JOURNAL OF BACTERIOLOGY, Aug. 1994, p. 5022–5032
0021-9193/94/$04.00+0

Vol. 176, No. 16

Structure, Expression, and Regulation of the *kilC* Operon of Promiscuous IncPα Plasmids

MICHELLE H. LARSEN AND DAVID H. FIGURSKI*

Department of Microbiology and Cancer Center, College of Physicians and Surgeons, Columbia University, New York, New York 10032

Received 10 March 1994/Accepted 11 June 1994

The *kil-kor* regulon was first identified on the broad-host-range IncPα plasmid RK2 by the presence of multiple *kil* loci (*kilA*, *kilB*, *kilC*, and recently *kilE*) that are lethal to *Escherichia coli* host cells in the absence of regulation by *kor* functions in various combinations. Whereas the *kilB* operon is required for mating-pair formation during conjugation, the functions encoded by the other *kil* loci are not known. They are not essential for replication or conjugal transfer, but their coregulation with replication and transfer genes indicates that they are likely to be important for RK2. In this report, we describe molecular and genetic studies on *kilC*. We determined the nucleotide sequence of the *kilC* region, which is located between the origin of vegetative replication (*oriV*) and transposon Tn*1* on RK2. Primer extension analysis identified the transcriptional start site and showed that a sequence corresponding to a strong σ^{70} promoter is functional. The abundance of RNA initiated from the *kilC* promoter is reduced in the presence of *korA* and *korC*, as predicted from genetic analysis of *kilC* regulation. The first gene of the *kilC* operon (*klcA*) is sufficient to express the host-lethal phenotype of the *kilC* determinant in the absence of *korA* and *korC*. By comparing RK2 to the related IncPα plasmids pUZ8 and R995, we determined that the Tn*1* transposon in RK2 interrupts a gene (*klcB*) immediately downstream of *klcA*. Thus, the *kilC* determinant is normally part of an autoregulated operon of three genes: *klcA*, *klcB*, and *korC*. *klcA* is predicted to encode a 15,856-Da polypeptide that is related to the ArdB antirestriction protein of the IncN plasmid pKM101, suggesting a role for *klcA* in the broad host ranges of IncPα plasmids. The predicted product of the uninterrupted *klcB* gene is a polypeptide of 51,133 Da that contains a segment with significant similarity to the RK2 regulatory proteins KorA and TrbA. Located 145 bp upstream of the *kilC* promoter is a 10th copy of the 17-bp *oriV* iteron sequence in inverted orientation relative to that of the other nine iterons of *oriV*. Iteron 10 is identical to the "orphan" iteron 1, and both have identical 6-bp flanking sequences that make them likely to be strong binding sites for the TrfA replication initiator protein. The locations and relative orientation of orphan iterons 10 and 1 raise the possibility that these iterons promote the formation of a DNA loop via protein-protein interactions by bound TrfA and lead us to propose that they demarcate the functional origin of replication. This analysis of the *kilC* region and our previous studies on the other *kil* loci of RK2 have revealed that the region between *oriV* and the *korABF* operon in wild-type IncPα plasmids is saturated by the *kilC*, *kilE*, and *kilA* loci arranged in four *kor*-regulated operons encoding a total of 12 genes.

The self-transmissible plasmids of incompatibility group P (IncPα and IncPβ) are known for their conjugal and replicative promiscuity among diverse bacteria (10, 40, 42, 55). Studies with the 60-kb IncPα plasmids RK2 (24) and RP4 (10) have revealed a unique regulatory network known as the *kil-kor* regulon (15, 16, 29). The eight operons of the *kil-kor* regulon are controlled by various combinations of transcriptional repressors encoded by *korA*, *korB*, and *korC* (15, 29, 55). The expression of several of the operons is further influenced by other regulatory functions encoded by *korE*, *korFI*, *korFII*, *trbA*, and *kfrA* (42, 55, 65). The functions of some of the proteins encoded by the *kil-kor* regulon are known (42). The *trfA* operon specifies the essential replication initiator function (TrfA), which binds to multiple copies of a 17-bp sequence (iteron) in the origin of replication (*oriV*), and a single-stranded DNA-binding protein (SSB). The *korA* operon includes the regulatory genes *korA*, *korB*, *korFI* and *korFII*, and another gene (*incC*) that may be involved in plasmid maintenance. The *kfrA* operon includes a single gene that expresses a site-specific DNA-binding protein with regulatory activity.

The discovery of the *kil* loci, so named because their unregulated expression in the absence of *kor* functions is lethal to *Escherichia coli* host cells, was the first indication of a complex regulon on IncPα plasmids (15). Five operons are found in four loci known as *kilA*, *kilB*, *kilC*, and *kilE* (15, 17, 29, 51, 58). The *kilB* operon has been shown to include at least 14 genes of the Tra2 region required for conjugal transfer (32, 58). The functions of *kilA*, *kilC*, and *kilE* have not yet been determined, but they are known not to be essential for plasmid replication or conjugal transfer (15, 29). The *kilA* locus consists of an operon of three genes (*klaA*, *klaB*, and *klaC*) (17), and the *kilE* locus contains two operons that specify a total of six genes (*kleA*, *kleB*, *kleC*, *kleD*, *kleE*, and *kleF*) (29).

Because the *kil* determinants are coregulated with important plasmid functions and because they are unique to the promiscuous plasmids of the IncPα group, we have suggested that they encode accessory functions that enhance the maintenance or host range of RK2 among diverse bacterial species (15–17). Here, we report molecular and genetic studies on the structure, expression, regulation, and possible functions of the RK2

* Corresponding author. Mailing address: Department of Microbiology and Cancer Center, College of Physicians and Surgeons, Columbia University, 701 West 168th St., New York, NY 10032. Phone: (212) 305-3425. Fax: (212) 305-1468. Electronic mail address: figurski@cuccfa.ccc.columbia.edu.

TABLE 1. Plasmids used in this study

Plasmid	Selective marker(s)	Relevant properties	Description	Reference or source
pCH8	Ap^r		Cloning vector with pMB1 replicon and polylinker followed by *rrnB* T1 and T2 transcriptional terminators	58
pCRII	Ap^r Km^r		Cloning vector for PCR-amplified DNA fragments	Invitrogen, San Diego, Calif.
pET-17b	Ap^r		Expression vector based on bacteriophage T7 φ10 promoter	Novagen, Madison, Wis.
pRK2161	Tp^r	*klcA*$^+$ *klcB'*	pSM1 replicon containing *kilC* region of plasmid RK2 (kb 14.3–9.5)	15
pRK2292	Tp^r	Φ(*catp-korA*$^+$)	pSM1 replicon with *korA* expressed constitutively from *cat* promoter	2
pRK2320	Cm^r	*klcA*$^+$ *klcB'*	P15A replicon containing *kilC* region of plasmid RK2 (kb 13.0–9.5) (Fig. 5)	28
pRK2462	Cm^r	Φ(*tetp-korC*$^+$)	P15A replicon with *korC* expressed constitutively from *tet* promoter	28
pRK21509	Ap^r		*kilC*-specific 384-bp *Sph*I fragment of pRK2320 cloned in pUC19 for sequencing	This study
pRK21510	Ap^r		*kilC*-specific 432-bp *Eco*RI-*Hin*dIII fragment of pRK21509 cloned in pUC18 for sequencing	This study
pRK21654	Ap^r		*Taq*I-*Bst*YI 247-bp fragment of pRK2320 cloned in pUC19 for sequencing	This study
pRK21655	Ap^r		*Taq*I-*Bst*YI 247-bp fragment of pRK2320 cloned in pUC18 for sequencing	This study
pRK21659	Ap^r Km^r	*kilCp*$^+$	pCRII with 300-bp *kilC* promoter fragment	This study
pRK21660	Ap^r	*kilCp*$^+$	pCH8 with *kilC* promoter fragment of pRK21659	This study
pRK21661	Cm^r Sp^r	*klcA*$^+$ *klcB'* Tn*1*::Ω	pRK2320 with transcriptional terminator of ΩSp at *Pst*I site in Tn*1* remnant of pRK2320 (Fig. 5)	This study
pRK21663	Cm^r Sp^s	*klcA*$^+$ *klcB*::Ω	pRK2320 with transcriptional terminator of ΩSp at *Sph*I site in *klcA* (Fig. 5)	This study
pRK21666	Cm^r Sp^r	*klcA*::Ω	pRK2320 with transcriptional terminator of ΩSp at *Sph*I site in *klcB'* (Fig. 5)	This study
pRK21669	Ap^r Km^r	*klcA2*	pCRII with *klcA* structural gene starting at codon 2	This study
pRK21670	Ap^r	Φ(T7 φ10*p*-T7 · Tag-*klcA2*)	pET-17b with *klcA* structural gene from codon 2 fused in frame to 34-codon ORF downstream of bacteriophage T7 φ10 promoter	This study
pUC18	Ap^r		Cloning vector	62
pUC19	Ap^r		Cloning vector	62
pUZ8	Hg^r Km^r Tc^r	Ap^s; lacks Tn*1*	IncPα plasmid (isolate from Spain)	22
pVEX1211	Sp^r	ΩSp	Cloning vector used as source of ΩSp transcriptional terminators	3
R995	Km^r Tc^r	Ap^s; lacks Tn*1*	IncPα plasmid (isolate from Hong Kong)	15
RK2	Ap^r Km^r Tc^r	*kilC*$^+$; contains Tn*1*	IncPα plasmid (isolate from United Kingdom)	24

kilC locus located between the origin of replication (*oriV*) and transposon Tn*1*. The nucleotide sequence also revealed a 10th copy of the 17-bp *oriV* iteron, and we discuss its significance and possible role in the replication of RK2.

