JOURNAL OF BACTERIOLOGY, Sept. 1995, p. 5247–5253 Vol. 177, No. 18
0021-9193/95/$04.00+0

Lateral Transfer of *rfb* Genes: a Mobilizable ColE1-Type Plasmid Carries the $rfb_{O:54}$ (O:54 Antigen Biosynthesis) Gene Cluster from *Salmonella enterica* Serovar Borreze

WENDY J. KEENLEYSIDE AND CHRIS WHITFIELD*

Department of Microbiology, University of Guelph, Guelph, Ontario, Canada N1G 2W1

Received 29 March 1995/Accepted 9 July 1995

Plasmid pWQ799 is a 6.9-kb plasmid isolated from *Salmonella enterica* serovar Borreze. Our previous studies have shown that the plasmid contains a functional biosynthetic gene cluster for the expression of the O:54 lipopolysaccharide O-antigen of this serovar. The minimal replicon functions of pWQ799 have been defined, and a comparison with nucleotide and protein databases revealed this replicon to be virtually identical to ColE1. This is the first report of involvement of ColE1-related plasmids in O-antigen expression. The replicon of pWQ799 is predicted to encode two RNA molecules, typical of other ColE1-type plasmids. RNAII, the putative replication primer from pWQ799, shares regions of homology with RNAII from ColE1. RNAI is an antisense regulator of DNA replication in ColE1-related plasmids. The coding region for RNAI from pWQ799 shares no homology with any other known RNAI sequence but is predicted to adopt a secondary structure characteristic of RNAI molecules. pWQ799 may therefore represent a new incompatibility group within this family. pWQ799 also possesses *cer*, *rom*, and *mob* determinants, and these differ minimally from those of ColE1. The plasmid is mobilizable in the presence of either the broad-host-range helper plasmid pRK2013 or the $IncI_1$ plasmid R64*drd*86. Mobilization and transfer of pWQ799 to other organisms provides the first defined mechanism for lateral transfer of O-antigen biosynthesis genes in *S. enterica* and explains both the distribution of related plasmids and coexpression of the O:54 factor with other O-factors in different *Salmonella* serovars. The base composition of the pWQ799 replicon sequences gives an average percent G+C value typical of *Salmonella* spp. In contrast, the percent G+C value is dramatically lower within $rfb_{O:54}$, consistent with the possibility that the cluster was acquired from an organism with a much lower G+C composition.

Lipopolysaccharide (LPS) is an integral component of the gram-negative outer membrane. It contains three structural domains. Lipid A forms the inner domain and serves to anchor the molecule in the outer leaflet of the outer membrane. The central portion consists of a core oligosaccharide region, while the outermost portion is composed of the O-polysaccharide. The O-polysaccharide (or O-antigen) extends from the cell surface, forming a polysaccharide layer. In the absence of a capsular polysaccharide layer, the O-antigen serves as an interface between the bacterial cell and its surrounding environment. The O-polysaccharide exhibits extensive structural variation (60), and this variation contributes significantly to the serological diversity of gram-negative bacteria.

The genus *Salmonella* consists of two species: *Salmonella enterica*, which is further subdivided into six subspecies, and *Salmonella bongori* (35). The *Salmonella* serotyping scheme is based on differences in the O- and H (flagellar filament)-antigens and has resulted in the description of more than 2,000 serotypes (35). Salmonellae are common inhabitants of the gastrointestinal tract in humans and animals, and many are facultative intracellular parasites. These organisms cause a variety of diseases, including acute gastroenteritis, septicemia, and typhoid (41). The O-polysaccharide has been shown to be involved in resistance to complement-mediated serum killing (reviewed in references 15 and 19) and, therefore, represents an important virulence factor in *Salmonella* spp.

* Corresponding author. Mailing address: Department of Microbiology, University of Guelph, Guelph, Ontario, Canada N1G 2W1. Phone: (519) 824-4120, ext. 3478. Fax: (519) 837-1802. Electronic mail address: cwhitfie@micro.uoguelph.ca.

The evolution of serological diversity in *Salmonella* spp. and in the family *Enterobacteriaceae* in general is a subject of considerable interest. The *rfb* gene cluster directs the synthesis of the O-polysaccharide repeating unit. It has been suggested that O-antigen structural diversity has resulted from repeated gene transfer and recombination events between and within *rfb* clusters. Adaptation to a particular niche is speculated to provide the selective pressure for continued lateral gene transfer (38, 39). The mechanisms involved in lateral gene transfer of *rfb* genes have not been defined.

While most *rfb* genes are located on the *Salmonella* chromosome, expression of the O:54 antigen requires the presence of a small plasmid (21, 36, 37). Studies in this laboratory have shown that the $rfb_{O:54}$ cluster of *S. enterica* serovar Borreze exists on a 6.9-kb plasmid, which we have designated pWQ799 (21). This cluster directs the synthesis of an *N*-acetylmannosamine-(Man*p*NAc)-containing disaccharide repeating unit with the structure: [→4)-β-D-Man*p*NAc-(1→3)-β-D-Man*p*NAc-(1→] (21). This structure is unlike that of any other *Salmonella* O-antigen. Similarly, the $rfb_{O:54}$ from *S. enterica* serovar Borreze is unique, in that it is the only plasmid-encoded *rfb* cluster in *Salmonella* spp. to be described.

The presence of $rfb_{O:54}$ on a naturally occurring plasmid raises interesting questions concerning the origin of this plasmid and has important implications for potential lateral gene transfer. We now report the characterization of the replicon-associated functions of pWQ799. The minimal replicon of pWQ799 is characteristic of a ColE1-like plasmid and, on the basis of a comparison with sequence databases, may represent a new incompatibility group. The plasmid possesses a functional *mob* region, and transfer of pWQ799 in the presence of

