JOURNAL OF BACTERIOLOGY, Dec. 1992, p. 7750–7756
0021-9193/92/237750-07$02.00/0

Vol. 174, No. 23

The Gene Coding for 3-Deoxy-*manno*-Octulosonic Acid Transferase and the *rfaQ* Gene Are Transcribed from Divergently Arranged Promoters in *Escherichia coli*

TONY CLEMENTZ†

Department of Microbiology, University of Lund, S-223 62 Lund, Sweden

Received 11 March 1992/Accepted 30 July 1992

The gene *kdtA* in *Escherichia coli* codes for 3-deoxy-D-*manno*-octulosonic acid transferase, the enzyme responsible for attachment of the two 3-deoxy-D-*manno*-octulosonic acid residues that constitute the link between lipid A and the core oligosaccharide of the lipopolysaccharide. Cloning and subsequent sequencing of the region upstream of *kdtA* revealed an open reading frame identified as the first gene (*rfaQ*) in an *rfa* gene cluster. The *kdtA* and *rfaQ* transcripts were identified, and the 5′ ends of the transcripts were mapped by primer extension. Two main, divergently arranged promoters were found. These promoters generated transcripts with 5′ ends separated by 289 bases. That the two divergent transcripts from the identified promoters represent the *kdtA* and *rfaQ* transcripts was confirmed by fusing different parts of the intergenic region between the promoterless *lacZ* and *phoA* genes in promoter-screening plasmid pCB267.

The anionic sugar 3-deoxy-D-*manno*-octulosonic acid (KDO) is an essential component of the lipopolysaccharide (LPS) in *Escherichia coli* and probably in most other gram-negative bacteria as well (21). As the proximal part of the inner core oligosaccharide, it links the polysaccharide portion of LPS to the membrane-anchored lipid A. Two molecules of KDO are sequentially transferred to lipid A by a KDO transferase with the activated sugar CMP-KDO as a substrate (6). The gene coding for the $(KDO)_2$ transferase (*kdtA*) has been cloned, sequenced, and shown to code for a 43-kDa polypeptide (9). The incorporation in extracts of *E. coli* of the two stereochemically distinct KDO residues is stimulated about 100-fold when *kdtA* is expressed from a high-copy-number plasmid. It is thus strongly believed that *kdtA* encodes both the first and the second KDO transferase activities. The gene *kdtA* is located adjacent to the *rfa* cluster mapping at about 81 min on the *E. coli* chromosome. The *rfa* gene cluster codes for several of the enzymes involved in the assembly of the LPS core region (23). Only a few of the enzymes encoded by *rfa* have been characterized at the biochemical level. Schnaitman et al. (1, 25) have provided evidence for an *rfa* operon (*rfaQGP-BIJYZK*) encoding 10 different polypeptides. This operon is suggested to have a complex organization with overlapping transcriptional units, each with its own promoter, and a major promoter at the 5′ end of the operon from which a transcript representing all 10 genes is initiated. Glycosyltransferases I and II, catalyzing the stepwise transfer of hexose sugars to the distal part of the core oligosaccharide, have been shown to be encoded by the *rfaG* and *rfaI* (formerly *rfaM* in *E. coli*) genes, respectively (1, 11). Several other genes of this operon have been identified as genes involved in the synthesis of the outer core region. The function of *rfaQ*, however, is still unknown.

Some type of correlation between cell growth and LPS biosynthesis seems likely because of the predominance of LPS in the outer monolayer of the outer membrane of most gram-negative bacteria. This possibility is also suggested by the observation that two central genes in lipid A biosynthesis, *lpxA* and *lpxB*, are transcriptionally coupled to *dnaE* (the catalytic subunit of DNA polymerase III) (10, 21). A variation in growth conditions has been shown to affect the core structure of LPS in *E. coli* (16). However, except for the observation that *rfaH* (formerly *sfrB* in *E. coli*) acts as a positive regulator of some of the *rfa* genes (20), not much is known about how LPS core biosynthesis is controlled. That the *rfa* gene cluster and the *kdtA* gene are located back-to-back makes the intergenic region between *kdtA* and *rfaQ* a possible location for a control region.

This paper describes the first characterization of this intergenic region by mapping the start of the *kdtA* and *rfaQ* transcripts and by sequencing *rfaQ*.

MATERIALS AND METHODS

Materials. [γ-^{32}P]ATP and [^{35}S]dATPαS were obtained from Amersham Corp. Redistilled phenol was purchased from IBI Technologies. Modified T7 DNA polymerase (Sequenase II) was a product of United States Biochemical Corp. Avian myeloblastosis virus (AMV) reverse transcriptase was obtained from Pharmacia. Amersham Hyperfilm-ECL and Hyperfilm β-max were used for autoradiography. Restriction enzymes and other enzymes used in cloning were obtained from Boehringer Mannheim. Nitrocefin was a product of Oxoid Ltd. *p*-Nitrophenyl phosphate and *o*-nitrophenyl-β-D-galactosidase were purchased from Sigma.

Bacterial strains and growth conditions. Cultures were grown in Luria broth, consisting of 10 g of NaCl, 5 g of yeast extract, and 10 g of tryptone per liter, or in minimal medium A (2). Required amino acids were added to minimal medium A at 40 mg/liter, and carbon sources were used at 0.2% (wt/vol). Ampicillin was added to a concentration of 100 μg/ml. The strains used in this study are derivatives of *E. coli* K-12, and their genotypes are shown in Table 1. *E. coli* CB806 and plasmid pCB267 were generous gifts from C. F. Beck.

Recombinant DNA techniques. Plasmids were prepared either by the modified boiling procedure of Holmes and Quigley (13) or by the Qiagen procedure (DIAGEN, Gmbh). DNA fragments to be used for cloning or as radiolabelled

† Electronic mail address: Tony.clementz.mikrbiol.lu.se.