MATERIALS AND METHODS

Nomenclature. Coordinates for the RK2 physical map are defined as the distance in kilobases from the unique *Eco*RI site (e.g., kb 10.4-to-11.9 region). If a relevant plasmid gene is not present in a bacterial strain, the genotype is designated with a superscript zero (e.g., *kor*^{0}). The genes of the *kilC* locus have been named according to the genetic nomenclature previously adopted for the *kil* loci of RK2 (17, 29). This system uses a three-letter genetic designation consisting of "*kl*" (for *kil*) followed by a letter indicating the specific *kil* locus. Accordingly, the genes of the *kilC* locus use the prefix *klc* (pronounced "klik") and are named *klcA* and *klcB*.

Bacteria and plasmids. The *E. coli* strains used in this study were BL21(DE3, pLysS) {F$^-$ *hsdS gal* [λD69 Φ(*lacUV5p*-T7 gene *1*)]} (54); DH5α (*supE44* Δ*lacU169 hsdR17 recA1 endA1 gyrA96 thi-1 relA1*) (18); MV10 (*thr-1 leuB6 thi-1 lacY1 tonA21 supE44 rfbD1* Δ*trpE5*) (23); and RP1894 (*korA*$^+$ *korB*$^+$ *korC*$^+$ *korE*$^+$ *korF*$^+$ *bla*$^+$ *trpE*$^+$) (64).

The plasmids used are described in Table 1. Unpublished plasmids were constructed as follows: pRK21509, by insertion of the *kilC*-specific *Sph*I fragment of pRK2320 into the *Sph*I site of pUC19; pRK21510, by ligation of the *kilC*-specific *Eco*RI-*Hin*dIII fragment of pRK21509 to *Eco*RI- and *Hin*dIII-digested pUC18; pRK21654 and pRK21655, by insertion of the *Taq*I-*Bst*YI fragment of pRK2320 into *Acc*I- and *Bam*HI-digested pUC19 and pUC18, respectively; pRK21659, by PCR amplification of the *kilC* promoter region using the oligonucleotide primers kilC-16 (5′-GTTC CATCCGCCCGCCTGGC-3′) and kilC-8 (5′-CAGCAT GCGCAGCATCGAG-3′), followed by ligation of the amplification product to pCRII; pRK21660, by ligation of the *kilC* promoter-containing *Eco*RI fragment from pRK21659 to *Eco*RI-digested pCH8; pRK21661, by ligation of ΩSp-containing *Pst*I fragment of pVEX1211 to *Pst*I-digested pRK2320; pRK21663, by replacement of the 384-bp *Sph*I fragment of pRK2320 with *Sph*I-cleaved pVEX1211, which contains ΩSp; pRK21666, by partial digestion of pRK21661 with *Sph*I and screening for loss of the spectinomycin

resistance (Sp^r)-encoding portion of ΩSp; pRK21669, by PCR amplification of *klcA* from pRK2320 using the oligonucleotide primers klcA-1 (5′-GACTGATGTTCAAATC CCCTCCCC-3′) and klcA-2 (5′-GCCCCTTAGTCGATG GCCCGGTAG-3′) followed by ligation of the amplification product to pCRII; and pRK21670, by ligation of the *klcA*-containing *Eco*RI fragment from pRK21669 to *Eco*RI-digested pET-17b. Plasmids pRK21659, pRK21660, pRK21661, pRK21663, and pRK21666 were constructed in a $korA^+$ $korC^+$ host.

Media. Media for growth of bacteria were LB (48), M9 (48), and M9 containing 0.5% Casamino Acids (Difco Laboratories, Detroit, Mich.). Antibiotics at the following concentrations were used: ampicillin, 100 μg/ml; chloramphenicol, 50 μg/ml; penicillin, 150 μg/ml; rifampin, 200 μg/ml; spectinomycin, 50 μg/ml; and trimethoprim, 50 μg/ml. To induce expression of T7 RNA polymerase in BL21(DE3, pLysS), the medium was supplemented with 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG).

DNA methodologies. Preparation of DNA was done according to an alkaline lysis protocol (1), including polyethylene glycol precipitation for DNA to be used as template for sequencing. Agarose gel and polyacrylamide gel electrophoresis (PAGE) have been described previously (17). DNA manipulations with restriction endonucleases, T4 DNA ligase, and T4 polynucleotide kinase were done according to the manufacturers' recommendations. Amplification of DNA by PCR was done with *Taq* DNA polymerase (47). DNA fragments were purified from gels by the crush-and-soak method (34). Transformation of *E. coli* was done by the method of Cohen et al. (9).

The complete nucleotide sequence for the *kilC* region was determined for both strands by primer walking with double-stranded template from pRK2161 and pRK2320 (Table 1). In addition, 70% of the *kilC* region sequence was confirmed by using double-stranded DNA template prepared from RK2. The nucleotide sequence was determined by the dideoxy chain termination method (49) by using either the Sequenase version 2.0 sequencing kit (U.S. Biochemicals) and labeling with [α-^{35}S]dATP or the Δ*Taq* Cycle Sequencing kit (U.S. Biochemicals). For the latter, radioactive labeling of the products was done by including [α-^{33}P]dATP in the reaction mixture or by using oligonucleotide primers end labeled with [γ-^{33}P]ATP. For sequencing of the high-molecular-weight plasmids RK2, R995, and pUZ8, it was necessary to use 3 μg of template DNA and the Δ*Taq* Cycle Sequencing kit. The nucleotide sequences of comparable regions of RK2, R995, and pUZ8 were determined with the primer kilC-27 (5′-GTCGCGG GCGTCCAGATACC3′), which hybridizes to the *orfX* region (28). To minimize band compressions arising from the high G+C content of RK2, sequencing reactions were also done with deaza-7-dGTP (4). To determine the sequence of the *kilC* fragment in pRK21654 and pRK21655, the reaction products were separated on 6% polyacrylamide gels containing 8 M urea and 40% formamide (1) to prevent band compression.

Computer analysis of the *kilC* nucleotide sequences and predicted polypeptides was done using an Excel spreadsheet to evaluate G+C usage (26), DNA Strider (33), MacVector (International Biotechnologies, New Haven, Conn.), a modified version of the MacTargSearch 1.1 program (39), and software from the Genetics Computer Group package (12). Protein alignments were done with the default settings for the BESTFIT algorithm of the Genetics Computer Group package.

RNA analysis. The transcription start point for the *kilC* operon was determined by primer extension analysis, as previously described (29, 58). A 300-bp fragment containing the *kilC* promoter was inserted into pCH8, which carries a transcriptional terminator downstream of the cloning site, to produce pRK21660 (described above). RNA was extracted from exponential-phase cultures. An oligonucleotide primer (35-mer) complementary to the transcribed region downstream of the promoter in pRK21660 (58) was end labeled with [γ-^{32}P]ATP, annealed to the RNA, and extended by avian myeloblastosis virus reverse transcriptase. The cDNA products were separated by electrophoresis in polyacrylamide sequencing gels and visualized by autoradiography. The 5′ end of the *kilC* transcript was mapped by comparing the primer extension product to the products of a sequencing reaction done with pRK21660 as the template and the end-labeled primer.

