb genes is under way. This would make it possible to detect and/or identify clinically important Salmonella and Shigella species upon arrival of specimens in the diagnostic microbiology laboratory within 1 working day.

ACKNOWLEDGMENTS

This work was supported by grant 16X-656 from the Swedish Medical Research Council.

We thank Lennart Lindqvist for helpful discussion, Sten Karshrink of the Clinical Research Center of Huddinge Hospital for photographic work, and Gordon Stevenson for performing the dot blot hybridization assays.

REFERENCES