TABLE 1. *E. coli* strains and plasmids

Strain or plasmid	Description	Source or reference
Strains		
CB806	Δ*lacZ lacY*$^+$ *galK phoA8 rpsL thi recA56*	C. F. Beck[a]
XL1-Blue	*recA lac endA1 gyrA96 thi hsdR17 supE44 relA1* (F′ *proAB lacI*^{q} *lacZ-M15*)	Stratagene
Plasmids		
pLC17-24	Complements *kdtA1* and *rfaG*	B. Bachmann[b]
pCL5	Contains a 1.3-kb *Hin*dIII-*Hin*dIII fragment from pLC17-24; see Fig. 1	This work
pCL3	Contains a 2.8-kb *Hin*dIII-*Bam*HI fragment from pLC17-24; see Fig. 1; complements *kdtA1*	9
pCB267	Promoterless *lacZ* and *phoA*; see Fig. 3A	C. F. Beck

[a] Albert-Ludwigs-Universität, Freiburg, Germany.
[b] *E. coli* Genetic Stock Center, Yale University, New Haven, Conn.

probes were isolated by agarose gel electrophoresis and purification with GeneClean (Bio 101). Sequencing with modified T7 DNA polymerase (Sequenase II) was done on nested deletions of pCL5 generated with exonuclease III and mung bean nuclease as described earlier (9). Both DNA strands were sequenced. Sequencing across the *rfaQ* internal *Hin*dIII site was accomplished with a subclone containing a 510-bp *Cla*I fragment from pLC17-24 (see Fig. 1). All processing of sequence data was done with the GCG sequence analysis software package (12).

Primer extension. RNA to be used for primer extension was isolated by the hot phenol method (22). Cultures were grown in minimal medium A supplemented with 0.2% glucose and 0.1% Casamino Acids and harvested in the mid-log phase. Primer extension was done as described by Ayer and Dynan (3) with 10 to 20 μg of total RNA, 0.2 pmol of ^{32}P-end-labelled primer, and 1 U of AMV reverse transcriptase. The primers used were as follows: Q_1, GCGTACT GATGACAGGAGTAGTTAA; Q_2, CCAGGCTCAACAGC TACCTGAATAC; K_1, GCCGTATCCAGATCAGCGGCTG AAT; and K_2, GCTTCAGTTTGGTCTGATCTGCCGC. Primers were 5′ end labelled with [γ-^{32}P]ATP (>5,000 Ci mmol^{-1}) and T4 polynucleotide kinase (2). The generated primer-extended products were separated on 6% polyacrylamide gels (6% [wt/vol] acrylamide, 0.3% [wt/vol] bisacrylamide) containing 7 M urea. Sequencing reactions done with the same primers as those used for primer extension were used as standards.

Assays for enzyme activities. When not otherwise stated, 1 ml of cells grown in Luria broth or minimal medium A supplemented with 0.2% glucose and 0.1% Casamino Acids was harvested in the mid-log phase, washed once, and then resuspended in ice-cold 0.1 M NaCl–10 mM Tris-HCl (pH 8.0). The cells were permeabilized by adding 2 drops of chloroform and 1 drop of 1% sodium dodecyl sulfate and vortexing the mixture vigorously for 15 s. The samples were then immediately frozen at −70°C.

Alkaline phosphatase and β-galactosidase activities were determined spectrophotometrically (24) with *p*-nitrophenyl phosphate and *o*-nitrophenyl-β-D-galactosidase, respectively, as substrates. β-Lactamase activity was measured spectrophotometrically with the chromogenic substrate nitrocefin as the substrate (17). Typically, 10 μl of permeabilized cell suspension was mixed with 10 mg of nitrocefin per ml in 50 mM potassium phosphate buffer (pH 7.0). Enzyme activity was measured during the linear increase in the A_{410} and expressed as the A_{410} min^{-1} mg of protein^{-1}.

Nucleotide sequence accession number. The GenBank accession number for *kdtA* and *rfaQ* is M86305.

RESULTS

The start of an open reading frame (Fig. 1) has been identified upstream of *kdtA* on a 2.8-kb *Hin*dIII-*Bam*HI fragment (9). This open reading frame (*rfaQ*) is oriented with a polarity opposite that of *kdtA*. A 1.3-kb *Hin*dIII-*Hin*dIII fragment adjacent to the 2.8-kb *Hin*dIII-*Bam*HI fragment (Fig. 1) was cloned from pLC17-24 (8, 9) into pBluescript KSII(+), generating plasmid pCL5. The restriction map of the cloned fragment was consistent with that of the same region reported by Austin et al. (1). Both strands of the cloned fragment were sequenced after nested deletions were generated in pCL5 as described in Material and Methods. The nucleotide sequence of the *Hin*dIII-*Hin*dIII fragment completes the sequence of *rfaQ*, starting 350 bp into the *Hin*dIII-*Bam*HI fragment of pCL3 (Fig. 1). This open reading frame could code for a polypeptide of 344 amino acid residues and with a molecular mass of 38.000 Da. The sequence of *rfaQGP* was recently published by Parker et al. (18); only the N-terminal part of *rfaQ* and the intergenic region between *rfaQ* and *kdtA* is shown in Fig. 2. Good agreement was found between the sequence published by Parker et al. and the sequence presented in this study. They do, however, differ at two positions. The sequence of Parker et al. has an insertion of an extra G (position 1547 in Fig. 2) and a deletion of a T (position 1729 in Fig. 2) compared with our sequence. The inserted base changes the reading frame of the N-terminal part of *rfaQ* and shifts the start codon, resulting in a slightly shorter polypeptide. A weak, but significant, sequence similarity can be found between the sequence of RfaQ and the carboxy-terminal sequence of a heptosyltransferase in *Salmonella typhimurium*, RfaC, derived from the partial DNA sequence of *rfaC* (15). Also, the sequence of the KDO transferase, KdtA (9), shows some similarities to these two enzymes (data not shown). No other significant sequence homologies could be found between *rfaQ* and any other sequence in the EMBL data base (release 27.0).