To assay for the effects of *korA* and *korC* on the relative abundance of *kilC* transcripts, primer extension analysis was done with two primers in the reaction mixture, as previously described (29, 58). One primer (35-mer) was specific for *kilC* RNA (described above); the second primer was specific for RNA initiated from the constitutive *bla* promoter on pRK21660. The level of *bla*-specific cDNA served as an internal control (58). The cDNA products were separated by electrophoresis in polyacrylamide sequencing gels and visualized by autoradiography. The relative levels of *kilC* RNA from kor^0, $korA^+$, $korC^+$, and $korA^+$ $korC^+$ strains were determined after normalizing the samples to *bla* RNA. Quantitation was done with a Betagen Betascope 603 blot analyzer.

Polypeptide analysis. The *klcA* coding region, beginning with the second codon, was fused in frame to a 34-codon open reading frame (ORF) whose transcriptional and translational initiation signals were provided by the bacteriophage T7 φ gene *10* in the plasmid vector pET-17b. This 34-codon ORF includes a 12-codon leader peptide of gene *10*. The *klcA* fusion was designated *klcA2*, and the resulting plasmid was pRK26669 (described above). The *klcA* gene fusion was induced in strain BL21(DE3, pLysS), which contains an IPTG-inducible T7 RNA polymerase gene, as previously described (54). Induction was accomplished as follows. Cells were grown in LB broth to an optical density at 600 nm of 0.6; T7 polymerase expression was then induced by adding IPTG to a final concentration of 0.4 mM. The culture was incubated at 37°C for 2 h; cells were then collected by centrifugation at 4,000 × *g* for 10 min. The pellet was resuspended in 1/20 of the original culture volume in 10 mM Tris-HCl–25 mM EDTA (pH 8.0), and cells were frozen at −70°C. Cells were thawed, Triton X-100 was added to a final concentration of 0.1%, and cells were incubated at room temperature for 15 min. $MgCl_2$ and DNase I were added to final concentrations of 10 mM and 20 μg/ml, respectively, and cells were incubated at room temperature for 15 min. After centrifugation at 6,000 × *g* for 15 min to remove debris, the extract was stored at −70°C. For Western (immunoblot) analysis, the chemiluminescence Amersham ECL kit and rainbow molecular weight markers were used. T7 · Tag monoclonal antibody directed against the 12-codon T7 gene *10* leader peptide was obtained from Novagen (Madison, Wis.).

Nucleotide sequence accession number. The nucleotide sequence and amino acid sequence data reported here have been submitted to the GenBank database under the accession number U05773.

RESULTS AND DISCUSSION

Determination and analysis of the nucleotide sequence of the *kilC* region. The *kilC* determinant, located in the region between the origin of replication (*oriV*) and transposon Tn*1* on RK2 (kb 12.2 to 10.5; Fig. 1), is lethal to *E. coli* host cells unless

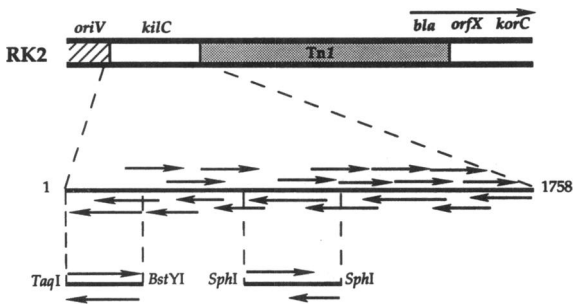

FIG. 1. Sequencing strategy for the *kilC* region. Shown here is the *kilC* region within a partial genetic and physical map of plasmid RK2. The kb 13-to-4.5 region of RK2 is expanded below the RK2 map. Arrows indicate regions sequenced by primer walking. The numbers refer to nucleotides in Fig. 2. The lowest portion of the figure shows regions that were subcloned for further sequencing. Genetic designations are described in the text.

korA and *korC* are present to regulate its expression (15, 64). We determined the nucleotide sequence of the *kilC* region (Fig. 2) for both strands from two independent *kilC* clones by primer walking (Fig. 1). In addition, 70% of the nucleotide sequence of this region was confirmed by sequencing directly from RK2 by PCR-based sequencing (1). Curiously, a small segment of the *kilC* region (nucleotides 1 through 250) was refractory to sequencing on one strand when the plasmid templates contained the complete *oriV-kilC* region. The problem was solved by subcloning a small fragment into pUC18 and pUC19 prior to sequencing (Fig. 1).

The nucleotide sequence was examined for ORFs. The entire *kilC* region contained only one complete ORF that is preceded by a good match to the Shine-Dalgarno sequence and exhibits the distinctive codon usage pattern that is characteristic of RK2 genes (17). This ORF occurs downstream of a sequence that is predicted to be a strong σ^{70} promoter (19, 21, 56). The putative promoter contains the palindromic sequences thought to be binding sites for the KorA and KorC repressors (42) that regulate the *kilC* host-lethal phenotype (15). Because the ORF seemed likely to be an RK2 gene, we named it *klcA*. The designation *klc* refers to *kilC*, in accordance with our previously established nomenclature for genes of the *kil* loci (17, 29).

klcA is predicted to encode a polypeptide product with a molecular mass of 15,856 Da. To detect a *klcA*-directed polypeptide product, the second codon of the putative coding region was fused in frame to a 34-codon ORF whose translation initiation signals were provided by the pET-17b vector. In this plasmid (pRK21670) the *klcA* fusion is situated downstream of the bacteriophage T7 ϕ10 promoter, which was induced as described in Materials and Methods. Sodium dodecyl sulfate (SDS)-PAGE and Western analysis of extracts from induced cells (Fig. 3) revealed a 19-kDa polypeptide, which corresponds well to the predicted full-length fusion product (19,146 Da), as well as a number of smaller polypeptides. The smaller polypeptides are most likely the result of degradation of the full-length product, since the monoclonal antibody recognizes the amino terminus of the fusion protein.

Immediately downstream of *klcA* is another large ORF that is interrupted by transposon Tn*1* sequences. This ORF also has a reasonable Shine-Dalgarno sequence and displays RK2-like codon usage, suggesting that it may have been an RK2 gene. We previously reported a large ORF (*orfX*) that begins at the other junction of Tn*1* and terminates just prior to *korC* (Fig. 1) (28). The codon usage of *orfX* indicated that it may be the 3′ end of an RK2 gene interrupted by Tn*1*. Comparison of the sequences at both ends of Tn*1* revealed a 5-bp duplication that is typical of Tn*1* transposition events (50). If Tn*1* and one copy of the duplicated sequence (5′-TTTTA-3′) are ignored, *orfX* is contiguous with and in frame to the interrupted ORF in the *kilC* region. To confirm that the incomplete *kilC* ORF and *orfX* are two parts of the same gene, we examined two other IncPα plasmids, R995 and pUZ8, that are closely related to RK2 but lack Tn*1* (60). Using a primer complementary to *orfX*, we sequenced the regions of R995 and pUZ8 that correspond to the putative Tn*1* insertion site in RK2. The sequences were identical for both plasmids and confirmed that the acquisition of Tn*1* by RK2 was the result of a simple insertion that resulted in the duplication of the 5-bp target without any other DNA rearrangements. Thus, the incomplete ORF of the *kilC* region appears to be part of an interrupted RK2 gene that includes *orfX*. We have named the complete gene *klcB*.

The intergenic region between *klcA* and *klcB* contains overlapping palindromic sequences. One sequence (i′ in Fig. 2) is related to a sequence found twice in the region upstream of the *kilC* promoter (as discussed below) and to a palindrome located between the *kleE* and *kleF* genes of the *kilE* locus (29). The i′ sequence overlaps a second large, nearly perfect, inverted repeat. The proximity of this sequence to the ribosome binding site for *klcB* raises the possibility of its involvement in the initiation of *klcB* translation.

Transcriptional start site and regulation of the *kilC* promoter. We used primer extension analysis to determine the transcriptional start site for the *kilC* promoter (*kilCp*) and to assay the effects of *korA* and *korC* on the level of transcription. We have shown previously that unregulated expression of strong promoters from the RK2 *kil-kor* regulon can interfere with plasmid maintenance or expression of plasmid-selective markers (3, 17). Plasmid pRK21660, which contains the *kilC* promoter, can be maintained in *E. coli* in the absence of *kor* functions because it has the *rrnB* transcriptional terminator downstream of the promoter cloning site. RNA was extracted from a strain containing pRK21660 and used for primer extension analysis with reverse transcriptase (Fig. 4). The major cDNA product corresponds to the expected transcription start site of the *kilC* promoter (Fig. 2), thus confirming that this predicted promoter is functional in *E. coli*. Sequence comparison of the *kilC* promoter with the *E. coli* σ^{70} promoter consensus sequence revealed that *kilCp*, like other promoters of the *kil-kor* regulon, is predicted to be a strong promoter (19, 21).