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Genotype or relevant property	Source or reference
Strains		
S. enterica serovar Borreze SA902282	54:f,g,s:-; contains plasmids of 96, 4.5, and 2.3 MDa; O:54$^+$	37
E. coli DH5α	K-12 ϕ80d *lacZ*ΔM15 *endA1 recA1 hsdR17* (r_K^- m_K^-) *supE44 thi-1 gyrA96 relA1* Δ(*lacZYA-argF*)U169 F$^-$	42
E. coli LE392	K-12 *hsdR574* (r_K^- m_K^-) *supE44 supF58 lacY1* or Δ(*lac-proAB*)*6 galK2 galT22 metB1 trpR55* λ$^-$ F$^-$	42
E. coli JM109Nal^r	K-12 *endA1 recA1 gyrA96 thi hsdR17* (r_k^- m_k^-) *relA1 supE44* λ$^-$ Δ(*lac-proAB*) [F′ *traD36 proAB lacI*^{q} ZΔM15] spontaneous nalidixic acid resistant	This study; 42
E. coli JE 2571-2	K-12 *leu thr str fla pil* Rif^r	5
Plasmids		
pRK2013	RK2 derivative Km^r Mob$^+$ Tra$^+$ ColE1	12 14
R64*drd*86	Transfer-derepressed mutant of IncI$_1$ plasmid R64 from *S. typhimurium*; Sm^r Tc^r	6
pGEM-7Zf(+)	Cloning vector Ap^r	Promega
pWQ799	Naturally occurring plasmid containing *rfb*$_{O:54}$	21
pWQ800	pWQ799/pGEM-7Zf(+) hybrid	21
pUC4K	Source of Km^r cassette	Pharmacia
pWQ802	pWQ799 with Km^r cassette from pUC4K inserted at *Eco*RI site	This study
pWQ806	1.75-kb *Pvu*II fragment from pWQ800, ligated to Km^r cassette	This study
pWQ807	4.0-kb *Hin*cII fragment from pWQ802, ligated to Km^r cassette	This study

tra functions provided by appropriate helper plasmids confers O:54 serological specificity on the recipient.

MATERIALS AND METHODS

Bacterial strains, growth conditions, and plasmids. The bacterial strains and plasmids used in this study are described in Table 1. *S. enterica* serovar Borreze is an isolate obtained from dust samples in a barn housing a commercial turkey flock in Manitoba, Canada (37). Bacteria were routinely grown at 37°C in Luria-Bertani broth, supplemented with antibiotics where required (ampicillin, 100 μg/ml; kanamycin, 30 μg/ml; nalidixic acid, 40 μg/ml).

DNA manipulation and analyses. Restriction enzyme digestion, ligation, and transformation reactions were performed as described elsewhere (42). Gel purification of restriction fragments and gene probes was done with the GENECLEAN kit (Bio/Can Scientific, Mississauga, Ontario, Canada). Plasmid DNA was purified by using QIAGEN columns (QIAGEN Inc., Chatsworth, Calif.), according to the manufacturer's instructions.

The DNA sequence of pWQ799 was determined by using a combination of pWQ799 subclones and synthetic oligonucleotides (MOBIX Inc., Hamilton, Ontario, Canada). Sequencing of both strands was performed by MOBIX Inc., using the Applied Biosystems cycle sequencing system with an automated fluorescent sequencer. Sequence data were edited and analyzed by using AssemblyLIGN and MacVector software (International Biotechnologies Inc., New Haven, Conn.). Nucleotide and amino acid sequence homology searches of the National Center for Biotechnology Information databases were done with the BLAST (stands for basic local alignment search tool) server analysis program (1). RNA secondary structure was analyzed with the PCFOLD program (65).

SDS-polyacrylamide gel electrophoresis (PAGE) of LPS. LPS samples were prepared by using the sodium dodecyl sulfate (SDS)-proteinase K whole-cell lysate method of Hitchcock and Brown (18). Samples were run in commercially prepared 1-mm-thick, 10 to 20% tricine gels from Novex (San Diego, Calif.). Silver-staining procedures (55) were as originally described.

Plasmid mobilization studies. Triparental mobilization studies of plasmids pWQ802, pWQ806, and pWQ807 were performed with the conjugative plasmid pRK2013 in *Escherichia coli* LE392 or R64*drd*86 in *E. coli* JE2571-2. A spontaneous nalidixic acid-resistant mutant of *E. coli* JM109 served as the recipient strain. Samples from overnight broth cultures of the recipient, helper, and donor strains were used for conjugal transfers. Viable counts of the donor and recipient strains were determined by dilution plating prior to mating experiments. Then 100 μl (each) of donor, recipient, and helper strains was used to inoculate fresh broth tubes. Mating was carried out in a static 37°C incubator for 3 h, and the mating mixtures were dilution plated on Luria-Bertani plates supplemented with kanamycin and nalidixic acid. Frequency of transfer was expressed as the proportion of recipient cells receiving a plasmid.

Nucleotide sequence accession number. The DNA sequence of pWQ799 has been entered in GenBank under accession number L39794.

RESULTS

Localization and subcloning of the plasmid replicon sequences of pWQ799. Initial studies of pWQ799 were done with pWQ800, a hybrid plasmid consisting of the *Eco*RI-linearized *Salmonella* plasmid, ligated to the commercial vector pGEM-7Zf(+) (21). pWQ799 itself contains no antibiotic resistance determinants. To facilitate analysis of the plasmid replicon sequences, a kanamycin resistance cassette was introduced into the unique *Eco*RI site of pWQ799, generating pWQ802 (Fig. 1B). This plasmid was stably maintained in *E. coli* K-12 hosts, and there is no obvious difference in the levels of pWQ802 and the native plasmid (data not shown). *E. coli* K-12(pWQ802) was agglutinated by anti-O:54 polyclonal antiserum, and SDS-PAGE analysis of purified LPS from this strain confirmed expression of full-length O:54 O-antigen (Fig. 2). Plasmid pWQ807 was obtained by ligating the 4.25-kb *Hin*cII fragment from pWQ802 to the kanamycin resistance cassette (Fig. 1B). This plasmid was also stably maintained in *E. coli* K-12 but was unable to direct the synthesis of the O:54 O-polysaccharide (Fig. 2), indicating that the deleted sequences contained at least part of *rfb*$_{O:54}$. Plasmid pWQ806 contains the 2.06-kb *Pvu*II fragment of pWQ800 ligated to the kanamycin resistance cassette (Fig. 1B). This plasmid represents the smallest subclone capable of replication. These subcloning experiments provided preliminary information on the location of the pWQ799 sequences involved in plasmid replication and localized the origin of replication to the DNA fragment contained in pWQ806.