Vector pCB267 (4), containing the promoterless *lacZ* and *phoA* genes in opposite orientations and separated by a cloning region (Fig. 3A), was used to analyze the intergenic region between *rfaQ* and *kdtA* for promoter activities. The *Cla*I-*Hin*dII fragment from pCL3 (Fig. 3B) was subcloned into pCB267 after new ends compatible with the *Hin*dII-*Sal*I sites in pCB267 were generated as described below for subcloning of the *Taq*I fragments (Fig. 3B). The plasmid construct was used to transform *E. coli* CB806 (*lac pho*). The transformants obtained yielded blue colonies when grown on Luria broth agar containing either 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) or 5-bromo-4-chloro-2-indolyl-phosphate (X-Pho), indicating bidirectional transcription originating in the cloned *Cla*I-*Hin*dII fragment.

Two *Taq*I fragments, contained within the above-described *Cla*I-*Hin*dII fragment and representing *kdtA* and *rfaQ*-proximal regions, respectively, were isolated from pCL3 (Fig. 3B) to locate more precisely the respective promoters. To clone the two *Taq*I fragments in pCB267 in specific orientations, we inserted the fragments into the *Cla*I

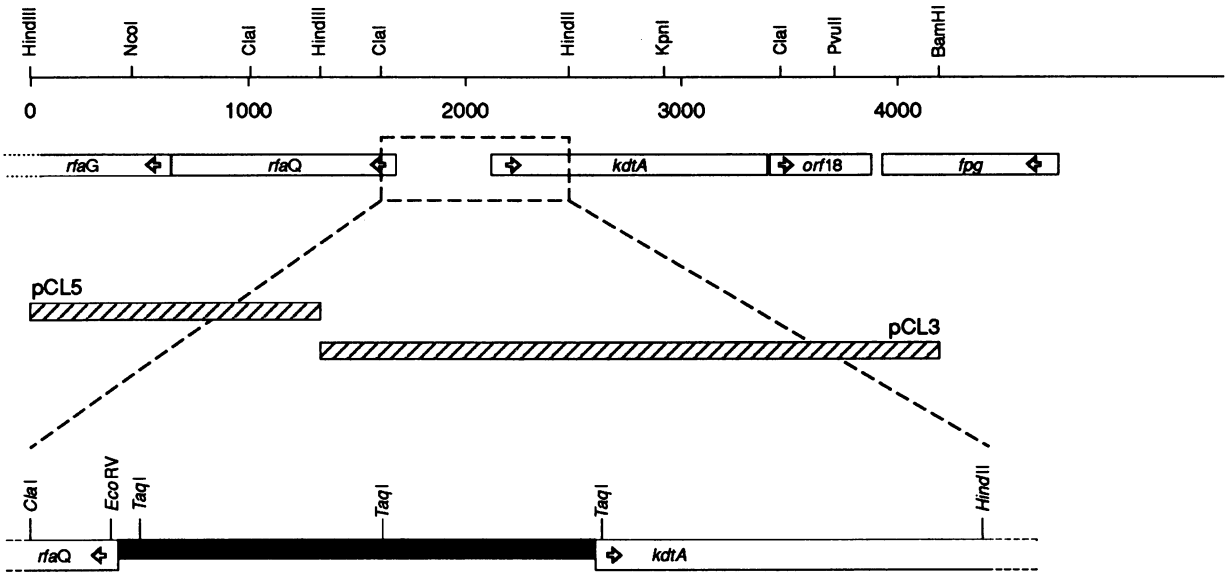

FIG. 1. Organization of genes surrounding *kdtA* in *E. coli*. The scale indicates distance in bases from the *rfaG* internal *Hin*dIII site. Open boxes represents open reading frames, and the orientation of each open reading frame is indicated by an arrow. *orf-18* encodes an 18-kDa polypeptide of unknown function (9). The two open reading frames upstream of *kdtA* were identified by comparing the restriction map of sequenced DNA with the restriction map of Austin et al. (1). The DNA fragments cloned into pCL3 and pCL5 are shown as hatched boxes. The *Cla*I-*Hin*dII region containing the intergenic region between *rfaQ* and *kdtA* is enlarged at the bottom of the figure.

site in the cloning cassette of pBluescript KSII(+). These two constructs were used to transform *E. coli* XL1-Blue. Plasmid DNA from several of the transformants was sequenced with a primer complementary to pBluescript KSII(+), enabling sequencing across one end of the insert. Clones containing plasmid DNA with the *rfaQ*-proximal

```
  1 ATAGAGCGCATTAATTTCCGGGTTTTCAGACAAAATAGGGATGGTGTCCTGATAAAGCAG  60
    TyrLeuAlaAsnIleGluProAsnGluSerLeuIleProIleThrAspGlnTyrLeuLeu

 61 CATATCGATTTTTGCATCAGGATAATTCTGCTTGAGCGTACTGATGACAGGAGTAGTTAA 120
    MetAspIleLysAlaAspProTyrAsnGlnLysLeuThrSerIleValProThrThrLeu

                 rfaQ
121 TAACATATCCCCATGATATCGCATCTTTATGACCAGGATTTTTCGAAATGGCTTTTCCAC 180
    LeuMetAspGlyHisTyrArgMet
                          rfaQp2 ↓↓
181 TAGCGACTCTTTTGTGTGATTGTCTGGTTAAGTTAAGCAGAAAAAAGCACGCTACCGCCC 240
                       rfaQp1 ↓
241 CAGGCTCAACAGCTACCTGAATACTGATAACGCACTTACTTTGACGTGCTAGTGTATCAT 300
  ↓ kdtAp2
301 TTCTTTCGCAGCAATCCGACCCGAATGATTTTTATACACAAAATATACTTTAATCATAAA 360

361 AATCAAATAGATAAAAAATAGCATTTCGATAATGTCAGTACGGCCTATGTAGATTTGAAA 420

421 ATGGGGCCGAGTTTCAATGAATCGCATAAAAATACAAAAAAAATGAAGATGTAGCGGCAG 480

481 ATCAGACCAAACTGAAGCTATTTAAGTCAAAAACGGGAAAAGTAATGGTAAAGCCACAGC 540
                ↓ kdtAp1          kdtA
541 TAAATACATAGAATCCCCAGCACATCCATAAGTCAGCTATTTACTATGCTCGAATTGCTT 600
                                    MetLeuGluLeuLeu
601 TACACCGCCCTTCTCTACCTTATTCAGCCGCTGATCTGGATACGGCTCTGGGTGCGCGGA 660
    TyrThrAlaLeuLeuTyrLeuIleGlnProLeuIleTrpIleArgLeuTrpValArgGly
```

