

Molecular Cloning, Characterization, and Nucleotide Sequence of the *rfc* Gene, Which Encodes an O-Antigen Polymerase of *Salmonella typhimurium*

L. VINCENT COLLINS[†] AND JIM HACKETT^{‡*}

Department of Microbiology, University of Adelaide,
Adelaide, South Australia 5001, Australia

Received 18 September 1990/Accepted 29 January 1991

The *rfc* gene of *Salmonella typhimurium* was located in a 1.75-kb *Hind*III fragment and restored wild-type lipopolysaccharide synthesis ability to both an older *rfc* point mutant and new *rfc::IS10* mutants. DNA sequencing of the *Hind*III fragment revealed an open reading frame which could encode a protein of 407 amino acids with an M_r of 47,472 and also revealed potential translation signals. Modulator codons accounted for 12.5% of the total codon content, providing a possible explanation for the nondetectability of the protein in subcellular systems. Secondary structure analysis suggested the presence of transmembrane β -sheet structures, implying a possible role for the protein in translocation of hydrophilic O-antigen-containing materials. *Salmonella* strains of groups A, B, and D1 contained *rfc*-homologous DNA, but strains of groups C1, C2, C3, D2, and E2 did not.

Three major genetic regions have been implicated as dedicated to the biosynthesis of the polysaccharide component of the lipopolysaccharide (LPS) of *Salmonella typhimurium*. The *rfa* cluster of genes is involved in core synthesis, and the *rfb* region is involved in the synthesis of the O-antigenic tetrasaccharide units (26). In addition, it is thought that genes of both the *rfa* and *rfb* clusters are involved in the transport and ligation of core-O-antigen complexes (26). The *rfc* region is thought to encode a polymerase responsible for the linking of the O-antigen tetrasaccharide units into long chains, giving rise to typical smooth LPS (26). Mutants of *S. typhimurium* which are defective in the O-antigen polymerase because of a mutation at the *rfc* locus produce LPS structures termed semirough (SR), which have at most one O-antigenic tetrasaccharide unit attached to any given core unit (30, 56). A bacteriophage sensitivity pattern characteristic of SR mutants has been described previously; such strains of *S. typhimurium* were resistant to the smooth LPS-specific phages 9NA and P22, sensitive to phage FO, and resistant to all phages specific for rough LPS (LPS without attached O antigen) except P22I (55).

Mutants of the SR LPS phenotype have been identified in *Salmonella* groups B and E (22, 30). In *S. typhimurium* (group B), the *rfc* locus has been located between *gal* and *trp* (30, 44), corresponding to a map position between 18 and 34 min. An analogous locus in some *Salmonella* strains of group D has been postulated on the basis of the observation that in hybrids between *Salmonella* groups B (O-4, 5, 12) and D (O-9, 12), involving transfer of *rfb* genes, the polymerase of one group could polymerize the O units of the other (24, 31, 51).

This paper describes the cloning of the *rfc* gene of *S. typhimurium* by complementation of *rfc* defects in spontaneous and *IS10*-derived SR mutants. A physical map of the DNA region encoding the *rfc* gene is presented, and attempts to visualize the gene product(s) are described. The distribution of *rfc*-homologous DNA in other *Salmonella* strains is also determined. The DNA fragment containing the gene is sequenced, and a putative *rfc* gene is identified.

MATERIALS AND METHODS

Materials. All reagents were of Analar grade. Enzymes active on DNA were obtained from either Amersham or Boehringer Mannheim and were used as directed. The four deoxyribonucleotide triphosphates (dATP, dCTP, dGTP, and dTTP), X-Gal (5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside) and isopropyl- β -D-thiogalactopyranoside were purchased from Boehringer Mannheim. M13 sequencing primer and [α -³²P]dCTP, at a specific activity of 1,700 Ci/mM, were obtained from BRESA (Adelaide, Australia). The -40 sequencing primer was obtained from New England BioLabs. L-[³⁵S]methionine (1,270 Ci/mM), [³⁵S]dATP (>1,000 Ci/mM), and Sequenase were purchased from Amersham.

Bacterial strains and media. The strains used are listed in Table 1. Transposons Tn1725 and Tn5 were contained on temperature-sensitive R factors in strain RU2901 (Tn1725) (50) or WR6016 (Tn5). Bacteria were routinely grown at 37°C in Difco nutrient broth at double strength, with 5 g of NaCl per liter, or on Difco nutrient agar. Antibiotics were added, when appropriate, at the following concentrations: ampicillin, 25 μ g/ml; chloramphenicol, 25 μ g/ml; kanamycin, 50 μ g/ml; tetracycline, 10 μ g/ml. Methionine assay medium (Difco) was reconstituted according to the manufacturer's instructions.

Bacteriophage sensitivity tests. Bacteria were grown to stationary phase and swabbed across a plate. Bacteriophages were spotted (ca. 10⁶ PFU in 5 μ l of nutrient broth) onto the bacterial swabs, and the plates were incubated at

* Corresponding author.

[†] Present address: Department of Biology, Washington University, St. Louis, MO 63130.

[‡] Present address: Department of Animal Sciences, University of Adelaide, Waite Agricultural Research Institute, Glen Osmond, South Australia 5064, Australia.