Analysis of pWQ799 minimal replicon sequences. Using synthetic oligonucleotides and pWQ799 subclones, we determined the nucleotide sequence of plasmid pWQ799 and compared the data with sequences in the National Center for Biotechnology Information databases. On the basis of computer analyses of possible coding regions in the plasmid and the results of these searches, the replicon sequences were identified and mapped (Fig. 1A). These analyses identified pWQ799 as a member of the ColE1-type plasmid group. There are a number of plasmids within this family; however, pWQ799 displayed the

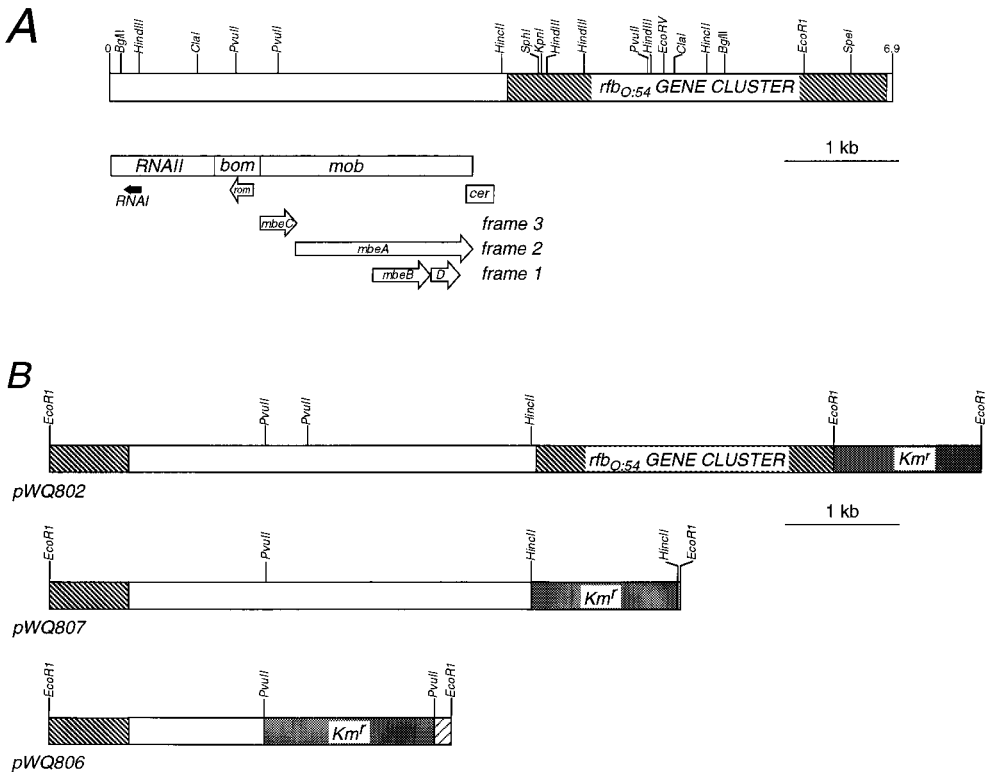

FIG. 1. Physical maps of pWQ799 and subclones. (A) Genetic organization and features of pWQ799 replicon sequences, based on sequence analysis and homology with ColE1. The position of the $rfb_{O:54}$ genes is only shown for clarity; the functions of the $rfb_{O:54}$ genes are presently unknown. *bom*, basis of mobility; *mob*, mobility; *cer*, resolution. (B) Physical maps of pWQ799 subclones. Note that the additional *Hin*cII site of pWQ807 comes from the kanamycin resistance (Km^r) cassette in pWQ802; the 1.2-kb *Pvu*II-*Eco*RI fragment in pWQ806 comes from pGEM-7Zf(+) in pWQ800.

highest homology with the prototype, the *E. coli* plasmid ColE1.

Replication of ColE1 and related plasmids has been extensively characterized (reviewed in references 8, 25, 32, and 34).

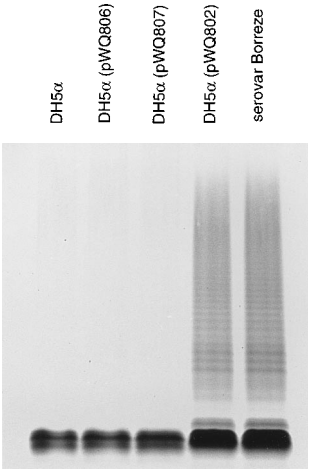

FIG. 2. Tricine–SDS-PAGE profiles of LPS from *S. enterica* serovar Borreze (containing pWQ799) and *E. coli* DH5α recombinants harboring pWQ799 subclones. Although the LPS core region is well resolved in the $O:54^+$ strains, the O-antigen bands are characteristically smeared, even when purified O:54 LPS is examined. This is presumably due to the chemistry of the O:54 repeating unit. Plasmids pWQ807 and pWQ806 are unable to direct synthesis of the O:54 polysaccharide, indicating that the deleted sequences contain $rfb_{O:54}$ genes.

Minimal replicons of these plasmids do not encode any proteins but instead contain determinants for two RNA molecules which control the replication process. RNAII is responsible for priming DNA synthesis and is transcribed from a site upstream of the origin of replication (*oriV*). Transcription terminates beyond *oriV*, at a site which appears to be plasmid specific. The RNAI determinant overlaps the 5′ end of RNAII and encodes a smaller antisense molecule. The origin of replication, *oriV*, is also part of the minimal replicon of these plasmids.

Analysis of the sequences involved in DNA replication revealed that the region encoding RNAII in pWQ799 shared extensive homology with that in ColE1 (Table 2). RNAII ho-

TABLE 2. Nucleotide sequence comparison of replication control determinants of ColE1 and pWQ799

ColE1 determinant	Position (nt) in:		% Identity
	ColE1	pWQ799	
Origin of replication (*oriV*)[a]	1196–1198	664–666	100
RNAII determinant[b]	586–661	72–278	73
	826–981	302–457	84
	1154–1220	622–688	83
Resolution (*cer*) region[c]	3686–3927	3143–3384	92
rom (RNAI modulator)[d]	1803–1614	1259–1070	96

[a] See reference 54.

[b] nt 641 to 1196 in ColE1 (9). Only blocks within the RNAII determinants showing significant homology are compared.

[c] nt 3686 to 4066 in ColE1 (50). Only the parts of *cer* showing significant homology are compared (see Fig. 4).

[d] See reference 9.