FIG. 2. DNA sequence of the intergenic region between *kdtA* and *rfaQ*. The nucleotides are numbered starting with the *rfaG* internal *Hin*dIII site. The predicted amino acid sequences of the N-terminal portions of KdtA and RfaQ are shown below the nucleotide sequences. The nucleotide corresponding to the 5′ end of each of the identified transcripts is indicated by an arrow. The nucleotides matching the consensus sequence of the −10 promoter regions $rfaQp_1$ and $kdtAp_1$ are underlined.

*Taq*I fragment in the two possible orientations were isolated. Eighteen clones containing the *kdtA*-proximal *Taq*I fragment all had the same orientation. The orientation not found would have had the β-galactosidase promoter from pBluescript KSII(+) in the same orientation as *kdtA*. A similar problem was encountered when nested deletions of pCL4 (9) were generated for the sequencing of *kdtA* (9). Deletions starting within this *Taq*I fragment could not be isolated when the β-galactosidase promoter in the vector was oriented in the same direction as *kdtA*.

The *Taq*I inserts were cut out from plasmid DNA by use of the flanking *Hin*dIII-*Sal*I sites in the remaining part of the cloning cassette and ligated into the corresponding sites in pCB267. The resulting plasmids (Fig. 3B) were used to transform *E. coli* CB806 (*lac pho*), and the plasmids were named as shown in Fig. 3B. Clones pCLP310/CB806, pCLP312/CB806, and pCLP313/CB806 were grown in Luria broth, and the activities of β-galactosidase (LacZ) and alkaline phosphatase (PhoA) were determined in early logarithmic cells. The activities of LacZ and PhoA were correlated with the activity of plasmid-encoded β-lactamase (Bla) to adjust for any change in plasmid copy number (14). The results show that the *Cla*I-*Hin*dII fragment contains divergent promoter activities and that these activities can be separated on each of the two *Taq*I fragments (Table 2). Similar results were obtained when cells were grown in minimal medium A. None of the *Taq*I fragments promoted as high an activity of the reporter genes as the intact *Cla*I-*Hin*dII fragment. This result may reflect the different lengths and compositions of the 5′-untranslated regions of the mRNAs generated from the different hybrid plasmids.

The location of the two promoters was more precisely established by mapping the 5′ ends of the two divergent transcripts by primer extension. Two primers (Q_1 and K_1) were chosen to be complementary to a 25-nucleotide region in the open reading frames of *rfaQ* and *kdtA*, respectively. A