A modification of the primer extension assay was used to determine the effects of *korA* and *korC* on the level of RNA initiated from the *kilC* promoter (*kilC* RNA). A second primer specific for RNA initiated from the constitutively expressed *bla* promoter of pRK21660 was used to provide an internal control for the amount of RNA, as described in Materials and Methods. Densitometric analysis showed that the abundance of *kilC* RNA relative to that of *bla* RNA was reduced more than 100-fold in the $korA^+$ $korC^+$ strain (containing pRK2292 and pRK2462) (Fig. 4 and data not shown). However, the $korA^0$ $korC^+$ strain (containing pRK2462) showed only a two- to threefold reduction in *kilC* RNA, and the level of *kilC* RNA was unaffected in the $korA^+$ $korC^0$ strain (containing pRK2292) (Fig. 4 and data not shown). Thus, while *korC* is able to reduce the level of RNA initiated at the *kilC* promoter, the tightest regulation occurs when both *korA* and *korC* are present. These results agree with genetic experiments showing that constitutively expressed *korC* (pRK2462), but not *korA* (pRK2292), can prevent the host-lethal phenotype expressed

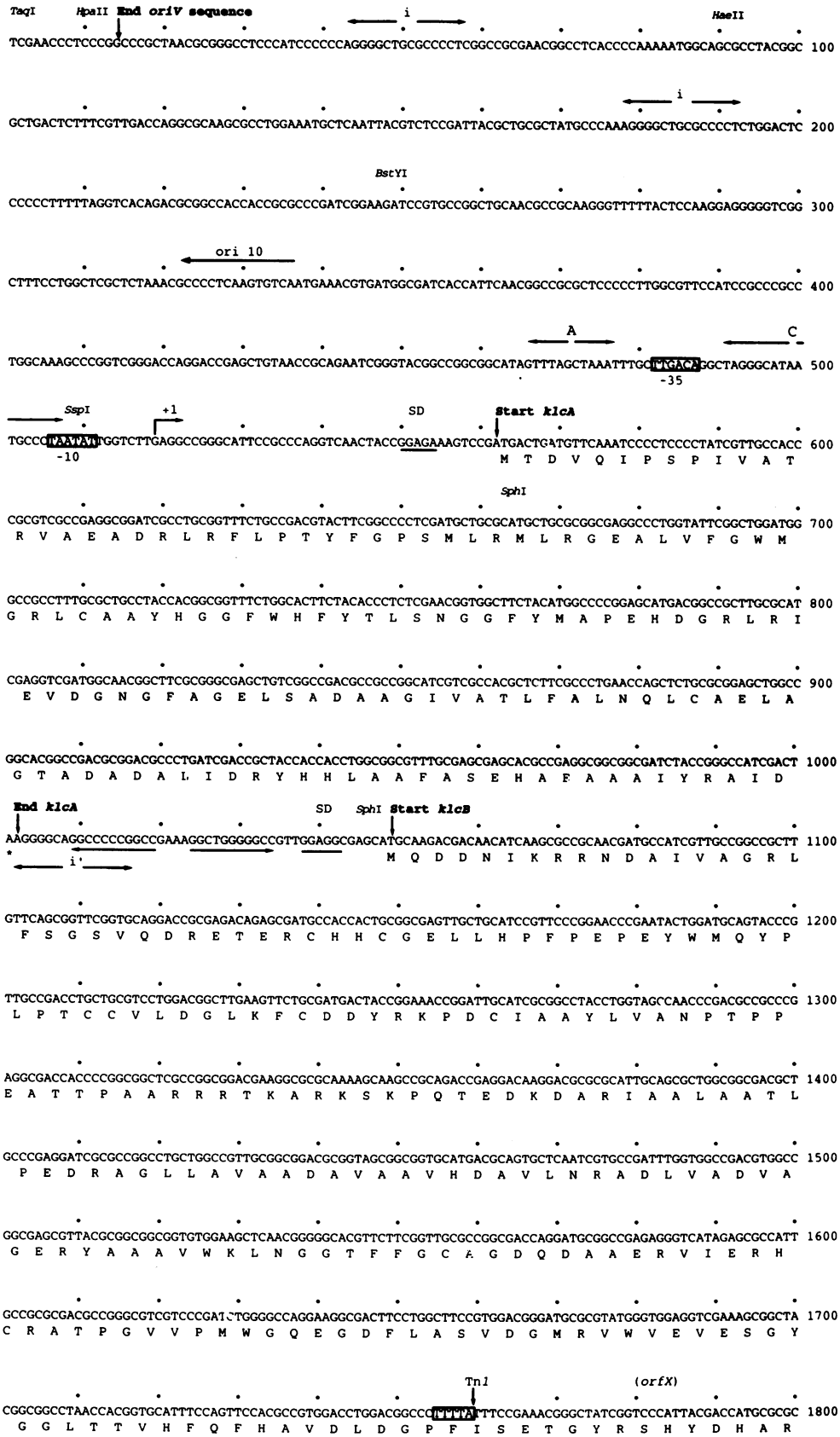
TaqI
HpaII
End oriV sequence
i
HaeII
TCGAACCCTCCCGGCCCGCTAACGCGGGCCTCCCATCCCCCCAGGGGCTGCGCCCCTCGGCCGCGAACGGCCTCACCCCAAAAATGGCAGCGCCTACGGC 100
i
GCTGACTCTTTCGTTGACCAGGCGCAAGCGCCTGGAAATGCTCAATTACGTCTCCGATTACGCTGCGCTATGCCCAAAGGGGCTGCGCCCCTCTGGACTC 200
BstYI
CCCCCTTTTTAGGTCACAGACGCGGCCACCACCGCGCCCGATCGGAAGATCCGTGCCGGCTGCAACGCCGCAAGGGTTTTTACTCCAAGGAGGGGGTCGG 300
ori 10
CTTTCCTGGCTCGCTCTAAACGCCCCTCAAGTGTCAATGAAACGTGATGGCGATCACCATTCAACGGCCGCGCTCCCCCTTGGCGTTCCATCCGCCCGCC 400
A
C
TGGCAAAGCCCGGTCGGGACCAGGACCGAGCTGTAACCGCAGAATCGGGTACGGCCGGCGGCATAGTTTAGCTAAATTTGCTTGACAGGCTAGGGCATAA 500
-35
SspI
+1
SD
Start klcA
TGCCGTAATATGGTCTTGAGGCCGGGCATTCCGCCCAGGTCAACTACCGGAGAAAGTCCGATGACTGATGTTCAAATCCCCTCCCCTATCGTTGCCACC 600
-10
M T D V Q I P S P I V A T
SphI
CGCGTCGCCGAGGCGGATCGCCTGCGGTTTCTGCCGACGTACTTCGGCCCCTCGATGCTGCGCATGCTGCGCGGCGAGGCCCTGGTATTCGGCTGGATGG 700
R V A E A D R L R F L P T Y F G P S M L R M L R G E A L V F G W M
GCCGCCTTTGCGCTGCCTACCACGGCGGTTTCTGGCACTTCTACACCCTCTCGAACGGTGGCTTCTACATGGCCCCGGAGCATGACGGCCGCTTGCGCAT 800
G R L C A A Y H G G F W H F Y T L S N G G F Y M A P E H D G R L R I
CGAGGTCGATGGCAACGGCTTCGCGGGCGAGCTGTCGGCCGACGCCGCCGGCATCGTCGCCACGCTCTTCGCCCTGAACCAGCTCTGCGCGGAGCTGGCC 900
E V D G N G F A G E L S A D A A G I V A T L F A L N Q L C A E L A
GGCACGGCCGACGCGGACGCCCTGATCGACCGCTACCACCACCTGGCGGCGTTTGCGAGCGAGCACGCCGAGGCGGCGGCGATCTACCGGGCCATCGACT 1000
G T A D A D A L I D R Y H H L A A F A S E H A F A A A I Y R A I D
End klcA
SD
SphI
Start klcB
AAGGGGCAGGCCCCCGGCCGAAAGGCTGGGGGCCGTTGGAGGCGAGCATGCAAGACGACAACATCAAGCGCCGCAACGATGCCATCGTTGCCGGCCGCTT 1100
M Q D D N I K R R N D A I V A G R L
i'
GTTCAGCGGTTCGGTGCAGGACCGCGAGACAGAGCGATGCCACCACTGCGGCGAGTTGCTGCATCCGTTCCCGGAACCCGAATACTGGATGCAGTACCCG 1200
F S G S V Q D R E T E R C H H C G E L L H P F P E P E Y W M Q Y P
TTGCCGACCTGCTGCGTCCTGGACGGCTTGAAGTTCTGCGATGACTACCGGAAACCGGATTGCATCGCGGCCTACCTGGTAGCCAACCCGACGCCGCCCG 1300
L P T C C V L D G L K F C D D Y R K P D C I A A Y L V A N P T P P
AGGCGACCACCCCGGCGGCTCGCCGGCGGACGAAGGCGCGCAAAAGCAAGCCGCAGACCGAGGACAAGGACGCGCGCATTGCAGCGCTGGCGGCGACGCT 1400
E A T T P A A R R R T K A R K S K P Q T E D K D A R I A A L A A T L
GCCCGAGGATCGCGCCGGCCTGCTGGCCGTTGCGGCGGACGCGGTAGCGGCGGTGCATGACGCAGTGCTCAATCGTGCCGATTTGGTGGCCGACGTGGCC 1500
P E D R A G L L A V A A D A V A A V H D A V L N R A D L V A D V A
GGCGAGCGTTACGCGGCGGCGGTGTGGAAGCTCAACGGGGGCACGTTCTTCGGTTGCGCCGGCGACCAGGATGCGGCCGAGAGGGTCATAGAGCGCCATT 1600
G E R Y A A A V W K L N G G T F F G C A G D Q D A A E R V I E R H
GCCGCGCGACGCCGGGCGTCGTCCCGATCTGGGGCCAGGAAGGCGACTTCCTGGCTTCCGTGGACGGGATGCGCGTATGGGTGGAGGTCGAAAGCGGCTA 1700
C R A T P G V V P M W G Q E G D F L A S V D G M R V W V E V E S G Y
Tn1
(orfX)
CGGCGGCCTAACCACGGTGCATTTCCAGTTCCACGCCGTGGACCTGGACGGCCGTTTATTTCCGAAACGGGCTATCGGTCCCATTACGACCATGCGCGC 1800
G G L T T V H F Q F H A V D L D G P F I S E T G Y R S H Y D H A R