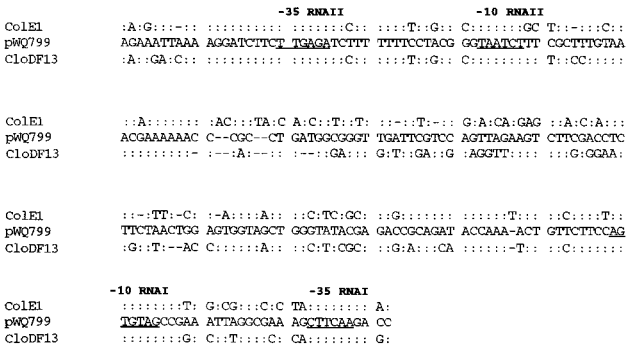

FIG. 3. Nucleotide sequence alignment of region encoding RNAI. The top sequence is from ColE1 (9) and begins at nt 587. The middle sequence is from pWQ799 and begins at nt 72. The bottom line represents the sequence from the ColE1-related plasmid CloDF13 (31) and begins at nt 932. Identical residues are indicated (:), as are gaps (-). Putative −10 and −35 sequences are underlined.

mology was highest at the 3′ end, in the region surrounding *oriV*. A conserved 28-bp region containing *oriV* was identified in pWQ799 (21a). The conserved region includes a poly(A) sequence typical of plasmid replicons of this type (2). Sequence homology in RNAII was somewhat lower at the 5′ end, particularly in the RNAI coding region (nucleotides [nt] 238 to 131 in pWQ799). A nucleotide alignment between the RNAI determinants of ColE1, the related *Enterobacter cloacae* plasmid CloDF13, and pWQ799 is shown in Fig. 3. The alignment indicates that the putative promoters for RNAII and RNAI are identical in all three plasmids. Two *dam* methylation sites were identified in the RNAII promoter region of pWQ799 (nt 84 to 87 and 95 to 98), at the same positions as for similar sites in ColE1 and CloDF13. These methylation sites are required for efficient CloDF13 plasmid replication (56). The intervening sequence between the two promoters diverges substantially; this corresponds to the RNAI coding region. All RNAI molecules examined so far possess a secondary structure resembling tRNA, with three stem-loop domains and a free single-stranded 5′ end (29). This structure is thought to be required for the activity of RNAI (26, 28). Despite the observed sequence divergence, the RNAI molecule encoded by pWQ799 is predicted by the PCFOLD algorithm (65) to adopt a similar conformation, with a free energy of −27.8 kcal (1 cal = 4.184 J). RNAI interacts with RNAII to negatively regulate primer formation by inhibiting processing of the preprimer by ribonuclease H (27, 52, 61). RNAI acts in *trans* and is the primary determinant of plasmid incompatibility (13, 23, 51, 53). On the basis of the sequence divergence of RNAI, pWQ799 represents a new incompatibility group among those ColE1-related plasmids for which the RNAI sequence has been determined. This is supported by the fact that pWQ799 and its derivatives are stably maintained in *E. coli* K-12 in the presence of the ColE1 derivative pGEM-7Zf(+) (21a).

Although RNAI and RNAII transcripts have not been confirmed, the extensive sequence conservation within the replicative region and the predicted secondary structure of RNAI, combined with the ability of pWQ799 subclone pWQ806 to be efficiently replicated, provide convincing evidence as to the function of this region.

High-copy-number plasmids are distributed randomly to daughter cells upon cell division (33). In order to maintain a high number of independently segregating units, a system is required for the efficient resolution of plasmid multimers. In ColE1, this occurs through site-specific recombination involving *cer*, a *cis*-acting plasmid-encoded sequence (49). This recombination event is dependent on the chromosomally encoded recombinases XerC (10) and XerD (4), in addition to the accessory proteins ArgR (46) and PepA (45), whose binding sites all lie within *cer*. XerCD-mediated plasmid resolution alone is insufficient to stably maintain ColE1 plasmids in recombination-proficient hosts. *cer* also encodes a transcript (designated Rcd) which is believed to contribute to plasmid stability (33). Accumulation of plasmid multimers appears to activate transcription from the Rcd promoter, with a concomitant inhibition of cell division. It is thought that this inhibition gives the cells time to resolve plasmid multimers. CloDF13 also possesses a *cis*-acting site of recombination (*parB*), although this region maps to a different plasmid location than *cer* (16). Alignment of the resolution region of pWQ799 with those of ColE1 and CloDF13 (Table 2; Fig. 4) revealed extensive homology between the pWQ799 and ColE1 nucleotide sequences (92%) and somewhat lower conservation in CloDF13. The −35 and −10 regions identified for Rcd, which fall within the ArgR binding site (46), are conserved in pWQ799, as is the proposed Rcd terminator (33). The pWQ799 *cer* homolog maps to the junction between the replicon sequences and the 5′ end of $rfb_{O:54}$ (Fig. 1A). The location of this sequence in relation to the other replicon sequences is identical to that of ColE1, overlapping the end of the mobilization region at the 5′ end (Fig. 1A) (discussed below).

The homology among all three resolution sequences diverges in the crossover region. This region contains the target sequence recognized by XerCD (11, 49). The target site consists of two well-conserved 11-bp sequences flanking a shorter variable region. Alignment with ColE1 indicates conservation of 9 of the 11 nt in the left-hand, or XerC, recognition site (4) but only 4 of the 11 nt in the right-hand, or XerD, recognition site (4). This may reflect differences in the sequence specificities or stringencies of the *Salmonella trans*-acting recombinases. Alternatively, the sequence divergence may suggest that *cer* is no longer an efficient target for the action of site-specific recombinases. Interestingly, *cer* has been deleted in pWQ806 (Fig. 1), yet the plasmid does not exhibit any obvious instability, nor does it display any obvious multimer formation when examined by agarose gel electrophoresis (21a).

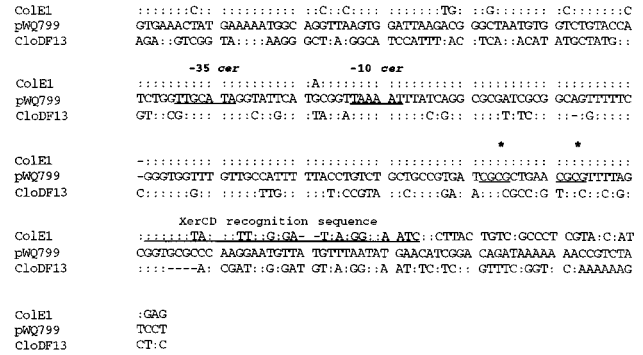

FIG. 4. Nucleotide sequence alignment of the plasmid resolution regions of ColE1 and CloDF13 with that of pWQ799. The top sequence represents ColE1 and begins at nt 3730 (9). The *cer* sequence of pWQ799 is shown in the middle, beginning at nt 3187. CloDF13 is shown on the bottom, beginning at nt 4981 (31). Identical bases are indicated (:), as are gaps (-). The −35 and −10 promoter regions are for the Rcd transcript encoded by *cer* in ColE1 (33). This is also the region where ArgR binds (46). A region of palindromic symmetry is indicated by the asterisks and corresponds to the Rcd terminator of ColE1. The XerCD recognition sequence (underlined) is from reference 11 and represents the crossover region, where the XerCD recombinases act.