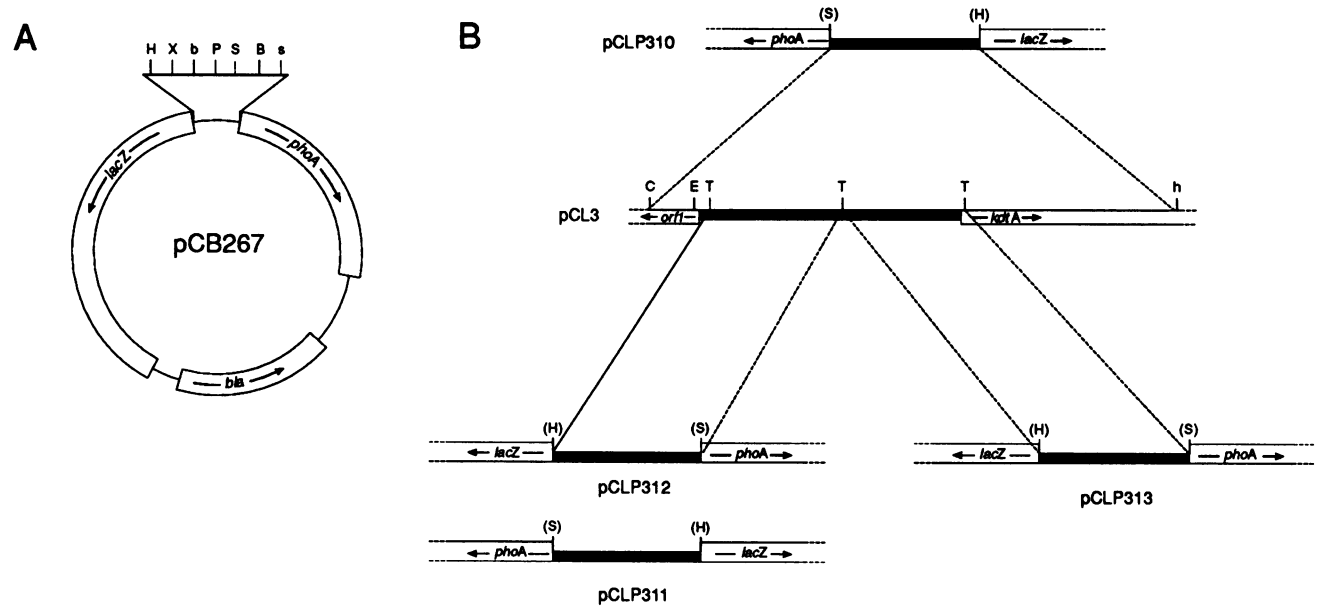

FIG. 3. Construction of subclones for promoter screening. (A) Promoter-screening vector constructed by Beck and Schneider (4) and containing the *lacZ* and *phoA* structural genes in opposite orientations. Both genes lack the promoter regions and are instead separated by a cloning cassette. The restriction sites in the cloning cassette are shown. The directions of transcription of the structural genes are indicated by arrows. (B) Subcloning of various parts of the intergenic region (solid box) of pCL3 into pCB267. Both the *Cla*I-*Hin*dII fragment and the *Taq*I fragments from pCL3 were subcloned into pCB267 after the generation of ends compatible with sites in the cloning cassette of pCB267. The ends were made compatible with *Hin*dIII and *Sal*I by subcloning of the fragments into the *Cla*I-*Hin*dII and *Cla*I sites, respectively, in pBluescript KSII(+). After transformation and selection, the fragments were cut out from the reisolated plasmids by use of the flanking *Bam*HI-*Sal*I sites present in the cloning cassette of pBluescript KSII(+). Abbreviations: (H) and (S), *Hin*dIII and *Sal*I sites originating from pBluescript KSII(+), respectively; H, *Hin*dIII; X, *Xba*I; b, *Bgl*II; S, *Sal*I; B, *Bam*HI; s, *Sma*I; C, *Cla*I; T, *Taq*I; h, *Hin*dII; E, *Eco*RV.

primer extension reaction with primer Q_1 and total RNA isolated from *E. coli* XL1/pCL3 generated two major primer-extended products (Fig. 4A). The shorter product corresponds to an *rfaQ* transcript with the 5′ end 80 to 81 nucleotides upstream of the start of *rfaQ* (bp 225–226 in Fig. 2), and the longer product corresponds to a transcript starting 140 nucleotides upstream of *rfaQ* (bp 284 in Fig. 2). Exchanging primer K_1 for Q_1 in the reaction generated one dominant primer-extended product (Fig. 4B), corresponding to a *kdtA* transcript with the 5′ end 24 nucleotides upstream of the start of *kdtA* (bp 562 in Fig. 2). The two *rfaQ* products and the *kdtA* product could also be detected with total RNA isolated from *E. coli* XL1/pBluescript KSII(+). These products thus reflect chromosomally encoded transcripts. The chromosomally encoded transcripts were present at roughly 1% of the corresponding plasmid-encoded transcripts, as estimated from the time of exposure needed for comparable band intensity.

Two additional primers (Q_2 and K_2) were used to verify the results obtained with primers Q_1 and K_1. Primer Q_2 was complementary to a region 97 to 121 bp upstream of *rfaQ*, thus excluding the detection of any transcript beginning 80 to 81 nucleotides upstream of *rfaQ* (the short product detected with Q_1). Primer K_2 was complementary to a region 88 to 113 bp upstream of *kdtA*. With primer Q_2 and total RNA from *E. coli* XL1/pCL3 in the primer extension reaction, a transcript that has a 5′ end identical to that of the longer transcript generated with primer Q_1 was detected (Fig. 5A), thus verifying that this is the 5′ end of a transcript and not a gel migration artifact due to strong secondary structures in the product. A primer-extended product obtained with primer K_2 and total RNA from *E. coli* XL1/pCL3 and corresponding to the 5′ end of a transcript starting 284 nucleotides upstream of the start of *kdtA* was found. This product was present at an amount severalfold lower than that of the major product detected with primer K_1 and total RNA from *E. coli* XL1/pCL3. The product found with primer K_2 could not be detected in primer extension reactions with primer K_1. Either AMV reverse transcriptase is unable to extend primer K_1 the 340 nucleotides necessary to reach the 5′ end of this longer transcript or the product found with primer K_1 represents a transcript separate from the one detected with primer K_2.

TABLE 2. Expression of β-galactosidase and alkaline phosphatase in *E. coli* CB806 containing promoter-screening plasmids

Plasmid[a]	% Activity[b] of:	
	β-Galactosidase	Alkaline phosphatase
pCLP310	100	100
pCLP312	25	11
pCLP313	4.0	69
pCB267	<0.8	11

[a] Activities were measured in cells grown in Luria broth with 100 μg of ampicillin per liter as described in Materials and Methods.

[b] Activity is correlated with β-lactamase activity and is presented relative to that of *E. coli* CB806/pCLP310.

DISCUSSION

An *rfa* operon containing genes involved in the assembly of the core oligosaccharide of LPS has been mapped and