TABLE 1. Bacterial strains and plasmids used in this work

Strain or plasmid	Genotype or relevant characteristics	Relevant phenotype	Source or reference
<i>E. coli</i> K-12			
DH1	<i>gyrA96 recA1 relA1 endA thi-1 hsdR17 supE44 λ⁻</i>		27
DS410	<i>minA minB rpsL</i>	Minicell producer	7
K38	HfrC	Host for T7 polymerase-promoter system	37
JM101	<i>supE Δ(lac-proAB) thi-1/F'(traD36 proAB lacI^q lacZ ΔM15)</i>	Host for M13 propagation	39
RU2901	Rts1::Tn1725	Donor of Tn1725 (Cm ^r)	50
WR6016	<i>thr leu thi recA/F'ts114lac::Tn5</i>	Donor of Tn5	41
CSR603	<i>uvrA uvrB recA</i>	Maxicell strain	38
<i>S. typhimurium</i> ^a			
C5	Wild type		IMVS ^b
SL901	<i>rfc-497 metA22 trpC flaA66 H1-b H2-e,n,x Fels 2⁻</i>	SR LPS produced	55
TV119	<i>rfb</i>	Rough LPS produced	46
LV56	As SL901, <i>galE</i>	SR LPS produced, galactose resistant	This study
LV203	As LV56, <i>recA</i>	As LV56, but recombination defective	This study
LV242	LV203/pADE200 <i>lamB⁺</i> Km ^r	As LV203, but λ sensitive and cryptic plasmid free	This study
TT521	<i>recA1 rpsL srl-202::Tn10</i>	Donor of <i>recA</i> allele	20
TT628	<i>pyrC7 strA1/F'ts114lac zzf-21::Tn10</i>	Donor of Tn10	20
J208	<i>rfc::IS10 F'lac zzf-21::Tn10</i>	Rfc	This study
J208-B1	<i>rfc::IS10</i>	Rfc	This study
J208-B2	<i>rfc::IS10</i>	Rfc	This study
Plasmids			
pHC79	Carries Ap ^r and Tc ^r , cosmid		16
pUC18	Carries Ap ^r		53
pSUP401	Carries Cm ^r and Km ^r		43
pGP1-2	Can express T7 polymerase		47
pBluescript KS ⁺	Carries T7 promoter and Ap ^r		Stratagene, Inc.
M13mp18	M13 cloning vector		28
M13mp19	M13 cloning vector		28
pPR11	<i>tolC::Tn10</i> ; source of IS10 DNA		29
pAMH70	<i>lamB⁺</i> , carries Ap ^r		14
pLG339	Source of Km ^r cassette		45
pADE200	<i>lamB⁺</i> , carries Km ^r		This study
pADE206	pUC18/ <i>rfc⁺</i> , carries Ap ^r		This study
pADE207	pSUP401/ <i>rfc⁺</i> , carries Cm ^r		This study

^a Other *Salmonella* strains used were *S. paratyphi* A, *S. typhi*, *S. derby*, *S. dublin*, *S. chester*, *S. strasbourg*, *S. newington*, *S. cholerae-suis*, *S. bonariensis*, and *S. virginia*. All were from IMVS.

^b IMVS: Institute of Medical and Veterinary Science, Adelaide, Australia.

37°C for 16 h. Strains were scored bacteriophage sensitive if a clear zone of lysis appeared.

Calculations of the efficiency of plaquing of bacteriophage on given strains were effected by mixing of 0.1-ml amounts of phage stock dilutions (in nutrient broth) with 0.1-ml amounts of bacteria prepared as described above. After 10 min at 37°C, the mixture was spread on plates. Incubation at 37°C for 16 h followed; plaques were then counted.

Construction of a lambda-sensitive *rfc* strain. Plasmid pAMH70 (*lamB⁺* [14]) was cut with *ScaI* and ligated with the end-filled 1.4-kb *MluI* fragment (which encodes Km^r) of pLG339 (45) to give pADE200. Maintenance of pADE200 and the cosmid vector pHC79 (Ap^r) (16) in the same cell was possible since the Ap^r determinant of pAMH70 had been inactivated during the cloning step.

The *rfc* strain SL901 was subjected to genetic manipulation in order to facilitate its use as a recipient in cosmid bank construction and to render it more receptive as a recipient in transformation than the unmodified strain. Initially, a *galE* mutation was introduced into SL901 (14) by initial selection for resistance to FO, followed by screening for susceptibility to galactose-induced lysis, which is a characteristic of *galE*

mutants (8). Subsequently, a galactose-resistant derivative of the *galE* strain was obtained by selection on nutrient agar with galactose (0.5% [wt/vol]). This galactose-resistant strain (LV56) retained the ability to make SR LPS in the presence of galactose and accordingly did not completely lack *galT* or *galK* activities (8). Strain LV56 then received, by phage ES18-mediated cotransduction with *srl-202::Tn10*, the *recA* allele of strain TT521. A Tc^s derivative, which had apparently lost the Tn10 element but remained *recA*, was selected on fusaric acid-containing medium (1) to give strain LV203.

Strain LV203 contained a large plasmid characteristic of *S. typhimurium* (19). In early experiments with this strain, problems were encountered with recombination between the resident plasmid and transformed plasmids and with purification of cloned DNA from the strain. Accordingly, strain LV203 was cured of the resident plasmid (3) and then received, by transformation, pADE200 (*lamB⁺*) to give the final strain, LV242. Strain LV242 was lambda sensitive, while SL901/pADE200 was lambda resistant.

None of the above alterations to the original SL901 strain altered the SR LPS phenotype. The *rfc* strain SL901 was

confirmed to have the phage sensitivity profile characteristic of SR strains (55), being 9NA^r, P22^r, P221^s, and FO^s. Strains LV203 and LV242 were 9NA^r, P22^r, P221^s, and FO^r on nutrient agar with glucose but were 9NA^r, P22^r, P221^s, and FO^s on nutrient agar with galactose.

Insertional inactivation of the *rfc* locus of C5. To facilitate complementation tests with a cloned *rfc* gene, IS10-generated SR insertion mutants of *S. typhimurium* C5 were isolated. An F' factor, which incorporated transposon Tn10 (F'114ts *lac*⁺ *zcf-21::Tn10*) was conjugated from strain TT628 into *S. typhimurium* C5, with selection for Tc^r on minimal medium at 30°C. Individual Tc^r colonies were screened for resistance to phage 9NA and for sensitivity to phage FO at 37°C. In this manner, strain J208 was identified as having an SR phage phenotype (9NA P22^r FO^s P221^s) but was found to be Lac⁺ on MacConkey lactose medium. This indicated that although an *rfc* locus had been inactivated, the F' factor had not been lost at the nonpermissive temperature. Derivatives of J208 which were Tc^s were isolated by plating on fusaric acid-containing medium (1). Two individual Tc^s isolates of J208, J208-B1, and J208-B2, were examined and were found to be Lac⁻, indicating that the F' factor had been lost concomitant with the loss of the tetracycline resistance of the Tn10 element. The *rfc* phenotype (SR phage type) was retained in both J208-B1 and J208-B2. As IS10-right (IS10-R) is a fully functional transposon (35), the above results suggested that IS10-R transposition from the F'114ts *lac*⁺ *zcf::Tn10* into the *rfc* locus might have occurred, as was confirmed by probing with a 2.8-kb *AccI* fragment of transposon Tn10 (from plasmid pPR11) containing part of the IS10-R element (29, 54). The probe detected a 3.2-kb *HindIII* fragment in J208-B1 and J208-B2 but did not react with *HindIII*-digested DNA of C5.