FIG. 2. Nucleotide sequence of *kilC* region. Nucleotides 1 to 14 and 1759 to 1800 were previously reported as part of the *oriV* (53) and *orfX* (28) regions, respectively. Landmark restriction sites are indicated above the sequence. Divergent arrows indicate inverted repeats. A and C show the predicted targets for KorA and KorC, respectively. One copy of the *oriV* iteron sequence (ori 10), a palindromic repeat that is related to the *oriV* iteron and occurs twice in the sequence (i), and the closely related sequence (i') are also indicated. The predicted amino acid sequences are shown in single-letter code below the nucleotide sequences of *klcA* and *klcB'*. The −10 and −35 regions of the *kilC* promoter are boxed, and the mRNA start site is indicated by +1 and an angled arrow. A 118-bp sequence suggested to contain the *kilC* promoter by Thomas et al. (56) differs from the sequence of the *kilC* promoter region shown in this figure as follows: by the inclusion of a T between nucleotides 504 and 505 and by the substitution of C, T, G, T, T, and C at nucleotides 514, 515, 517, 522, 573, and 574, respectively. Potential Shine-Dalgarno (SD) sequences for ribosome binding (52) are underlined. Tn*1* shows the site of insertion in *klcB*, and the box at Tn*1* indicates nucleotides that were duplicated upon insertion of the transposon.

by *kilC* (data not shown). In addition, it was previously shown that *korA* alone cannot control the lethal phenotype of *kilC* (15) but that a requirement for *korA* is observed when *korC* is limiting (28, 64).

Molecular and genetic studies have indicated that KorA and KorC are DNA-binding proteins that act as transcriptional repressors (42). The *kilC* promoter contains a putative KorC-binding site that overlaps the −10 region and a predicted KorA target that occurs 5 bp upstream of the −35 region (Fig. 2). The identical arrangement and spacing of KorA and KorC binding sites occur in the *kleA* and *kleC* promoters of the *kilE* locus, which are also regulated by *korA* and *korC* (29). However, both the *kleA* and *kleC* promoters are significantly affected by both *korA* and *korC* individually (29). The primer extension analysis and the genetic evidence that *korA* alone cannot control the lethal phenotype of *kilC* suggest that the interactions of KorA and KorC with the *kilC* promoter differ from their interactions with the *kleA* and *kleC* promoters.

Structure of the intact *kilC* operon. Our results show that the uninterrupted *kilC* operon of IncPα plasmids consists of three genes (*klcA*, *klcB*, and *korC*) followed by a potential Rho-independent transcription terminator (Fig. 5). The presence of *korC* demonstrates that the intact *kilC* operon is autoregulated. In RK2, the operon is disrupted by a transposon Tn*1* insertion in *klcB* and the expression of *korC* depends on the strong β-lactamase promoter within Tn*1* (28). The interruption of the *kilC* operon by Tn*1* may be significant with respect to the genetic structure of RK2. It is possible that the constitutive expression of *korC* and/or the disruption of *klcB* resulted in the selection of other compensatory mutations in RK2. Thus, while RK2 and RP4 are the paradigms for the IncPα group of promiscuous plasmids, they must nevertheless be considered mutants of a prototypic IncPα plasmid.

Unregulated expression of *klcA* is lethal to *E. coli* host cells. The *kilC* region present on plasmid pRK2320 is lethal to *E. coli* host cells in the absence of *korA* and *korC* (15). To identify the determinant responsible for the lethal (Kil$^+$) phenotype, we inserted the strong transcriptional terminators from the ΩSp fragment (45) at three positions in the *kilC* region on pRK2320 (Fig. 5): (i) downstream of the *kilC* region (Tn*1*::Ω), (ii) within *klcB* near its 5′ end (*klcB*::Ω), and (iii) within *klcA* (*klcA*::Ω). The mutant plasmids were constructed in the presence of *korA* and *korC* and then tested for their ability to prevent the production of viable colonies following transformation of a strain lacking *korA* and *korC* (Kil$^+$ phenotype). Like the pRK2320 parental plasmid, the Tn*1*::Ω mutant (pRK21661)

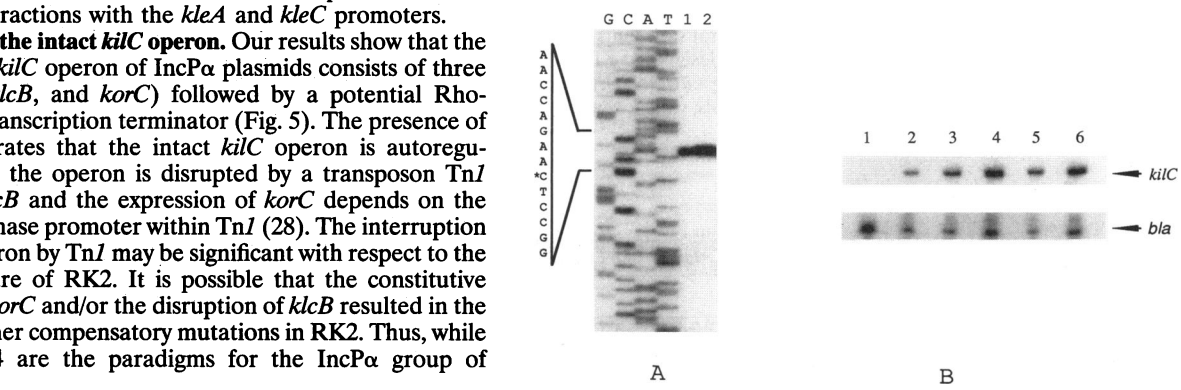

FIG. 3. Polypeptide product of a *klcA* fusion. *klcA* was fused to a 34-codon ORF downstream of the bacteriophage T7 ϕ10 promoter. After induction, the polypeptide was separated by SDS-PAGE, transferred to nitrocellulose, probed with T7 · Tag antibody, and visualized by chemiluminescence. Lanes: 1, pET-17b; 2, pRK21669. The arrow on the right indicates the full-length *klcA* fusion protein. The numbers on the left show the positions and sizes of the following protein markers: ovalbumin (46 kDa), carbonic anhydrase (30 kDa), trypsin inhibitor (21.5 kDa), lysozyme (14.3 kDa), and aprotinin (6.5 kDa).