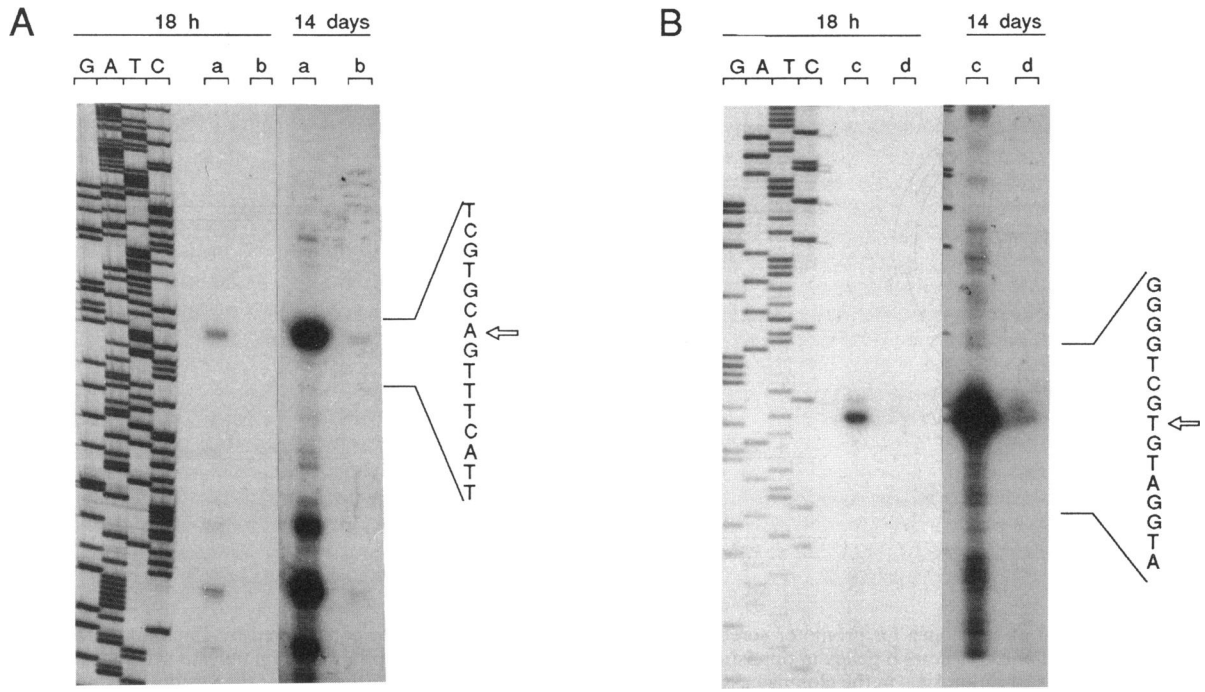

FIG. 4. Autoradiogram of ^{32}P-labelled primer extension products obtained with primers Q_1 and K_1. Primer extension reactions were done as described in Materials and Methods. Total RNA was isolated from *E. coli* XL1/pBluescript KSII(+) (lanes b and d) or from *E. coli* XL1/pCL3 (lanes a and c). The sequencing ladders obtained with the primer used in each primer extension experiment are shown in lanes A, G, T, and C. The arrows indicate the nucleotide corresponding to the 5′ end of the dominating primer extension product. The time of exposure is indicated above the lanes. (A) A 5′-end-labelled 25-mer (primer Q_1) was hybridized to 10 μg (lane a) or 20 μg (lane b) of total RNA. (B) A 5′-end-labelled 25-mer (primer K_1) was hybridized to 10 μg (lane c) or 20 μg (lane d) of total RNA.

characterized by restriction enzyme analysis and by insertion mutagenesis (1, 25). Restriction enzyme analysis of the 1.3-kb *Hin*dIII-*Hin*dIII region upstream of *kdtA* identifies this region as part of the beginning of the *rfaQGP* gene cluster. The first gene in this cluster, *rfaQ* (25), codes for a polypeptide with an apparent molecular weight of 42,000 (1), in fair agreement with the calculated molecular mass of 39 kDa obtained from the nucleotide sequence. Mutations in *rfaQ* cause a deep rough phenotype of LPS (1), but the function of RfaQ is still unknown. The observed similarity among the amino acid sequences of RfaQ, heptosyltransferase in *S. typhimurium* (RfaC), and KDO transferase (KdtA) supports the notion that RfaQ is involved in the biosynthesis of the inner core region. However, a definite functional assignment of the gene must await the isolation and identification of the gene product.

Recently, a locus (*lsi*) involved in the biosynthesis of lipooligosaccharide in *Neisseria gonorrhoeae* was sequenced (19). The amino acid sequence of an open reading frame coding for a polypeptide of 39 kDa shows similarities to the amino acid sequences of RfaQ, RfaC, and KdtA. Mutations in *lsi* produce a defective lipooligosaccharide, but the enzymatic function of the polypeptide has not been identified.

That *rfaQGP* is transcribed in a direction opposite that of *kdtA* (9, 25) suggests the presence of at least two promoters in the intergenic region between *rfaQ* and *kdtA*. This suggestion is verified by the restoration of both β-galactosidase and alkaline phosphatase activities in *E. coli* CB806 by pCLP310, a plasmid containing the entire intergenic region inserted between the promoterless reporter genes *lacZ* and *phoA* in pCB267.

5′ end mapping of transcripts generated from the intergenic region revealed two major, divergently arranged transcripts. The primer-extended product from primer Q_1, indicating an *rfaQ* transcript 60 nucleotides shorter than the longer transcript, is most likely an artifact. The 5′ end of this shorter product coincides with a DNA region from which it was very difficult to obtain a nucleotide sequence because of premature termination of the sequencing reaction. It is thus not surprising that AMV reverse transcriptase, used in the primer extension reaction, also would terminate prematurely at this region. The possibility that it is the longer transcript that is an artifact was ruled out by use of a primer excluding the shorter transcript. However, the primer extension experiments cannot definitely exclude the possibility that two alternative start sites exist for the *rfaQ* transcript. A second transcript was also identified in the direction of *kdtA*. The 5′ end of this transcript is separated by only 22 nucleotides from the 5′ end of the longer *rfaQ* transcript, meaning that the promoter regions for the longer *rfaQ* transcript ($rfaQp_1$) and for the longer *kdtA* transcript ($kdtAp_2$) overlap. Expression from $kdtAp_1$ does not require the $kdtAp_2$ region located in *cis*, since the *kdtA*-proximal *Taq*I fragment, containing only $kdtAp_1$, is sufficient for promoting the expression of alkaline phosphatase in pCB267. This fact is consistent with the earlier observation that KDO transferase activity is overexpressed from a plasmid containing *kdtA* and 147 bp of the upstream region but not from a plasmid containing *kdtA* and only 35 bp of the upstream region (9).