Attempts to detect plasmid-encoded polypeptides. Insertion-encoded polypeptides were sought in minicells containing various plasmids (6), in *Escherichia coli* K-12 maxicells after UV irradiation (38), or in an in vitro system for expression of plasmid genes (5, 57). The 1.75-kb *HindIII* fragment was cloned, in either orientation, into pBluescript KS⁺, and cells bearing the clones and the T7 polymerase-encoding plasmid pGP1-2 were treated to cause induction of transcription from the T7 promoter (47).

DNA methods. The DNA methods used are described by Maniatis et al. (27). These include the preparation of chromosomal and plasmid DNA, the cutting of DNA with restriction endonucleases, ligation of DNA, visualization of DNA on agarose gel electrophoresis, and Southern blotting. High-stringency washing was usually employed (twice with shaking at 37°C for 30 min in 2× SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate] containing 0.1% [wt/vol] sodium dodecyl sulfate [SDS] and then two further washes in 0.1× SSC plus 0.1% [wt/vol] SDS at 65°C). For hybridization tests at lower stringency, the final two washes were performed at 50 or 55°C. Probes were labeled with [α -³²P]dCTP by using random oligonucleotide primers. For DNA sequencing, the dideoxy-chain termination procedure of Sanger et al. (40) was modified to encompass the use of Sequenase (modified T7 DNA polymerase) in place of the Klenow enzyme (48). A 17-mer (-20) M13 sequencing primer was usually employed. A 17-mer (-40) M13 sequencing primer was used to determine the sequence of the 1.75-kb *HindIII* fragment at the junction with the polylinker in M13mp18. Synthetic deoxyoligonucleotide sequences were used as primers to sequence DNA regions not overlapped by the M13 subclones of pADE206.

The initial cosmid cloning of *S. typhimurium* DNA in-

TABLE 2. Efficiency of plaquing of various bacteriophages on various strains

Strain	Efficiency of plaquing ^a of:			
	P22	9NA	FO	P221
C5	<u>1</u>	<u>1</u>	<u>1</u>	<0.25 × 10 ⁻⁶
LV242	<0.25 × 10 ⁻⁸	<0.25 × 10 ⁻⁸	5	<u>1</u>
J208-B1	<0.25 × 10 ⁻⁸	<0.25 × 10 ⁻⁸	4	<u>1</u>
LV242/pADE206	1	1	1	<0.25 × 10 ⁻⁶
J208-B1/pADE206	1	1	1	<0.25 × 10 ⁻⁶

^a Bacteria were mixed with phage at various dilutions, and plaques were counted after overnight incubation. The efficiencies of plaquing of P22, 9NA, and FO on C5 were taken as unity (underlined). Similarly, the titer of P221 on LV242 was the reference efficiency of plaquing for this phage (underlined).

involved the ligation of partially *Sau3A*-cut total bacterial DNA with *Bam*HI-cut pHC79. The packaging mixture was obtained from Amersham.

Oligodeoxynucleotides were synthesized with reagents purchased from Applied Biosystems or Ajax Chemicals (acetonitrile). Synthesis was performed on an Applied Biosystems 381A DNA synthesizer. Oligonucleotides were routinely of a purity such that no further purification was required.

SDS-PAGE. SDS-polyacrylamide gel electrophoresis (PAGE) was performed on 11 to 20% (wt/vol) polyacrylamide gradients as described previously (23). LPS for analysis by SDS-PAGE followed by silver staining was prepared and analyzed as previously described (17).

Nucleotide sequence accession number. The sequence of the 1,750-bp *HindIII* *rfc*⁺ fragment reported herein has been deposited with GenBank under accession number M60066.

RESULTS AND DISCUSSION

Cloning of the *S. typhimurium* *rfc* locus. Strain LV242 (lambda sensitive and *rfc*, Table 1) was used as the recipient in transfection of an *S. typhimurium* cosmid bank prepared with vector pHC79. Individual cosmid clones from the bank were screened on nutrient agar with galactose for sensitivity to bacteriophage 9NA. Of 300 clones screened, one was 9NA^s, P22^s, FO^s, and P221^r. This phage phenotype was consistent with the production of smooth LPS (55), and the contained cosmid was therefore considered to complement the genetic defect in the SR mutant. *HindIII* fragments of this cosmid were subcloned into the *HindIII* sites of pUC18 or pSUP401. The same 1.75-kb fragment in either vector (pADE206 and pADE207, respectively) complemented the *rfc* defects in strains LV242 and J208-B1 (*rfc::IS10*) (Table 1), as judged by phage efficiency of plaquing tests (Table 2). A restriction map of the 1.75-kb *HindIII* fragment is shown in Fig. 1.

Plasmid pADE206 was subjected to transposon insertion mutagenesis with either Tn1725 or Tn5, and the *rfc* status of the plasmids obtained was assessed by phage sensitivity observations after transformation into SL901 (*rfc*) and J208-B1. The locations of individual transposon insertion points were mapped (Fig. 1). Plasmids pADE212 and pADE213 both contained Tn1725 insertions which did not affect *rfc* activity, while noninactivating Tn5 insertions were mapped in plasmids pADE208 and pADE210. Transposon insertions which inactivated the *rfc* gene were located in plasmids pADE214, pADE215, pADE216, and pADE217 (all Tn1725 inserts) and plasmids pADE209 and pADE211 (Tn5 inserts) (Fig. 1). In summary, transposon insertions to the left of the