Detection of RNAI modulator (*rom*) in pWQ799. Replication of ColE1-related plasmids is controlled by the binding of RNAI to RNAII. In some plasmids, this interaction is regulated by the product of the *rom* gene (stands for RNAI [one] modulator; previously called *rop* [repressor of primer]). In ColE1, *rom* maps immediately downstream of *oriV* and encodes a small (63 amino acid residues) acidic protein, which enhances the inhibitory activity of RNAI in vitro and in vivo. This effect is mediated through Rom-dependent stabilization of the RNAI-RNAII complex (reviewed in references 8 and 34). The *rom* gene product is not essential but exerts some control over plasmid copy number. *rom* is conserved in ColE1 and pWQ799, and the homologs share 96% identity (Table 2). Since the frequency of generation of plasmid-free daughter cells is related to the number of plasmid molecules at septation (48), expression of this protein would provide an additional mechanism for maintaining plasmid stability in pWQ799-containing cells.

Characterization of the *mob* region and conjugal transfer of pWQ799. ColE1-type plasmids are generally mobilizable in the presence of an appropriate conjugation system, supplied by another plasmid. Approximately 30% of the ColE1 genome is devoted to mobilization functions (*mob*). The *mob* region is conserved in the closely related plasmids ColA and ColK. The four genes within this region show complex organization, with two *mob* genes (*mbeB* and *mbeD*) being entirely overlapped by a third (*mbeA*). pWQ799 shows an identical arrangement (Fig. 1A). Nucleotide sequence analysis revealed 95% identity with the entire *mob* region of ColE1. Extensive sequence homology was also detected with the *bom* (stands for basis of mobility) site, which maps immediately upstream of *mob* and contains the origin of transfer (*oriT*). It is the *bom* region which is recognized by the Mob proteins.

The extensive sequence conservation of the mobilization determinants suggested that pWQ799 might be mobilizable if transfer functions were provided in *trans*. Triparental matings using pRK2013 as the helper plasmid indicated that both pWQ802 and pWQ807 (Fig. 1B) were mobilizable. Average efficiency of transfer was 5%. pWQ806 contains the replication sequences of pWQ799 and *oriT* but none of the *mob* genes and could not be mobilized. Since pRK2013 is itself a Mob^+ ColE1 derivative (14), the function of the pWQ799 Mob proteins was further examined by using the conjugative plasmid R64*drd*86 ($IncI_1$ [6]) in place of pRK2013. Plasmids belonging to groups IncF, IncI, and IncP have been previously shown to mobilize ColE1 derivatives with high efficiency (58), provided that the nonconjugative plasmid encodes the *mob* genes and *oriT*. Under these conditions, plasmids pWQ802 and pWQ807 were again mobilized, although the transfer efficiencies were approximately 10-fold lower. pWQ806 could not be mobilized. Together, these data provide convincing evidence that the transfer was mediated by donation and dependent on the Mob proteins of pWQ799.

In all cases, transfer of pWQ802 conferred an O:54 serotype on the recipient *E. coli* K-12 strain (21a).

G+C content. The average G+C content of pWQ799 is 46%. A base composition plot of the entire sequence revealed that the $rfb_{O:54}$ sequences are uniformly richer in A and T residues (21a). Calculating the percent G+C of the $rfb_{O:54}$ sequences and the remaining plasmid sequences separately, the values were 39 and 54%, respectively. These percent G+C values are consistent throughout the regions comprising either *rfb* or the remaining plasmid sequences. An average G+C content of 50 to 53% is typical for *Salmonella* genomic DNA (39). Three small areas rich in A and T residues were identified in the plasmid replication and mobilization regions of pWQ799. The first two correspond to the promoter regions of RNAII (nt 75 to 120) and RNAI (nt 290 to 235). The third corresponds to *oriT* (nt 820 to 935). The base composition of the latter may reflect the requirement for the local DNA strand separation associated with conjugation (2).

DISCUSSION

The O:54 antigen is coexpressed with a variety of different O-antigens in *S. enterica*, and its expression requires a small plasmid (36). These observations were explained by the presence of a complete and functional $rfb_{O:54}$ gene cluster on the 6.9-kb plasmid, pWQ799, from *S. enterica* serovar Borreze (21). Studies in our laboratory indicate that the plasmids found in different O:54 isolates fall into four different clonal groups on the basis of size and polymorphism in *Eco*RI restriction sites (21a). However, all of the plasmids hybridize with pWQ799, and pWQ799 has been studied as the prototype. Plasmid-encoded functions are involved to various extents in the biosynthesis of O-antigen in a number of other bacterial species. The form I LPS O-antigen of *Shigella sonnei* exists on a large virulence plasmid (22, 57, 64). *Shigella dysenteriae* requires the plasmid-encoded *rfp* determinant in addition to chromosomal *rfb* genes, for serotype 1 O-antigen expression (47, 59). One strain of *E. coli* O111 requires a plasmid for O-antigen synthesis (40), although other strains have chromosomal rfb_{O111} (40). *S. enterica* serovar Dublin is reported to require the product of a plasmid-encoded *rfc*-like gene (20). Expression in Δ*rfb E. coli* K-12 indicates that the $rfb_{O:54}$ cluster on pWQ799 is sufficient for O:54 synthesis (21). Most *rfb* genes are located on the chromosome of *S. enterica*, and $rfb_{O:54}$ from serovar Borreze is the only plasmid-encoded *rfb* cluster described for *Salmonella* spp. pWQ799 is, to our knowledge, the first naturally occurring *rfb*-associated replicon to be characterized in any detail.

O-antigens are structurally (and therefore serologically) diverse. It has been suggested that the evolution of the O-antigen in *S. enterica* diversity has resulted from repeated gene transfer and recombination events involving *rfb* (39). Evidence for intra-*rfb* recombination has been reported for *S. enterica* serogroup D2, and it has been proposed that flanking H-repeats, resembling insertion sequences, may be involved (63). Horizontal transfer of *rfb* DNA has also been implicated in the genesis of a new epidemic isolate of *Vibrio cholerae* serogroup O139 (3). This transfer is believed to have occurred between an environmental isolate and an El Tor-like O1 strain, giving rise to a novel strain, which has the ability to cause epidemic cholera. Although insertion sequence elements are also speculated to have mediated this recombination event, a critical unresolved aspect in this process is a mechanism for lateral *rfb* gene transfer between the donor and recipient strains. The finding that pWQ799 belongs to the ColE1 family of small mobilizable plasmids provides the first direct demonstration of one such mechanism. A similar argument has been proposed for the generation of flagellar antigen diversity in *Salmonella* spp. Salmonellae are normally diphasic, producing two antigenic forms of the flagellin protein. However, some strains are characterized by a more complex serovar, producing three or more distinct flagellins. The flagellin genes (*fliC*) are typically chromosomally encoded, but in at least one triphasic strain, a third *fliC* gene has been found on a large plasmid, giving rise to the possibility of lateral gene transfer and homologous recombination with the chromosomal genes (44).