Many examples of functionally related genes with divergently arranged promoters show coordinated regulation (5). This type of gene organization seems to be a widespread method for two sets of genes to be transcribed and regulated

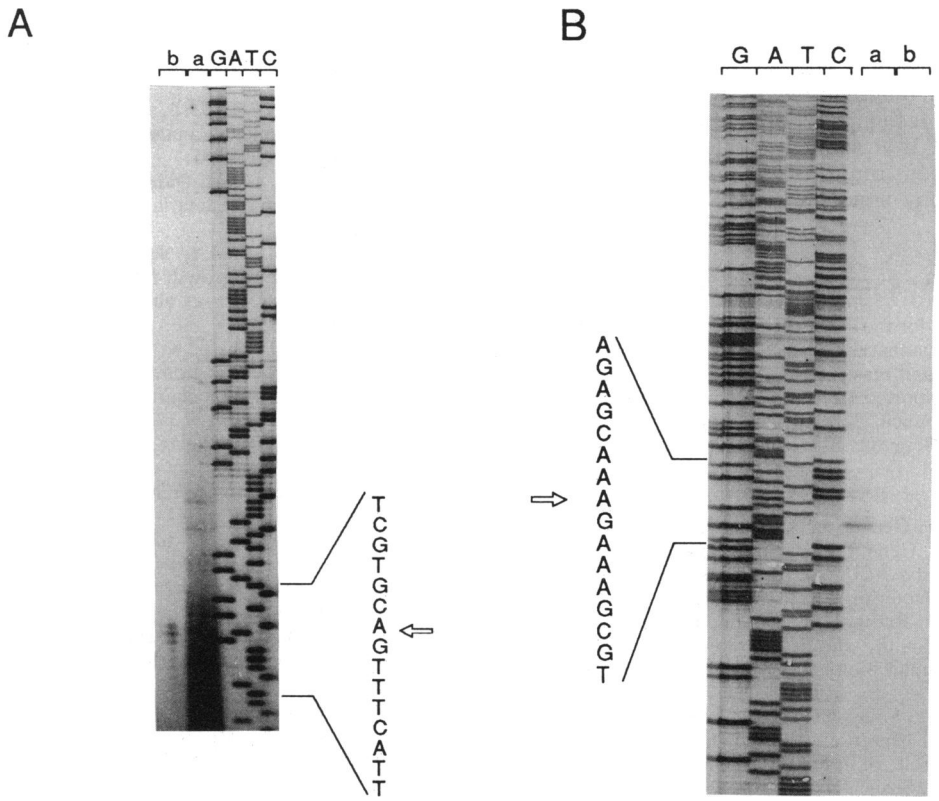

FIG. 5. Autoradiogram of ^{32}P-labelled primer extension products obtained with primers Q_2 and K_2. Primer extension reactions were done as described in Materials and Methods. Total RNA was isolated from *E. coli* XL1/pCL3. The sequencing ladders obtained with the primer used in each primer extension experiment are shown in lanes A, G, T, and C. The arrows indicate the nucleotide corresponding to the 5′ end of the dominating primer extension product. (A) A 5′-end-labelled 25-mer (primer K_2, lane a, or primer Q_2, lane b) was hybridized to 10 μg of total RNA. (B) A 5′-end-labelled 25-mer (primer K_2, lane a, or primer Q_2, lane b) was hybridized to 10 μg of total RNA.

from a single control region. The function of divergent promoters is in most cases poorly understood, but several types of mechanisms can be involved. The proximity of the two promoters may affect the frequency of initiation of transcription at one or both promoters. If regulatory proteins are involved, they may bind to a centrally located control region and regulate transcription in both directions. Since *rfa* gene products enter LPS biosynthesis after the synthesis of $(KDO)_2$-lipid IV_A (6, 7), a tightly coordinated expression of *rfa* and *kdtA* seems plausible. A sequence similar to the Pribnow box is found in the −10 regions of $kdtAp_1$, $kdtAp_2$, and $rfaQp_1$. However, no similarity to the so-called "consensus" sequence of *E. coli* σ^{70} promoters is found in the −35 regions (Fig. 6). The start sites of the major *rfaQ* and *kdtA* transcripts are separated by 289 bp, a fact that argues against any interference in the binding of RNA polymerase at the two promoter regions. However, it does not exclude the possibility that indirect interference, e.g., the binding of RNA polymerase to one promoter, may affect the topology of the neighboring DNA and the frequency of initiation at the adjacent promoter. It is also possible that the coordinated expression of $rfaQp_1$ and $kdtAp_1$ is mediated by a control element in the DNA region between these two promoters.

The intergenic region between *rfaQ* and *kdtA* could be a possible location for the regulation of the suggested *rfaQGP* operon and *kdtA*, e.g., coordinated control of the production of enzymes involved in the different steps of LPS core biosynthesis. The observed gene arrangement cannot be an

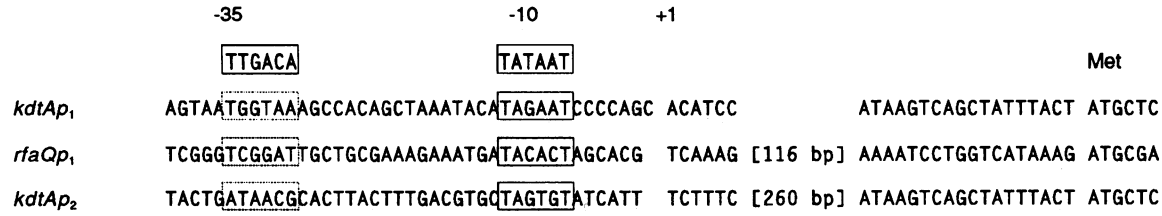

FIG. 6. Promoter regions identified in the intergenic region between *kdtA* and *rfaQ*. Sequences are aligned with the start of transcription at +1. The consensus sequences for the −10 and −35 regions of σ^{70} promoters are shown above the aligned sequences. The −10 and −35 regions in the aligned sequences are indicated within boxes, assuming 17-base spacing between the −10 and −35 regions. The numbers of bases not shown between the start of transcription and the start codon (ATG) are given within brackets.

unconditional requirement for promoter activity, since the expression of unrelated reporter genes from each of the individual promoters can be detected even when the promoters are separated on different fragments. However, it should be noted that these results are observed with promoters on a multicopy plasmid. If an unidentified factor is required for controlled expression from this region, it could be titrated out in this multicopy situation, with a subsequent loss of regulation.