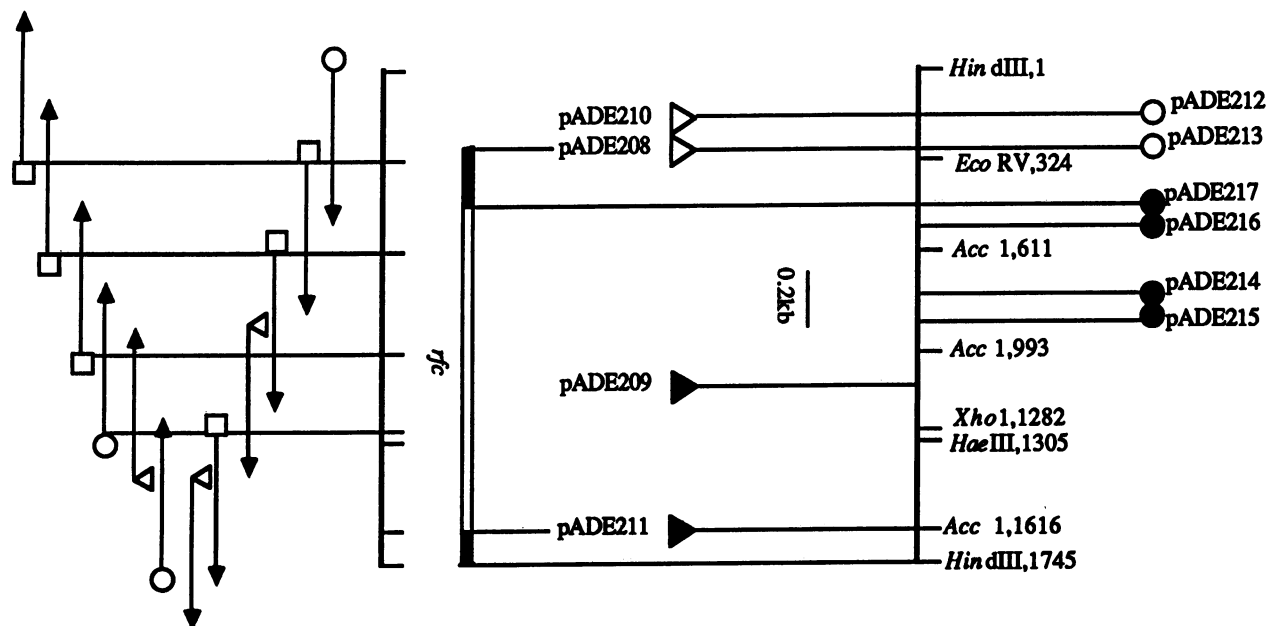


FIG. 1. Restriction enzyme and transposon insertion map of the 1.75-kb *Hind*III fragment of pADE206 and sequencing strategy for the gene. The site coordinates (see Fig. 5) are the first bases in the restriction sites, regardless of the enzyme cleavage points. Plasmids carrying transposon insertions located in this fragment are numbered, and the transposon insertion sites are indicated as follows: ○, *Tn*1725 insertions not affecting *rfc*; ●, *Tn*1725 insertions inactivating *rfc*; ▷, *Tn*5 insertions not affecting *rfc*; ►, *Tn*5 insertions inactivating *rfc*. The maximum and minimum limits of the *rfc* gene are shown. Indicated restriction fragments of the insert were subcloned in either M13mp18 or M13mp19 for dideoxy sequencing. The sequencing start point of each clone is indicated by a circle for M13mp18 derivatives and by a square for M13mp19 derivatives. The direction and extent of the sequencing carried out on each clone are represented by arrows. The DNA regions preceded by a triangle were sequenced with a specifically synthesized oligodeoxynucleotide primer.

single *Eco*RV site had no effect on the activity of the *rfc* gene, whereas transposon insertions to the right of the *Eco*RV site all resulted in inactivation of the *rfc* gene (Fig. 1). The transposon insertion mutagenesis therefore indicated that the size of the *rfc* gene was between 1.1 and 1.5 kb.

The cloned 1.75-kb *Hind*III fragment of pADE206 was used to probe *Hind*III-digested total DNA of C5, J208, J208-B1, and J208-B2 (Fig. 2). As expected, a 1.75-kb *Hind*III fragment of C5 hybridized strongly to the probe (Fig. 2, lane 1). A *Hind*III fragment of 3.2 kb with homology to the probe was observed for each of the mutants (Fig. 2, lanes 2 to 4), in line with the suggestion (see Materials and Methods) that an *IS*10 element was inserted in the chromosomal *rfc* region in each case.

LPS phenotype analysis by SDS-PAGE. Silver-stained SDS-PAGE gels of the LPS of control strains, the *rfc* strains, and their derivatives are shown in Fig. 3. The smooth strain C5 (lane 2) showed the typical ladderlike pattern, presumably representing stepwise elongation of core moieties by addition of O-antigen units. The *rfb* mutant TV119 (lane 1) showed low-molecular-weight LPS species corresponding to core LPS alone, while in the *rfc* mutant SL901 (lane 3), the TV119 doublet did not appear but a doublet of lower mobility was seen (12, 33). Strain J208-B1 (lane 5) displayed the same SR-type LPS as SL901. Strains SL901/pADE206 (lane 4) and J208-B1/pADE206 (lane 6) made LPS similar to that of wild-type *S. typhimurium* C5. The LPS of J208-B1/pADE215 (*rfc*) (lane 7) was identical to that of J208-B1, while J208-B1/pADE212 (*rfc*⁺) (lane 8) made smooth LPS.

In summary, the *rfc* status of strains, as shown by SDS-PAGE LPS profiles, was in agreement with the indications given by the phage sensitivity profiles.

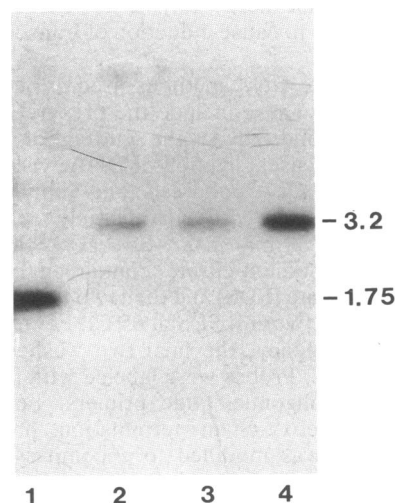


FIG. 2. *IS*10 insertions in the C5 *rfc* gene. Whole DNA of C5 (lane 1), J208 (lane 2), J208-B1 (lane 3), and J208-B2 (lane 4) was digested with *Hind*III and electrophoresed on a 0.8% (wt/vol) agarose gel. Following transfer to nitrocellulose, the filter was probed for homology with the [α -³²P]dCTP-labeled 1.75-kb *Hind*III fragment of pADE206, washed, and subjected to autoradiography. The numbers on the right are the sizes, in kilobases, of the indicated fragments.

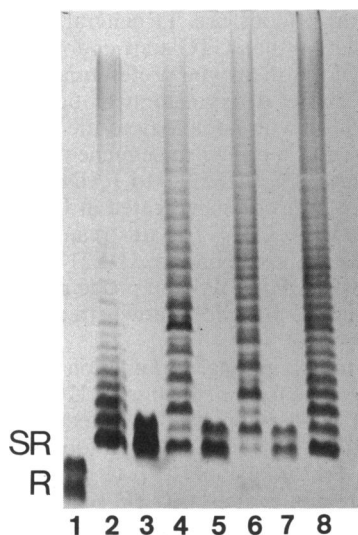


FIG. 3. LPS profiles of various strains in silver-stained SDS-PAGE. Bacteria were lysed with SDS and treated with proteinase K, and the residual material was analyzed by SDS-PAGE with LPS-specific silver staining. Lanes and relevant genotypes: 1, TV119 (*rfb*); 2, C5 (wild type); 3, SL901 (*rfc*); 4, SL901/pADE206 (*rfc*⁺); 5, J208-B1 (*rfc*::IS10); 6, J208-B1/pADE206 (*rfc*⁺); 7, J208-B1/pADE215 (*rfc*); 8, J208-B1/pADE212 (*rfc*⁺). R, Location of rough (core) LPS; SR, location of semirough LPS.