There are a number of plasmids within the ColE1 family, and many aspects of the biology of ColE1 plasmids are known. Characterization of pWQ799 replication functions is confined

to sequence comparison with those plasmids for which mechanisms have been determined. RNAI analysis indicates that pWQ799 represents a new incompatibility group among these plasmids, but despite this divergence, pWQ799 displayed striking homology with the prototype plasmid ColE1, from *E. coli*. The overall similarities strongly support identical replicon function in pWQ799. Plasmids of the ColE1 family carry a number of different genes. However, pWQ799 is unique among these plasmids in that it does not contain genes for bacteriocin production and immunity, or antibiotic resistance (17, 24), but instead encodes a functional O-antigen biosynthetic gene cluster. ColE1 plasmids are not self-transmissible but can be mobilized by a conjugative plasmid. This aspect is particularly important with respect to the involvement of pWQ799 in lateral transfer of O-antigen genes. The function of the pWQ799 *mob* genes has been directly demonstrated. The plasmid is efficiently mobilized when transfer functions are provided in *trans* by the broad-host-range conjugative plasmid pRK2013 or by the $IncI_1$ conjugative plasmid R64*drd*86. Since pWQ799 contains a complete $rfb_{O:54}$ gene cluster, transfer of the plasmid is sufficient to alter O-serotype specificity. *Salmonella* O:54 serovars express other O-factors in addition to O:54 (21, 36, 62), indicating that conjugal plasmid transfer occurs with smooth as well as rough recipient strains.

The low percent G+C value of $rfb_{O:54}$ is not unexpected. Most *rfb* clusters analyzed to date possess low percents G+C, relative to the average for the species. It has been suggested that this may reflect a lateral transfer of the *rfb* cluster from a strain with a lower G+C composition (39). The same explanation can be proposed for pWQ799, and the observed differences in percent G+C between $rfb_{O:54}$ and the replicon are consistent with recent acquisition of the $rfb_{O:54}$ cluster. However, it is also conceivable that the AT bias seen in many LPS biosynthesis genes may reflect other requirements, for example, a need to maintain low translational levels (30, 43, 60). The manner in which $rfb_{O:54}$ was acquired by the ColE1-like plasmid remains a matter of speculation. Analysis of the DNA flanking $rfb_{O:54}$ failed to reveal any sequences resembling insertion sequence elements, although a 6-bp inverted repeat was identified. Inverted repeat A (GGGTTC) occurs at nt 40 to 35 (data not shown), upstream of the RNAII promoter region, while inverted repeat A′ (CCCAAG) is located at nt 3373 to 3378 (Fig. 4). By alignment with ColE1, the latter position falls within the XerC recognition site (4). It is, therefore, possible that an insertion was mediated by a site-specific recombination process involving two unrelated DNA molecules, rather than a multimer of the pWQ799 progenitor. This would also provide an explanation for the considerable sequence divergence in the XerD recognition site of pWQ799. A similar explanation has been suggested for the presence of two genes in another ColE1-type plasmid from *Salmonella* spp. Plasmid NTP16 is a small, multicopy nonconjugative plasmid from *S. enterica* serovar Typhimurium and is related to *Salmonella* plasmid NTP1 (24). NTP16 carries transposon-derived sequences including genes for ampicillin and kanamycin resistance. It also carries two open reading frames, ORF1 and ORF2, which map immediately adjacent to the NTP16 *cer* equivalent. The functions of these open reading frames are unknown, but their percents G+C are extremely low compared with that of the surrounding DNA. The plasmid is, therefore, thought to have inherited both open reading frames simultaneously, and the mechanism for this inheritance is speculated to involve the *cer* analog, *nmr* (7).

In summary, we have shown that $rfb_{O:54}$ is carried on a ColE1-type replicon and that this plasmid is mobilizable when supplied with the appropriate transfer functions in *trans*. This ability provides a mechanism for the generation of antigenic diversity and may not be confined to *Salmonella* spp.

ACKNOWLEDGMENTS

We thank G. Harauz for help with analysis of the RNA secondary structure, P. Krell for stimulating discussions, and D. Bradley for generously providing strains.

This work was supported by funding to C.W. from the National Sciences and Engineering Research Council of Canada.