FIG. 4. Transcriptional start site of the *kilC* operon and effects of *korA* and *korC* on the abundance of RNA initiated from the *kilC* promoter. The ^{32}P-end-labeled 35-mer primer was annealed to RNA extracted from cells containing pRK21660 (*kilCp*$^+$) and extended by avian myeloblastosis virus reverse transcriptase. The labeled DNA products were separated in polyacrylamide sequencing gels and visualized by autoradiography, as described in Materials and Methods. The reference sequences were obtained by the PCR-based dideoxynucleotide chain termination method (1) with the ^{32}P-end-labeled 35-mer primer and plasmid pRK21660. (A) The primer extension product (lanes 1 and 2) from a pRK21660-containing strain is shown next to the reference sequence. An asterisk indicates the transcriptional start site. The sample in lane 1 is a 1:2 dilution of that in lane 2. (B) To determine the effects of *korA* and *korC* on the level of RNA initiated from the *kilC* promoter, RNA was extracted from strains containing pRK21660 in the presence and absence of *korA* and *korC*. The ^{32}P-end-labeled 35-mer and *bla* primers were annealed to RNA extracted from the various strains and extended by avian myeloblastosis virus reverse transcriptase. Primer extension products are shown for the RNA for strains containing pRK21660 (*kilCp*$^+$) and the following plasmids: pRK2292 (*korA*$^+$) and pRK2462 (*korC*$^+$) (lane 1), pRK2462 (*korC*$^+$) (lane 2), pRK2292 (*korA*$^+$) diluted 1:2 (lane 3), pRK2292 (*korA*$^+$) undiluted (lane 4), no other plasmid (*kor*^{0}) and diluted 1:2 (lane 5), and *kor*^{0} undiluted (lane 6).

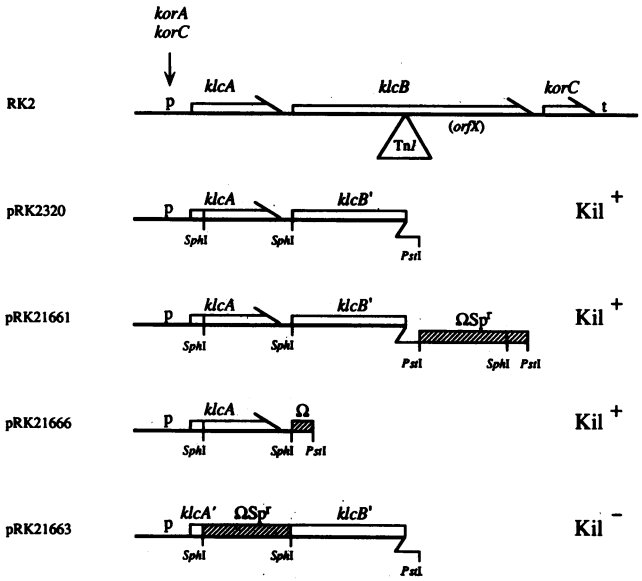

FIG. 5. Mapping the Kil$^+$ host-lethal determinant in the *kilC* operon. The *kilC* operon is depicted in the top line. The position of transposon Tn*1* in RK2 is indicated by the triangle below the line. The *kilC* promoter (P) and its negative regulation by *korA* and *korC* (vertical arrow) are indicated. The locations of transcriptional terminators (Ω or ΩSp^r) are shown for pRK21661, pRK21663, and pRK21666. Plasmids were used to transform *kor*^{0} and *korA*$^+$ *korC*$^+$ strains. Plasmids that produced viable transformants of the *kor*^{0} strain at 1,000-fold-less efficiency than that of the *korA*$^+$ *korC*$^+$ strain are designated Kil$^+$; plasmids that gave rise to viable transformants of both strains with approximately equal efficiencies are designated Kil$^-$. The strains were transformed with a plasmid lacking *kilC* to correct for competence differences of the *kor*^{0} and *kor*$^+$ strains, which were never more than twofold.

produced at least 1,000-fold fewer colonies of the *kor*^{0} strain than of an equally competent *korA*$^+$ *korC*$^+$ strain. Thus, as expected, the Tn*1*::Ω plasmid conferred a Kil$^+$ phenotype. In contrast, the *klcA*::Ω insertion mutant (pRK21663) was able to produce viable transformants of both the *korA*$^+$ *korC*$^+$ and *kor*^{0} strains with equal efficiency. Thus, it was defective in host cell killing (Kil$^-$). To determine if *klcA* is sufficient for host lethality, we examined the *klcB*::Ω mutant (pRK21666) and found that it conferred a Kil$^+$ phenotype in the absence of *korA* and *korC*. We conclude that the portion of *klcB* present in the *kilC* region is not required for the Kil$^+$ phenotype and that *klcA* is lethal to *E. coli* host cells in the absence of *korA* and *korC*.

KlcA is related to the ArdB antirestriction protein of IncN plasmid pKM101. The *klcA* product is predicted to be an acidic polypeptide of 146 amino acids with a molecular mass of 15,856 Da. Computer analysis of the GenBank database revealed that the predicted *klcA* product has significant relatedness (31% identity; 56% similarity) to the 16.5-kDa ArdB protein of the IncN plasmid pKM101 (Fig. 6A). ArdB is encoded by the leading region of pKM101 that is transferred early during conjugation and has been shown to possess an antirestriction activity directed primarily against type I restriction endonucleases (7). Like ArdB, the predicted KlcA polypeptide shows no significant similarity to other known proteins with antirestriction activity (ArdA of pKM101 [6], Ard of ColIb-P9 [11, 46], and the product of the 0.3 gene of bacteriophage T7 [13]). The regulation of *ardB* is strikingly similar to the regulation of *klcA*. The *ardB* promoter is inhibited by dual transcriptional repressors, one of which is encoded by the *ardB* operon (7). However, a 9-amino-acid motif found in the various antirestriction proteins and hypoth-

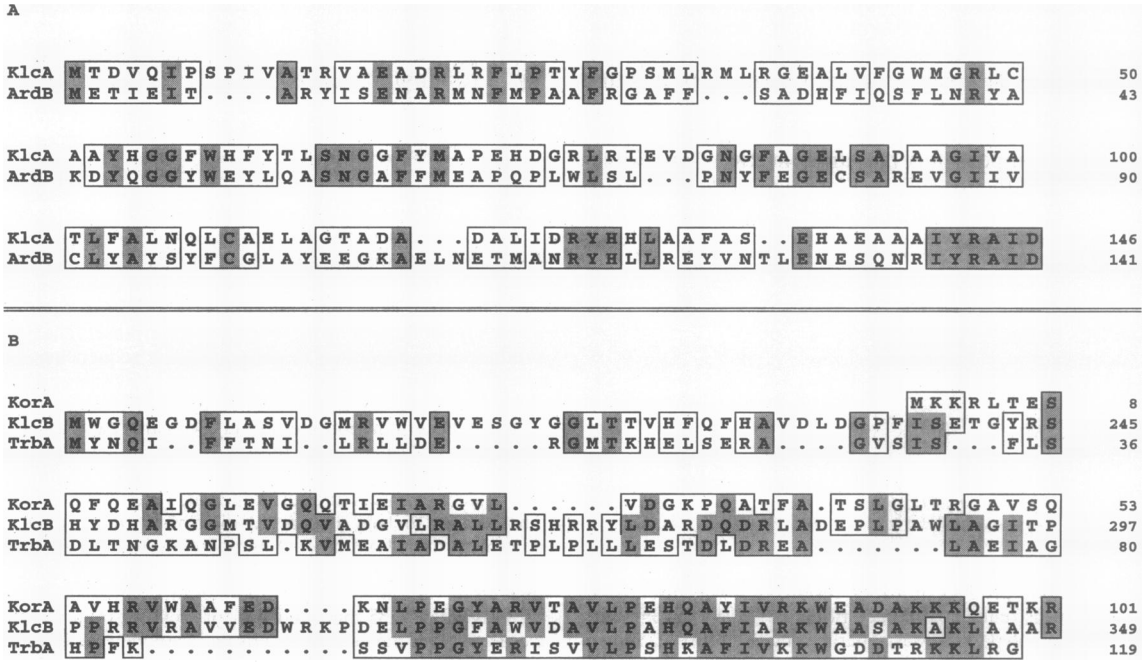

FIG. 6. Relationship of KlcA to the ArdB antirestriction protein of plasmid pKM101 (A) and KlcB to KorA and TrbA of RK2 (B). Comparison of the predicted polypeptides was done with the default settings of the BESTFIT program of the Genetics Computer Group software package. Identities (dark shading) and similarities (unshaded boxes) are indicated. The number at the right of each line indicates the position of the last amino acid on that line. In these alignments, KlcA and ArdB show 31% identity and 56% similarity, KlcB and KorA show 37% identity and 60% similarity, and KlcB and TrbA show 32% identity and 56% similarity.