Distribution of the *rfc* locus among *Salmonella* strains. A 0.4-kb *AccI* fragment internal to the *rfc* region (Fig. 1) was purified from pADE206, labeled, and used for high-stringency probing of *AccI*- and *EcoRV*-digested chromosomal DNA of various strains (Fig. 4). Homologous *AccI* fragments identical in size to the 0.4-kb *AccI* probe were detected in *Salmonella paratyphi* A (lane 3), *Salmonella typhi* (lane 4), *Salmonella derby* (lane 5), *Salmonella dublin* (lane 6), and *Salmonella chester* (lane 7), in addition to the parental *S. typhimurium* (lane 2), but no homologous DNA was seen in *Salmonella strasbourg* (lane 8), *Salmonella newington* (lane 9), or *E. coli* K-12 DH1 (lane 1). Similarly, homology with the probe was observed for the *EcoRV*-cut total DNA of *S. typhimurium* (lane 11), *S. paratyphi* A (lane 12), *S. typhi* (lane 13), *S. derby* (lane 14), *S. dublin* (lane 15) and *S. chester* (lane 16), although the hybridizing fragments differed in size among strains. The *EcoRV*-generated DNA fragments of *S. strasbourg* (lane 17), *S. newington* (lane 18), and *E. coli* K-12 DH1 (lane 10) again showed no homology to the *rfc* probe. In summary, the cloned *rfc* gene showed homology with the genomic DNA of *Salmonella* strains of groups A, B, and D1 but not of groups D2 or E2. The cloned *rfc* gene did not probe total DNA from *Salmonella choleraesuis* (group C1), *Salmonella bonariensis* (group C2), or *Salmonella virginia* (group C3) (data not shown). These results did not change when the stringency of hybridization was reduced by effecting the final two washes at 50 or 55°C.

These results suggest that among *Salmonella* strains of serogroups A, B, and D1, the *rfc* gene may be conserved, in line with previous work implying that the O-antigen polymerases of *Salmonella* strains of groups B and D1 accepted LPS units of either serotype as substrates for polymerization (24, 31, 51). In contrast, a group E *Salmonella* serovar appeared not to express an enzyme capable of polymerization of LPS subunits characteristic of a group B serovar (32). Similarly, hybrids in which the group B *rfb* region was

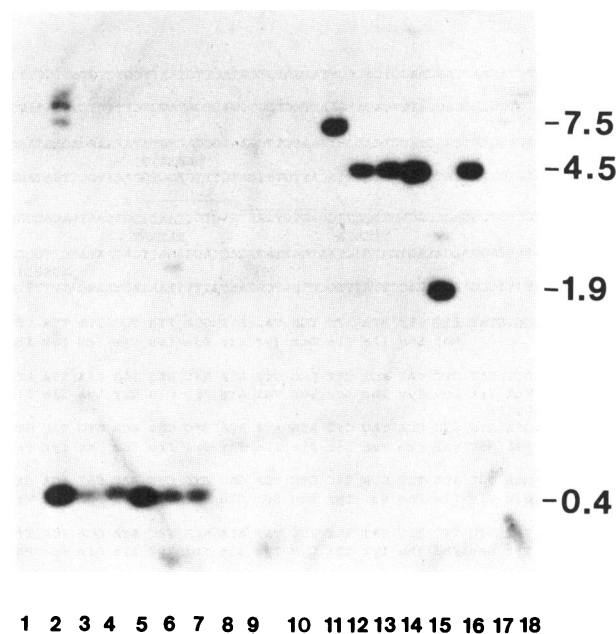


FIG. 4. Homology of cloned *S. typhimurium rfc* with DNA from other *Salmonella* strains. The internal 0.4-kb *AccI* fragment of pADE206 was labeled with [α -³²P]dCTP by oligonucleotide primer labeling. This probe was used in Southern hybridization to test for homology with *AccI*- and *EcoRV*-generated restriction fragments of total DNA of *E. coli* K-12 and a variety of *Salmonella* strains. Lanes: 1 and 10, *E. coli* K-12; 2 and 11, *S. typhimurium*; 3 and 12, *S. paratyphi* A; 4 and 13, *S. typhi*; 5 and 14, *S. derby*; 6 and 15, *S. dublin*; 7 and 16, *S. chester*; 8 and 17, *S. strasbourg*; 9 and 18, *S. newington*. *AccI* was used for lanes 1 to 9; *EcoRV* was used for lanes 10 to 18. The numbers on the right are the sizes, in kilobases, of the indicated fragments.

transferred into *Salmonella* strains of group C1 or C2 did not express polymerized LPS of the B serotype (25, 30).

***rfc* nucleotide sequence determination and analysis.** The entire 1.75-kb *HindIII* fragment of pADE206 was cloned in both directions into the *HindIII* site of M13mp18 (Fig. 1). In addition, DNA fragments of pADE206 were subcloned in the multiple cloning site of either M13mp18 or M13mp19 (53) (Fig. 1). The complete nucleotide sequence of the fragment was determined in both directions (Fig. 5). Potential transcriptional and translational signals are described below.

Analysis of the nucleotide sequence revealed a potential promoter structure between nucleotides (nt) 278 and 306 (Fig. 5). Within this segment, a potential -10 region reads TAatAT (4 of 6 bases as the consensus TATAAT) (15, 36) and a potential -35 region reads TaGAtA (4 of 6 bases identical to the consensus TTGACA) (15). The spacing between these putative -10 and -35 sequences is 17 bp, which is optimum for promoter activity (15, 36). The spacer sequence between the -10 and -35 regions displays two direct repeats of the 10-bp series GAGTGTTAAT, the second of which overlaps the putative -10 region (Fig. 5).

A potential open reading frame (ORF), ORF1, with the translational start codon ATG at nt 461 is preceded by a potential ribosome-binding site (ctaaAGGc) (10, 42) at a spacing of 5 nt. A second possible ORF (ORF2) starts with GTG at nt 641 and is preceded by a putative ribosome-binding site (gtGcAGaa) at a spacing of 7 nt (Fig. 5). ORF1 and ORF2 read in the same frame and are predicted to encode proteins with M_r s of 47,472 and 40,511, respectively.