REFERENCES

1. **Altschul, S. F., W. Gish, W. Miller, E. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
2. **Astill, D. S. J., P. A. Manning, and M. W. Heuzenroeder.** 1993. Characterization of the small cryptic plasmid, pIMVS1, of *Salmonella enterica* ser. Typhimurium. Plasmid **30:**258–267.
3. **Bik, E. M., A. E. Bunschoten, R. D. Gouw, and F. R. Mooi.** 1995. Genesis of the novel epidemic *Vibrio cholerae* O139 strain: evidence for horizontal transfer of genes involved in polysaccharide synthesis. EMBO J. **14:**209–216.
4. **Blakely, G., G. May, R. McCulloch, L. K. Arciszewska, M. Burke, S. T. Lovett, and D. J. Sherrat.** 1993. Two related recombinases are required for site-specific recombination at *dif* and *cer* in *E. coli* K12. Cell **75:**351–361.
5. **Bradley, D. E.** 1980. Determination of pili by conjugative bacterial drug resistance plasmids of incompatibility groups B, C, H, J, K, M, V, and X. J. Bacteriol. **141:**828–837.
6. **Bradley, D. E.** 1984. Characteristics and function of thick and thin conjugative pili determined by transfer-derepressed plasmids of incompatibility groups I_1, I_2, I_5, B, K and Z. J. Gen. Microbiol. **130:**1489–1502.
7. **Cannon, P. M., and P. Strike.** 1992. Complete nucleotide sequence and gene organization of plasmid NTP16. Plasmid **27:**220–230.
8. **Cesarini, G., M. Helmer-Citterich, and L. Castagnoli.** 1991. Control of ColE1 plasmid replication by antisense RNA. Trends Genet. **7:**230–235.
9. **Chan, P. T., H. Ohmori, J.-I. Tomizawa, and J. Lebowitz.** 1985. Nucleotide sequence and gene organization of ColE1 DNA. J. Biol. Chem. **260:**8925–8935.
10. **Colloms, S. D., P. Sykora, G. Szatmari, and D. J. Sherratt.** 1990. Recombination at ColE1 *cer* requires the *Escherichia coli xerC* gene product, a member of the lambda integrase family of site-specific recombinases. J. Bacteriol. **172:**6973–6980.
11. **Cornet, F., I. Mortier, J. Patte, and J.-M. Louarn.** 1994. Plasmid pSC101 harbors a recombination site, *psi*, which is able to resolve plasmid multimers and to substitute for the analogous chromosomal *Escherichia coli* site *dif*. J. Bacteriol. **176:**3188–3195.
12. **Ditta, G., D. Stanfield, D. Corbin, and D. R. Helinski.** 1980. Broad host range DNA cloning system for Gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. Proc. Natl. Acad. Sci. USA **77:**7347–7351.
13. **Dooley, T. P., J. Tamm, and B. Polisky.** 1985. Isolation and characterization of mutants affecting functional domains of ColE1 RNAI. J. Mol. Biol. **186:**87–96.
14. **Figurski, D. H., and D. R. Helinski.** 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided *in trans*. Proc. Natl. Acad. Sci. USA **76:**1648–1652.
15. **Frank, M. M., K. Joiner, and C. Hammer.** 1987. The function of antibody and complement in the lysis of bacteria. Rev. Infect. Dis. **9:**S537–S545.
16. **Hakkaart, M. J. J., P. J. M. van den Elzen, E. Veltkamp, and H. J. J. Nijkamp.** 1984. Maintenance of multicopy plasmid Clo DF13 in *E. coli* cells: evidence for site-specific recombination at *parB*. Cell **36:**203–209.
17. **Hardy, K. G.** 1975. Colicinogeny and related phenomena. Bacteriol. Rev. **39:**464–515.
18. **Hitchcock, P. J., and T. M. Brown.** 1983. Morphological heterogeneity among *Salmonella* lipopolysaccharide chemotypes in silver-stained polyacrylamide gels. J. Bacteriol. **154:**269–277.
19. **Joiner, K. A.** 1985. Studies on the mechanism of bacterial resistance to complement-mediated killing and the mechanism of action of bactericidal antibody. Curr. Top. Microbiol. Immunol. **121:**99–133.
20. **Kawahara, K., T. Hamaoka, S. Suzuki, M. Nakamura, S. Y. Murayama, T. Arai, N. Terakado, and H. Danbara.** 1989. Lipopolysaccharide alteration mediated by the virulence plasmid of *Salmonella*. Microb. Pathog. **7:**195–202.
21. **Keenleyside, W. J., M. Perry, L. MacLean, C. Poppe, and C. Whitfield.** 1994. A plasmid-encoded $rfb_{O:54}$ gene cluster is required for biosynthesis of the O:54 antigen in *Salmonella enterica* serovar Borreze. Mol. Microbiol. **11:**437–448.

21a. **Keenleyside, W. J., and Whitfield, C.** Unpublished data.

22. **Kopecko, D. J., O. Washington, and S. B. Formal.** 1980. Genetic and physical evidence for plasmid control of *Shigella sonnei* form I cell surface antigen. Infect. Immun. **29:**207–214.
23. **Lacatena, R. M., and G. Cesarini.** 1983. Interaction between RNAI and the