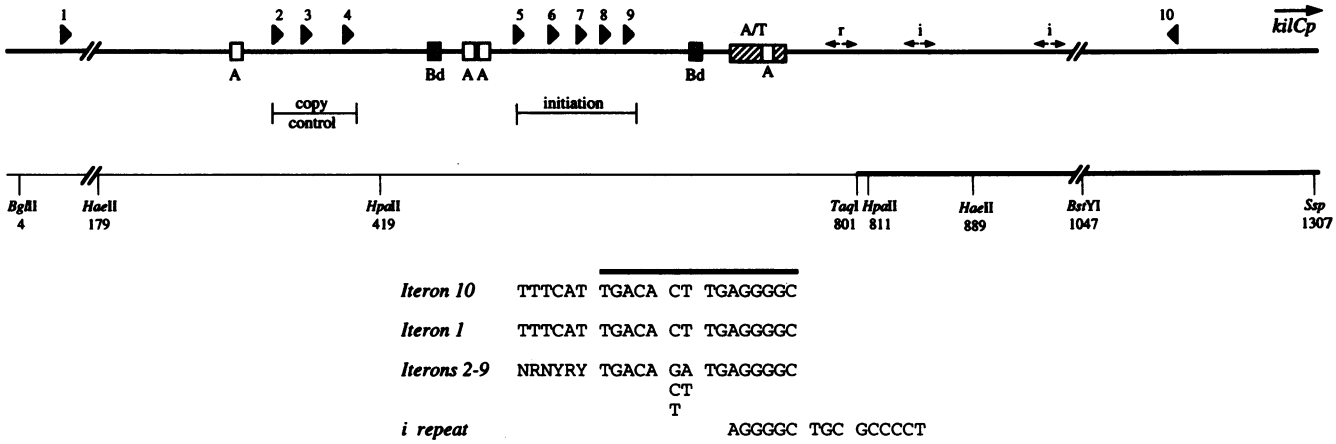

FIG. 7. Physical map of the *oriV* region of plasmid RK2. The 10 17-bp iterons are depicted as numbered arrowheads above the top line; relative orientation is indicated by the direction of the arrowhead. Iterons 1 and 10 are orphan iterons, the cluster of iterons 2 to 4 is involved in control of copy number, and iterons 5 to 9 are essential for initiation of replication (42). Double slashes crossing the lines mark regions of discontinuity in the physical map. *kilCp* is the *kilC* promoter, with the arrow indicating the direction of transcription. Divergent arrows show a previously reported palindromic sequence (r) (53) and two identical palindromes (i) reported in this work. Open boxes show predicted DnaA-binding sites (31, 42). An AT-rich region (42) is indicated by the shaded rectangle labeled A/T. Horizontally striped boxes designated Bd show the positions of predicted bending sequences (14, 30). The line below the map indicates distance in base pairs and landmark restriction sites. The thin portion of the line shows the region whose sequence was known (53, 61), while the thick portion indicates a 520 bp-region of the nucleotide sequence reported here. The *Taq*I site at position 801 corresponds to the *Taq*I site at position 1 in Fig. 2. The nucleotide sequences of the iterons and the related i palindrome are shown below the physical map. The sequence corresponding to the 17-bp iteron is indicated by a thick line, and the variation at positions 6 and 7 in iterons 2 to 9 is also shown. The 6-bp sequence essential for TrfA binding at the iteron (43) is shown to the left of the sequence. Iterons 1 and 10 are the only two iterons that are identical in this region. The 6-bp consensus sequence for iterons 2 to 9 is indicated (N, any base; R, A or G; Y, C or T).

esized to be an interaction site for antirestriction proteins with restriction endonucleases (7) is not present in KlcA.

To determine if the *kilC* region expresses an antirestriction activity, we tested RK2, R995, pUZ8, and the high-copy-number *kilC* plasmid derivatives shown in Fig. 5 for the ability to alleviate restriction of unmodified bacteriophage λ by the type I K-restriction system of *E. coli* (*Eco*K). We found no evidence for antirestriction activity in this assay, since the number of plaques varied no more than twofold on the various strains (data not shown). There are several possible explanations for the failure to detect antirestriction activity. (i) Antirestriction may be expressed only during conjugal transfer of the plasmid into the recipient cell. Once regulation of the *kilC* operon has been established, the expression of *klcA* may be insufficient to inhibit restriction. It seems likely that a plasmid, once established in a new host, would not interfere with the security conferred by the host restriction system. (ii) KlcA may be an antirestriction function, but its activity may be specific for certain restriction endonucleases of other bacterial hosts. (iii) Finally, despite its similarity to ArdB, KlcA may have a function other than antirestriction. Genetic analysis of the conjugal transfer properties of an RK2 mutant with *klcA* deleted will help distinguish among these possibilities.

KlcB, KorA, and TrbA are related. The predicted product of the intact *klcB* gene (including *orfX* [28]) is a polypeptide of 461 amino acids with a molecular mass of 51,133 Da. We found that a segment of KlcB is related to the RK2 proteins KorA (5) and TrbA (25) (Fig. 6B). The similarity to TrbA extends into the KlcB segment encoded by sequences upstream of the Tn*1* insertion. The relationship of KorA, TrbA, and the *orfX*-encoded portion of *klcB* has also been noted previously (25). The similarity is particularly striking in the terminal 35 amino acids of the smaller KorA and TrbA proteins (exhibiting, respectively, 36% identity and 60% similarity with KorA and 32% identity and 56% similarity with TrbA). KorA and TrbA (both 11.3 kDa) are regulatory proteins that affect RK2 gene expression. The similarity of KlcB to these proteins raises the possibility that it is involved in regulation, although the larger size of KlcB may indicate additional functions. The common domain is likely to have an important function, perhaps in protein-protein interactions. Previous genetic studies with *korA* have shown that its regulatory activity is sensitive to deletions in the C-terminal coding region. Loss of the two C-terminal amino acids of *korA* results in a protein with reduced repressor activity, and loss of the nine C-terminal residues lead to an inactive repressor (5).

A 10th *oriV* iteron in the *kilC* upstream region: implications for *oriV* structure and function. The 450-bp region upstream of the *kilC* promoter is contiguous with the origin of replication (*oriV*) but lies outside the region shown to be essential for initiation of replication (53, 57) (Fig. 2 and 7). The region contains five small ORFs, but none is preceded by a Shine-Dalgarno sequence (51) nor do they display RK2-like codon usage (reference 17 and data not shown).

Our analysis of the sequence suggests that the *kilC* upstream region is part of the RK2 origin of replication. The previously reported *oriV* region (Fig. 7) includes multiple copies of a 17-bp sequence (iteron) that is bound by the TrfA replication initiator protein (42). The minimal functional origin of replication contains five closely spaced iterons that are essential for initiation. Another set of three iterons is involved in copy number control. Separated from this cluster by 262 bp is an individual "orphan" iteron having no known function. All nine iterons occur in the same orientation. We found that the *kilC* upstream region contains a 10th copy of the *oriV* iteron in the opposite orientation (Fig. 7). Iteron 10 is an orphan iteron located 490 bp from the five essential iterons and 145 bp upstream of the −35 region of the *kilC* promoter (Fig. 2). The