TABLE 3. Codon usage within the presumed *rfc* gene

Codon ^a - amino acid	No. of times used	% of total codons	Codon ^a - amino acid	No. of times used	% of total codons
TTT-Phe	27	6.6	CTT-Leu	12	2.9
TTC-Phe	19	4.7	CTC-Leu	2	0.5
TTA-Leu	11	2.7	CTA*-Leu	5	1.2
TTG-Leu	7	1.7	CTG-Leu	4	1.0
TCT-Ser	5	1.2	CCT-Pro	8	2.0
TCC-Ser	5	1.2	CCC-Pro	1	0.2
TCA-Ser	11	2.7	CCA-Pro	6	1.5
TCG-Ser	3	0.7	CCG-Pro	2	0.5
TAT-Tyr	18	4.4	CAT-His	4	1.0
TAC-Tyr	18	4.4	CAC-His	0	0.0
TAA	0	0.0	CAA-Gln	1	0.2
TAG	0	0.0	CAG-Gln	3	0.7
TGT-Cys	2	0.5	CGT-Arg	1	0.2
TGC-Cys	4	1.0	CGC-Arg	0	0.0
TGA	0	0.0	CGA*-Arg	0	0.0
TGG-Trp	3	0.7	CGG*-Arg	0	0.0
ATT-Ile	17	4.2	GTT-Val	12	2.9
ATC-Ile	2	0.5	GTC-Val	5	1.2
ATA*-Ile	24	5.9	GTA-Val	8	2.0
ATG-Met	12	2.9	GTG-Val	6	1.5
ACT-Thr	5	1.2	GCT-Ala	10	2.5
ACC-Thr	3	0.7	GCC-Ala	0	0.0
ACA-Thr	9	2.2	GCA-Ala	11	2.7
ACG-Thr	3	0.7	GCG-Ala	3	0.7
AAT-Asn	11	2.7	GAT-Asp	6	1.5
AAC-Asn	6	1.5	GAC-Asp	1	0.2
AAA-Lys	17	4.2	GAA-Glu	9	2.2
AAG-Lys	4	1.0	GAG-Glu	7	1.7
AGT-Ser	2	0.5	GGT-Gly	4	1.0
AGC-Ser	3	0.7	GGC-Gly	3	0.7
AGA*-Arg	9	2.2	GGA*-Gly	5	1.2
AGG*-Arg	3	0.7	GGG*-Gly	5	1.2

^a Modulating codons (see text) are marked with asterisks.

the *rfc* gene (18) was calculated to be 0.42. This value also places the *rfc* gene in the range of poorly expressed genes.

The G+C content of the 1,750-bp *Hind*III fragment was 34.9%, and that of the putative Rfc-encoding region was

33.5%. Considering the average G+C content of *Salmonella* strains, 52%, the *rfc* gene is extremely A+T rich. Similar A+T richness has been observed for the *S. typhimurium* *rfbJ*, *rfbE*, and *rfbS* genes; it has been proposed that these genes evolved outside of the species *Salmonella* (52). The percentage of A+T in the *rfc* gene promoter (nt 278 to 460) region is 67.2%. A zone of exceptionally high A+T content (78%) exists between nt 420 and the start codon at nt 461. It is possible that this region represents a site for the binding of regulatory proteins.

Analysis of plasmid-encoded polypeptides. Insert-encoded labeled polypeptides of pADE206 and pADE207 could not be visualized when the plasmids were present in *E. coli* K-12 minicells or in a maxicell strain of *E. coli* K-12 or when translation products were sought in a cell-free *E. coli* K-12-derived coupled transcription-translation system (see Materials and Methods). In all of these systems, the β -lactamase or acetyltransferase proteins encoded by control plasmids were clearly visible (data not shown). The 1.75-kb *Hind*III fragment was inserted behind the T7 phage promoter in vitro. When *E. coli* K38 was transformed with this plasmid, followed by induction of the T7 polymerase from pGP1-2 in the same cell, no detectable labeled translation product was seen. A possible explanation for these results is suggested by the relative abundance of rare codons in the presumptive coding sequence (see above). In addition, it is possible that protein synthesis systems which use *E. coli* K-12 host strains or subcellular components of *E. coli* K-12 are not suitable for expression of the protein product of the *rfc* gene, if *Salmonella*-specific cofactors or positive regulatory elements are required for gene expression. When the entire cloned *rfb* region of *S. typhimurium* was introduced into *E. coli* K-12, biosynthesis of the polymerized O-polysaccharide of *S. typhimurium* was not detected, since *rfc* function was not provided by the *E. coli* K-12 strain (2). The Southern hybridization data (see above) indicated that there were no DNA sequences in *E. coli* K-12 homologous to the *rfc* gene. Since *rfc* performs no known function in *E. coli*, it is possible that expression of the *rfc* gene of *S. typhimurium* is suboptimal in *E. coli*.

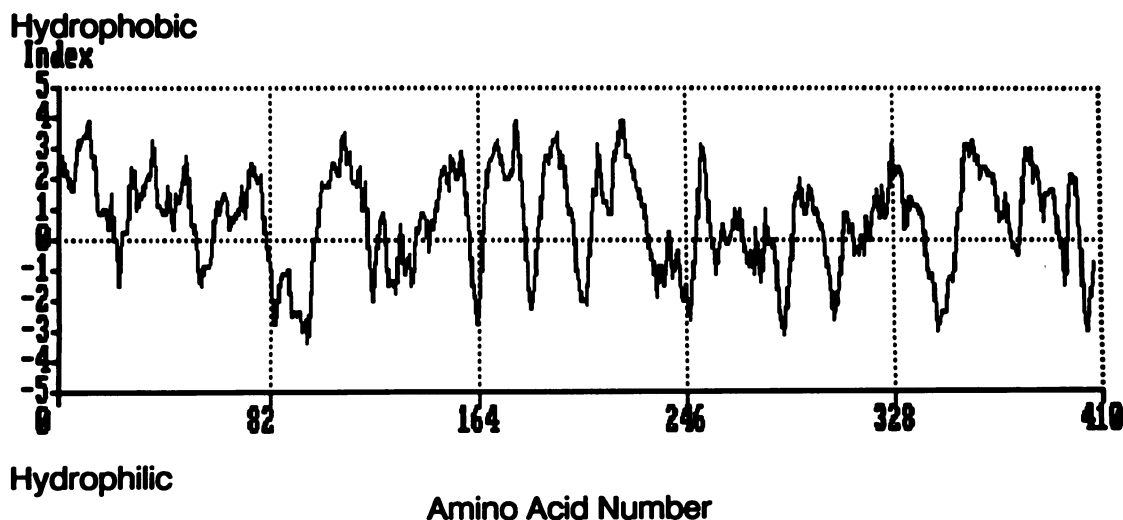


FIG. 6. Hydropathic plot of the Rfc protein. The 407-amino-acid sequence of the predicted Rfc protein was analyzed according to the method of Kyte and Doolittle (21) with a window of 9 amino acids. The average hydropathic index for this protein is 0.7.