primer precursor in the regulation of ColE1 replication. J. Mol. Biol. **170:** 635–650.
24. **Lambert, C. M., C. J. Wrighton, and P. Strike.** 1987. Characterization of the drug resistance plasmid NTP16. Plasmid **17:**26–36.
25. **Lin-Chao, S., and S. N. Cohen.** 1991. The rate of processing and degradation of antisense RNAI regulates the replication of ColE1-type plasmids in vivo. Cell **65:**1233–1242.
26. **Lin-Chao, S., T.-T. Wong, K. J. McDowall, and S. N. Cohen.** 1994. Effects of nucleotide sequence on the specificity of *rne*-dependent and RNase E-mediated cleavages of RNA I encoded by the pBR322 plasmid. J. Biol. Chem. **269:**10797–10803.
27. **Masukata, H., and J. Tomizawa.** 1990. A mechanism of formation of a persistent hybrid between elongating RNA and template DNA. Cell **62:**331–338.
28. **Masukata, H., and J.-I. Tomizawa.** 1984. Effects of point mutations on formation and structure of the RNA primer for ColE1 DNA replication. Cell **36:**513–522.
29. **McDowall, K. J., S. Lin-Chao, and S. N. Cohen.** 1994. A + U content rather than a particular nucleotide order determines the specificity of RNase E cleavage. J. Biol. Chem. **269:**10790–10796.
30. **Morona, R., M. Mavris, A. Fallarino, and P. A. Manning.** 1994. Characterization of the *rfc* region of *Shigella flexneri*. J. Bacteriol. **176:**733–747.
31. **Nijkamp, H. J. J., R. De Lang, A. R. Stuitje, P. J. M. Van Den Elzen, E. Veltkamp, and A. J. Van Putten.** 1986. The complete nucleotide sequence of the bacteriocinogenic plasmid CloDF13. Plasmid **16:**135–160.
32. **Nijkamp, H. J. J., B. van Gemen, M. J. J. Hakkaart, A. J. van Putten, and E. Veltkamp.** 1985. Stable maintenance of plasmid CloDF13: structural and functional relationships between replication control, partitioning, and incompatibility. p. 283–298. *In* D. R. Helinski, S. N. Cohen, D. Clewell, D. Jackson, and A. Hollander (ed.), Plasmids in bacteria. Plenum Press, N.Y.
33. **Patient, M. E., and D. K. Summers.** 1993. ColE1 multimer formation triggers inhibition of *Escherichia coli* cell division. Mol. Microbiol. **9:**1089–1095.
34. **Polisky, B.** 1988. ColE1 replication control circuitry: sense from antisense. Cell **55:**929–932.
35. **Popoff, M. Y., J. Bockemühl, and A. McWhorter-Murlin.** 1992. Supplement 1991 (no. 35) to the Kauffmann-White scheme. Res. Microbiol. **143:**807–811.
36. **Popoff, M. Y., and L. Le Minor.** 1985. Expression of antigenic factor O:54 is associated with the presence of a plasmid in *Salmonella*. Ann. Inst. Pasteur/Microbiol. (Paris) **136B:**169–179.
37. **Poppe, C., and K. A. McFadden.** 1992. Elimination of a 4.5 MDal plasmid converts *S. borreze* to *S. agona*., p. 37–43. Reports and communications of the Symposium on Salmonella and Salmonellosis. World Veterinary Poultry Association, Ploufragan/Saint-Brieuc, France.
38. **Reeves, P. R.** 1992. Variation in O-antigens, niche-specific selection and bacterial populations. FEMS Microbiol. Lett. **100:**509–516.
39. **Reeves, P. R.** 1994. Biosynthesis and assembly of lipopolysaccharide. New Compr. Biochem. **27:**281–317.
40. **Riley, L. W., L. Junio, L. B. Libaek, and G. K. Schoolnik.** 1987. Plasmid-encoded expression of lipopolysaccharide O-antigenic polysaccharide in enteropathogenic *Escherichia coli*. Infect. Immun. **55:**2052–2056.
41. **Rubin, R. H., and I. Weinstein.** 1977. Salmonellosis: microbiologic, pathogenic and clinical features. Stratton Intercontinental Medical Book Corp., New York.
42. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
43. **Sirisena, D. M., P. M. MacLachlan, S.-L. Liu, A. Hessel, and K. E. Sanderson.** 1994. Molecular analysis of the *rfaD* gene, for heptose synthesis, and the *rfaF* gene, for heptose transfer, in lipopolysaccharide synthesis in *Salmonella typhimurium*. J. Bacteriol. **176:**2379–2385.
44. **Smith, N. H., and R. K. Selander.** 1991. Molecular genetic basis for complex flagellar antigen expression in a triphasic serovar of *Salmonella*. Proc. Natl. Acad. Sci. USA **88:**956–960.
45. **Stirling, C. J., S. D. Colloms, J. F. Collins, G. Szatmari, and D. J. Sherratt.** 1989. *xer*B, an *Escherichia coli* gene required for plasmid ColE1 site-specific recombination, is identical to *pepA*, encoding aminopeptidase A, a protein with substantial similarity to bovine lens leucine aminopeptidase. EMBO J. **8:**1623–1627.
46. **Stirling, C. J., G. Szatmari, G. Stewart, M. C. M. Smith, and D. J. Sherratt.** 1988. The arginine repressor is essential for plasmid-stabilizing site-specific recombination at the ColE1 *cer* locus. EMBO J. **7:**4389–4395.
47. **Sturm, S., B. Jann, K. Jann, P. Fortnagel, and K. N. Timmis.** 1986. Genetic and biochemical analysis of *Shigella dysenteriae* 1 O antigen polysaccharide biosynthesis in *Escherichia coli* K-12: structure and functions of the *rfb* gene cluster. Microb. Pathog. **1:**307–324.
48. **Summers, D. K.** 1991. The kinetics of plasmid loss. Trends Biotechnol. **9:**273–278.
49. **Summers, D. K., and D. J. Sherratt.** 1984. Multimerisation of high copy number plasmids causes instability: ColE1 encodes a determinant essential for plasmid monomerization and stability. Cell **36:**1097–1103.
50. **Summers, D. K., and D. J. Sherratt.** 1988. Resolution of ColE1 dimers requires a DNA sequence implicated in the three-dimensional organization of the *cer* site. EMBO J. **7:**851–858.
51. **Tomizawa, J.** 1985. Control of ColE1 plasmid replication: initial interaction of RNAI and the primer transcript is reversible. Cell **40:**527–535.
52. **Tomizawa, J.** 1986. Control of ColE1 plasmid replication: binding of RNA I to RNA II and inhibition of primer formation. Cell **47:**89–97.
53. **Tomizawa, J., and T. Itoh.** 1981. Plasmid ColE1 incompatibility determined by interaction of RNAI with primer transcript. Proc. Natl. Acad. Sci. USA **78:**6096–6100.
54. **Tomizawa, J.-I., H. Ohmori, and R. E. Bird.** 1977. Origin of replication of colicin E1 plasmid DNA. Proc. Natl. Acad. Sci. USA **74:**1865–1869.
55. **Tsai, G. M., and C. E. Frasch.** 1982. A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels. Anal. Biochem. **119:**115–119.
56. **van Putten, A. J., R. de Lang, E. Veltkamp, H. J. J. Nijkamp, P. van Solingen, and J. A. van den Berg.** 1986. Methylation-dependent transcription controls plasmid replication of the CloDF13 *cop-1* (Ts) mutant. J. Bacteriol. **168:**728–733.
57. **Viret, J.-F., S. J. J. Cryz, A. B. Lang, and D. Favre.** 1993. Molecular cloning and characterization of the genetic determinants that express the complete *Shigella* serotype D (*Shigella sonnei*) lipopolysaccharide in heterologous live attenuated vaccine strains. Mol. Microbiol. **7:**239–252.
58. **Warren, G. J., A. J. Twigg, and D. J. Sherratt.** 1978. ColE1 plasmid mobility and relaxation complex. Nature (London) **274:**259–261.
59. **Watanabe, M., and K. N. Timmis.** 1984. A small plasmid in *Shigella dysenteriae* 1 specified one or more functions essential for O-antigen production and bacterial virulence. Infect. Immun. **43:**391–396.
60. **Whitfield, C., and M. A. Valvano.** 1993. Biosynthesis and expression of cell surface polysaccharides in gram-negative bacteria. Adv. Microb. Physiol. **35:**135–246.
61. **Wong, E. M., and B. Polisky.** 1985. Alternative conformations of the ColE1 replication primer modulate its interaction with RNAI. Cell **42:**959–966.
62. **Xiang, S.-H., A. M. Haase, and P. R. Reeves.** 1993. Variation of the *rfb* gene clusters in *Salmonella enterica*. J. Bacteriol. **175:**4877–4884.
63. **Xiang, S.-H., M. Hobbs, and P. R. Reeves.** 1994. Molecular analysis of the *rfb* gene cluster of a D2 *Salmonella enterica* strain: evidence for its origin from an insertion sequence-mediated recombination event between group E and D1 strains. J. Bacteriol. **176:**4357–4365.
64. **Yoshida, Y., N. Okamura, J. Kato, and H. Watanabe.** 1991. Molecular cloning and characterization of form I antigen genes of *Shigella sonnei*. J. Gen. Microbiol. **137:**867–874.
65. **Zuker, M., and P. Steigler.** 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res. **9:**133–148.