ACKNOWLEDGMENTS

I thank Pia Persson for the sequencing of pCL5 and Lars Rutberg for comments on the manuscript. I also thank C. F. Beck for his kind gift of *E. coli* CB806 and plasmid pCB267.

This work was supported by grants from the Swedish Natural Science Research Council (K-KU 9228), Stiftelsen Lars Hiertas Minne, and Magnus Bergvalls Stiftelse.

REFERENCES

1. **Austin, E. A., J. F. Graves, L. A. Hite, C. T. Parker, and C. A. Schnaitman.** 1990. Genetic analysis of lipopolysaccharide core biosynthesis by *Escherichia coli* K-12: insertion mutagenesis of the *rfa* locus. J. Bacteriol. **172:**5312–5325.
2. **Ausubel, F. M., R. Brent, R. E. Kingsdon, D. D. Moore, J. G. Saidman, J. A. Smith, and K. Struhl (ed.).** 1989. Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
3. **Ayer, D. E., and W. S. Dynan.** 1988. Simian virus 40 major late promoter: a novel tripartite structure that includes intragenic sequences. Mol. Cell. Biol. **8:**2021–2033.
4. **Beck, C. F., and K. Schneider.** 1987. New expression vectors for identifying and testing signal structures for initiation and termination of transcription. Methods Enzymol. **153:**452–461.
5. **Beck, C. F., and R. A. J. Warren.** 1988. Divergent promoters, a common form of gene regulation. Microbiol. Rev. **52:**318–326.
6. **Brozek, K. A., K. Hosaka, A. D. Robertson, and C. R. H. Raetz.** 1989. Biosynthesis of lipopolysaccharide in *Escherichia coli*. Cytoplasmic enzymes that attach 3-deoxy-D-*manno*-octulosonic acid to lipid A. J. Biol. Chem. **264:**6956–6966.
7. **Brozek, K. A., and C. R. H. Raetz.** 1990. Biosynthesis of lipid A in *Escherichia coli*. Acyl carrier protein-dependent incorporation of laurate and myristate. J. Biol. Chem. **265:**15410–15417.
8. **Clarke, L., and J. Carbon.** 1976. A colony bank containing synthetic ColEI hybrid plasmids representative of the entire *Escherichia coli* genome. Cell **9:**91–99.
9. **Clementz, T., and C. R. H. Raetz.** 1991. A gene coding for 3-deoxy-D-*manno*-octulosonic-acid transferase in *Escherichia coli*. Identification, mapping, cloning, and sequencing. J. Biol. Chem. **266:**9687–9696.
10. **Coleman, J., and C. R. H. Raetz.** 1988. First committed step of lipid A biosynthesis in *Escherichia coli*: sequence of the *lpxA* gene. J. Bacteriol. **170:**1268–1274.
11. **Creeger, E. S., and L. I. Rothfield.** 1979. Cloning of genes for bacterial glycosyltransferases. I. Selection of hybrid plasmids carrying genes for two flucosyltransferases. J. Biol. Chem. **254:**804–810.
12. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
13. **Holmes, D. S., and M. Quigley.** 1981. A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. **114:**193–197.
14. **Klotsky, R.-A., and I. Schwartz.** 1987. Measurement of *cat* expression from growth-rate-regulated promoters employing β-lactamase activity as an indicator of plasmid copy number. Gene **55:**141–146.
15. **MacLachlan, P. R., S. K. Kadam, and K. E. Sanderson.** 1991. Cloning, characterization, and DNA sequence of the *rfaLK* region for lipopolysaccharide synthesis in *Salmonella typhimurium* LT2. J. Bacteriol. **173:**7151–7163.
16. **Nelson, D., A. J. Bathgate, and I. R. Poxton.** 1991. Monoclonal antibodies as probes for detecting lipopolysaccharide expression in *Escherichia coli* from different growth conditions. J. Gen. Microbiol. **137:**2741–2751.
17. **O'Callaghan, C. H., A. Morris, S. M. Kirby, and A. H. Shingler.** 1972. Novel method for detection of β-lactamases using a chromogenic cephalosporin substrate. Antimicrob. Agents Chemother. **1:**283–288.
18. **Parker, C. T., E. Pradel, and A. Schnaitman.** 1992. Identification and sequence of the lipopolysaccharide core biosynthetic genes *rfaQ*, *rfaP*, and *rfaG* of *Escherichia coli* K-12. J. Bacteriol. **174:**930–934.
19. **Petricoin, E. F., III, R. J. Danaher, and D. C. Stein.** 1991. Analysis of the *lsi* region involved in lipooligosaccharide biosynthesis in *Neisseria gonorrhoeae*. J. Bacteriol. **173:**7896–7902.
20. **Pradel, E., and C. A. Schnaitman.** 1991. Effect of *rfaH* (*sfrB*) and temperature on expression of *rfa* genes of *Escherichia coli* K-12. J. Bacteriol. **173:**6428–6431.
21. **Raetz, C. R. H.** 1990. Biochemistry of endotoxins. Annu. Rev. Biochem. **59:**129–170.
22. **Resnekov, O., L. Rutberg, and A. Gabain.** 1990. Changes in the stability of specific mRNA species in response to growth stage in *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA **87:**8355–8359.
23. **Rick, P. D.** 1987. Lipopolysaccharide biosynthesis, p. 648–662. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 1. American Society for Microbiology, Washington, D.C.
24. **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.
25. **Schnaitman, C. A., C. T. Parker, J. D. Klena, E. L. Pradel, N. B. Pearson, K. E. Sanderson, and P. R. Maclachlan.** 1991. Physical maps of the *rfa* loci of *Escherichia coli* K-12 and *Salmonella typhimurium*. J. Bacteriol. **173:**7410–7411.