The putative Rfc protein. In the following, the protein of a predicted M_r of 47,472 is regarded as the Rfc protein.

The hydropathic plot of Rfc, made according to the methods of Kyte and Doolittle (21) (Fig. 6), indicated that the overall protein was hydrophobic in character, with a mean index of +0.7. Many alternating hydrophobic-hydrophilic domains were indicative of spatial phase transitions for various regions of the protein. The secondary structure of the 407-amino-acid Rfc was predicted by the methods of Chou and Fasman (4). The structure embodied a predominance of β -sheet structure and β -turn- β segments with very few α -helical sections. The presence of several hydrophobic domains, interposed with hydrophilic stretches, in the predicted secondary structure was indicative of transmembrane domains. These regions were composed almost exclusively of β -sheet structure with several β -turn- β domains. The most common configuration of transmembrane protein is that of hydrophobic α -helical bundles. However, bacterial porins, which form aqueous channels in the outer membrane, are made up of an assembly of β -sheets (34). A macromolecular structure consisting of parallel interacting transmembrane β -sheets is conceivable for the Rfc protein but is highly speculative until the existence and cellular localization of the protein are confirmed.

Computer homology searches of both the nucleotide and amino acid sequences, using the Genbank and EMBL Gene/Protein Sequence Databases, failed to uncover any prokaryotic genes with significant homology to *rfc*.

ACKNOWLEDGMENTS

L.V.C. thanks the University of Adelaide for financial support. J.H. is supported by the National Health and Medical Research Council of Australia.

We thank Helena Cerin and Jan Gunter for assistance in the laboratory.

REFERENCES

- Bochner, B. R., H. Huang, G. L. Schieven, and B. N. Ames. 1980. Positive selection for loss of tetracycline resistance. *J. Bacteriol.* **143**:926-933.
- Brahmbhatt, H. N., P. Wyk, N. B. Quigley, and P. R. Reeves. 1988. Complete physical map of the *rfb* gene region cluster encoding biosynthetic enzymes for the O antigen of *Salmonella typhimurium* LT2. *J. Bacteriol.* **170**:98-102.
- Cerin, H., and J. Hackett. 1989. Molecular cloning and analysis of the incompatibility and partition functions of the virulence plasmid of *Salmonella typhimurium*. *Microb. Pathog.* **7**:85-99.
- Chou, P. Y., and G. D. Fasman. 1974. Prediction of protein conformation. *Biochemistry* **13**:222-245.
- Collins, J. 1979. Improved *in vitro* transcription/translation of plasmid encoded genes. *Gene* **6**:29-42.
- Dougan, G., and M. Kehoe. 1984. The minicell system for studying expression from plasmid DNA. *Methods Microbiol.* **17**:233-257.
- Dougan, G., and D. J. Sherratt. 1977. The transposon TnI as a probe for studying ColEI structure and function. *Mol. Gen. Genet.* **151**:151-160.
- Fukasawa, T., and H. Nikaido. 1961. Galactose-sensitive mutants of *Salmonella typhimurium*. II. Bacteriolysis induced by galactose. *Biochim. Biophys. Acta* **48**:470-483.
- Gilbert, W. 1976. Starting and stopping sequences for the RNA polymerase, p. 193-205. In R. Losick and M. Chamberlin (ed.), *RNA polymerase*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Gold, L., D. Pribnow, T. Schneider, S. Shinedling, B. W. Singer, and G. Stormo. 1981. Translational initiation in prokaryotes. *Annu. Rev. Microbiol.* **35**:365-403.
- Gold, L., and G. Stormo. 1987. Translational initiation, p. 1302-1307. In F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umberger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
- Goldman, R. C., and L. Leive. 1980. Heterogeneity of antigenic side-chain length in lipopolysaccharide from *Escherichia coli* O111 and *Salmonella typhimurium* LT2. *Eur. J. Biochem.* **107**:145-153.
- Grosjean, H., and W. Fiers. 1982. Preferential codon usage in prokaryotic genes: the optimal codon-anti-codon energy and the selective codon usage in efficiently expressed genes. *Gene* **18**:199-209.
- Harkki, A., H. Karkku, and E. T. Palva. 1987. Use of lambda vehicles to isolate *ompC-lacZ* fusions in *Salmonella typhimurium* LT2. *Mol. Gen. Genet.* **209**:607-611.
- Hawley, D. K., and W. R. McClure. 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res.* **11**:2237-2255.
- Hohn, B., and J. Collins. 1980. A small cosmid for efficient cloning of large DNA fragments. *Gene* **11**:291-298.
- Hone, D., R. Morona, S. Attridge, and J. Hackett. 1987. Construction of defined *galE* mutants of *Salmonella* for use as vaccines. *J. Infect. Dis.* **156**:167-174.
- Ikemura, T., and H. Ozeki. 1982. Codon usage and transfer RNA contents: organism-specific codon-choice patterns in reference to the isoacceptor contents. *Cold Spring Harbor Symp. Quant. Biol.* **47**:1087-1097.
- Jones, G. W., D. K. Rabert, D. M. Svinarich, and H. J. Whitfield. 1982. Association of adhesive, invasive, and virulent phenotypes of *Salmonella typhimurium* with autonomous 60-megadalton plasmids. *Infect. Immun.* **38**:476-486.
- Kleckner, N., J. Roth, and D. Botstein. 1977. Genetic engineering *in vivo* using translocatable drug-resistance elements. *J. Mol. Biol.* **116**:125-159.
- Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**:105-132.
- Losick, R., and P. W. Robbins. 1967. Mechanism of e^{15} conversion studied with a bacterial mutant. *J. Mol. Biol.* **30**:445-455.
- Lugtenberg, B., J. Meijers, R. Peters, P. van der Hoek, and L. van Alphen. 1975. Electrophoretic resolution of the major outer membrane proteins of *Escherichia coli* K-12 into four bands. *FEBS Lett.* **58**:254-258.
- Mäkelä, P. H. 1965. Inheritance of the O-antigens of *Salmonella* groups B and D. *J. Gen. Microbiol.* **41**:57-65.
- Mäkelä, P. H. 1966. Genetic determination of the O antigens of *Salmonella* groups B (4, 5, 12) and C₁ (6, 7). *J. Bacteriol.* **91**:1115-1125.
- Mäkelä, P. H., and B. A. D. Stocker. 1984. Genetics of lipopolysaccharide, p. 59-137. In E. T. Rietschel (ed.), *Handbook of endotoxin*, vol. 1. Elsevier Science Publishing, Amsterdam.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Messing, J., and J. Vieira. 1982. A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. *Gene* **19**:269-276.
- Morona, R., and P. Reeves. 1981. Molecular cloning of the *tolC* locus of *Escherichia coli* K-12 with the use of transposon Tn10. *Mol. Gen. Genet.* **184**:430-437.
- Naide, Y., H. Nikaido, P. H. Mäkelä, R. G. Wilkinson, and B. A. D. Stocker. 1965. Semirough strains of *Salmonella*. *Proc. Natl. Acad. Sci. USA* **53**:147-153.
- Nurminen, M., C. E. Hellerqvist, V. V. Valtonen, and P. H. Mäkelä. 1971. The smooth lipopolysaccharide character of 1,4,(5),12 and 1,9,12 transductants formed as hybrids between groups B and D of *Salmonella*. *Eur. J. Biochem.* **22**:500-505.
- Nyman, K., M. Plosila, L. Howden, and P. H. Mäkelä. 1979. Genetic determination of lipopolysaccharide: locus of O-specific unit polymerase in group E of *Salmonella*. *Zentralbl. Bakteriol. Hyg. Abt. 1 Orig. Reihe A* **243**:355-362.
- Palva, E. T., and P. H. Mäkelä. 1980. Lipopolysaccharide heterogeneity in *Salmonella typhimurium* analyzed by sodium dodecyl sulfate/polyacrylamide gel electrophoresis. *Eur. J. Bio-*

- chem. **107**:137–143.
34. **Popot, J.-L., and D. M. Engelman.** 1990. Membrane protein folding and oligomerization: the two stage model. *Biochemistry* **29**:4031–4034.
 35. **Roberts, D. E., B. C. Hoopes, W. R. McClure, and N. Kleckner.** 1985. *IS10* transposition is regulated by DNA adenine methylation. *Cell* **43**:117–130.
 36. **Rosenberg, M., and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. *Annu. Rev. Genet.* **13**:319–353.
 37. **Russel, M., and P. Model.** 1984. Replacement of the *fip* gene of *Escherichia coli* by an inactive gene cloned on a plasmid. *J. Bacteriol.* **159**:1034–1039.
 38. **Sancar, A., A. M. Hack, and W. D. Rupp.** 1979. Simple method for identification of plasmid-coded proteins. *J. Bacteriol.* **137**:692–696.
 39. **Sanger, F., A. R. Coulson, B. G. Barrell, A. J. H. Smith, and B. A. Roe.** 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J. Mol. Biol.* **143**:161–178.
 40. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
 41. **Sansonetti, P. J., D. J. Kopecko, and S. B. Formal.** 1981. *Shigella sonnei* plasmids: evidence that a large plasmid is required for virulence. *Infect. Immun.* **34**:75–79.
 42. **Shine, J., and L. Dalgarno.** 1974. The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to non-sense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci. USA* **71**:1342–1346.
 43. **Simon, R., U. Priefer, and A. Pühler.** 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in gram-negative bacteria. *Biotechnology* **1**:784–791.
 44. **Stocker, B. A. D., and P. H. Mäkelä.** 1971. Genetic aspects of biosynthesis and structure of *Salmonella* lipopolysaccharide, p. 369–438. In G. Weinbaum, S. Kadis, and A. Ajl (ed.), *Microbial toxins*, vol. IV. Academic Press, Inc., New York.
 45. **Stoker, N. G., N. F. Fairweather, and B. G. Spratt.** 1982. Versatile low-copy-number plasmid vectors for cloning in *Escherichia coli*. *Gene* **18**:335–341.
 46. **Subbiah, T. V., and B. A. D. Stocker.** 1964. Rough mutants of *Salmonella typhimurium*. I. Genetics. *Nature (London)* **201**:1298–1299.
 47. **Tabor, S., and C. C. Richardson.** 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA* **82**:1074–1078.
 48. **Tabor, S., and C. C. Richardson.** 1987. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. USA* **84**:4767–4771.
 49. **Tinoco, J., Jr., P. N. Borer, B. Dengler, M. D. Levine, O. C. Uhlenbeck, D. M. Gothers, and J. Gralla.** 1973. Improved estimation of secondary structure in ribonucleic acids. *Nature (London)* **246**:40–41.
 50. **Ubben, D., and R. Schmitt.** 1986. Tn1721 derivatives for transposon mutagenesis, restriction mapping and nucleotide sequence analysis. *Gene* **41**:145–152.
 51. **Valtonen, M. V., M. Plosila, V. V. Valtonen, and P. H. Mäkelä.** 1975. Effect of the quality of lipopolysaccharide on mouse virulence of *Salmonella enteritidis*. *Infect. Immun.* **12**:828–832.
 52. **Verma, N. K., and P. R. Reeves.** 1989. Identification and sequence of *rfbS* and *rfbE*, which determine antigenic specificity of group A and D salmonellae. *J. Bacteriol.* **171**:5694–5701.
 53. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an m13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**:259–268.
 54. **Way, J. C., M. A. Davis, D. Morisato, D. E. Roberts, and N. Kleckner.** 1984. New Tn10 derivatives for transposon mutagenesis and for construction of *lacZ* operon fusions by transposition. *Gene* **32**:369–379.
 55. **Wilkinson, R. G., P. Gemski, Jr., and B. A. D. Stocker.** 1972. Non-smooth mutants of *Salmonella typhimurium*: differentiation by phage sensitivity and genetic mapping. *J. Gen. Microbiol.* **70**:527–554.
 56. **Yuasa, R., K. Nakane, and H. Nikaido.** 1970. Structure of cell wall lipopolysaccharide from *Salmonella typhimurium*: structure of lipopolysaccharide from a semirough mutant. *Eur. J. Biochem.* **15**:63–71.
 57. **Zubay, G.** 1973. The *in vitro* synthesis of proteins in microbial systems. *Annu. Rev. Genet.* **7**:267–287.