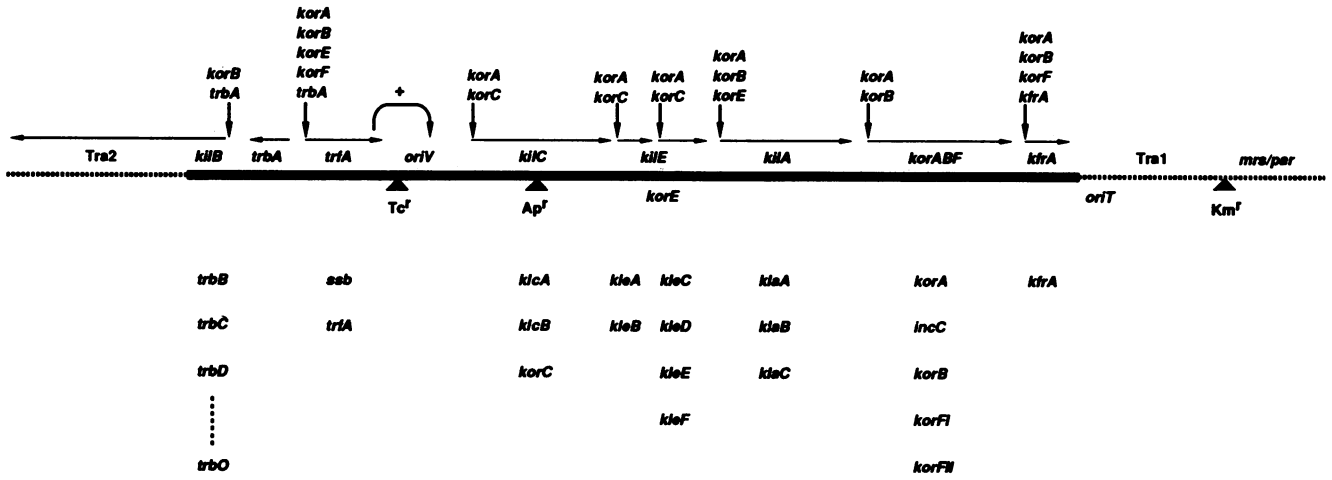

FIG. 8. The *kil-kor* regulon of IncPα plasmids. The 60-kb circular genome of RK2 is presented as a linear map. The thick line (representing about 26 kb) indicates the segment that contains the *kor*-regulated promoters of the *kil-kor* regulon. The regions depicted by dashed lines are not drawn to scale. Relative positions of specific loci, operons, and genes are shown immediately above and below the line. The curved arrow shows the positive interaction of the replication initiator function on *oriV*. The direction of transcription is designated by horizontal arrows. Downward arrows indicate determinants involved in negative regulatory interactions. The following genetic designations (42) are not described in the text: Tra1 and Tra2, two large regions required for conjugal transfer; *oriT*, the origin of conjugal transfer; and *mrs/par*, a locus involved in plasmid stability. The positions of genes for tetracycline resistance (Tc^r), ampicillin resistance (Ap^r) carried by transposon Tn*1*, and kanamycin resistance (Km^r) are indicated (arrowheads). Listed below the map are the individual genes of the *kor*-regulated operons. The Tra2 operon encodes at least 14 genes (*trbB* through *trbO*). The *kilB* locus includes *trbB* and *trbC* (58). (Note that *trbB*, *trbC*, and *trbD* were previously designated *klbA*, *klbB*, and *klbC*, respectively [58].)

region also shows two copies of a palindromic sequence (Fig. 2, sequence i) whose arms are identical to the 5′-AGGGGC-3′ segment of the *oriV* iteron sequence (Fig. 2 and 7). Recent studies have demonstrated that palindromic sequences based on a 9-bp segment of the 22-bp origin iteron of plasmid R6K bind a dimeric form of the R6K replication initiator protein (63). In this regard, methylation protection assays of the RK2 iterons have demonstrated that the portion of the iteron sequence contained within the palindromic i repeat is in close contact with TrfA (43). Thus, the i repeats may be binding sites for TrfA. Finally, we have noticed that the *kilC* upstream region contains unusual clustering of purine and pyrimidine tracts.

The nine previously identified iterons are not identical. Positions 6 and 7 of the iteron sequence can be CT (iterons 1, 3, 5, and 7), GA (iterons 2, 4, 6, 8), or TT (iteron 9) (Fig. 7) (42). Iteron 10 was found to be identical to iteron 1. In addition, iterons 1 and 10 are the only iterons to have identical 6-bp sequences adjacent to the left end of the iteron (Fig. 7). DNA-binding studies with purified TrfA protein have demonstrated unequivocally that the 6-bp nonrandom flanking sequences are essential for TrfA binding to the iterons and, along with the 17-bp iteron, constitute the TrfA-binding unit (43). Iteron 1 has been shown by footprint analysis to bind TrfA (44). Cloned iteron 10 expresses strong incompatibility with a mini-RK2 plasmid (data not shown), indicating that it too is a functional TrfA-binding site. TrfA has been shown to have the greatest affinity for iterons with the sequence Py-Pu-Py immediately adjacent to the left end of the iteron sequence (43). Because both iterons 1 and 10 contain the sequence CAT at this location (Fig. 7), they are likely to be high-affinity TrfA-binding units.

Our finding of a second orphan iteron flanking the origin region leads us to propose that TrfA protein binds to iterons 1 and 10 and promotes the formation of a DNA loop that encompasses the remaining eight iterons known to be involved in replication initiation and control. There is strong genetic and molecular evidence for the interaction of TrfA proteins bound to iterons on different plasmid molecules (27). Intermolecular coupling has been proposed as a mechanism for RK2 replication control, in which the linking of plasmids via *oriV* regions bound by TrfA effectively blocks initiation of replication and controls copy number (27). TrfA proteins stably bound to iterons 1 and 10 of the *oriV* region on a single plasmid may therefore be expected to interact similarly, and the result of such an intramolecular interaction would be the formation of a DNA loop. It is common for regions involved in looping to exhibit intrinsic bends or to contain binding sites for DNA-bending proteins (20). We have found that the essential five-iteron cluster of the *oriV* region is situated between two sets of sequences predicted to have an intrinsic bend (Fig. 7) (14, 30, 59). While we have no direct evidence for bending, it is possible that these sequences facilitate looping.

The involvement of a replication initiator in the formation of DNA loops containing an origin of replication has been demonstrated for other plasmids. In plasmid P1, RepA-mediated looping of the replication origin to a second set of iterons is thought to provide an additional level of copy number regulation beyond intermolecular coupling (8, 41). In plasmid R6K, the arrangement of the single and partial iterons of the α and β origins of replication, respectively, and the multiple iterons of the γ origin bear striking resemblance to the RK2 *oriV* region, except that the predicted loops are larger in R6K (35). Studies have indicated that the R6K π initiator protein mediates α-γ and β-γ interactions that lead to the initiation of replication (35, 36, 38). Thus, looping of DNA may be important to the function and regulation of a replication origin. In RK2, the orphan iterons 1 and 10 are not essential for replication (57). However, the formation of a loop in the *oriV* region might serve to alter the structure of the DNA to allow more stringent regulation of copy number by intermolecular

coupling or to enhance the efficiency of *oriV* utilization in *E. coli* and diverse hosts.

Significance. Having completed the nucleotide sequence of *kilC*, we now know that the region between *oriV* and the *korA* operon on RK2 specifies 12 genes in four *kor*-regulated operons that constitute the *kilC*, *kilE*, and *kilA* determinants (Fig. 8) (17, 29). This cluster of genes is not essential for plasmid replication or conjugal transfer. Nevertheless, the genes of the *kil* loci are coregulated with replication and transfer functions as part of the *kil-kor* regulon (Fig. 8). This is a unique feature of the IncPα broad-host-range plasmids, and we have suggested that it is likely to have significance for the success of these plasmids in nature (15–17, 29). While natural plasmids are stably maintained by their hosts, remarkably little is understood about the interaction of plasmid and host functions. The host-lethal phenotypes exhibited by the cloned *kil* loci in the absence of *kor* functions (15, 29) indicate that several RK2 proteins expressed by the *kilC-kilE-kilA* region interact with vital components of the host cells. An understanding of these interactions should provide new insights into the relationship of a plasmid with its host.

ACKNOWLEDGMENTS

We thank Yun Bai for assistance with the genetic analysis, Chun-Huey Chang for help with the nucleotide sequencing in the early stages of this work, and Robert Pohlman for the synthesis of several sequencing primers. We are grateful to John Anderson for assistance in the identification of potential bending sequences in the *oriV-kilC* region.

This study was supported by NIH grants R01-GM29085 and R01-GM26863 to D.H.F. and Cancer Center support grant CA13696 to Columbia University. M.H.L. was partially supported by NIH training grants CA09503 and AI07161.

REFERENCES

1. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1989. Current protocols in molecular biology. J. Wiley & Sons, New York.
2. **Ayres, E. K., S. Saadi, H. C. Schreiner, V. J. Thomson, and D. H. Figurski.** 1991. Differentiation of lethal and nonlethal, *kor*-regulated functions in the *kilB* region of broad host-range plasmid RK2. Plasmid **25:**53–63.
3. **Ayres, E. K., V. J. Thomson, G. Merino, D. Balderes, and D. H. Figurski.** 1993. Precise deletions in large prokaryotic genomes by vector-mediated excision (VEX): the *trfA* gene of promiscuous plasmid RK2 is essential for replication in several gram-negative hosts. J. Mol. Biol. **230:**174–185.
4. **Barr, P., R. Thayer, R. Najarian, F. Seela, P. Laybourn, and D. Tolan.** 1986. 7-Deaza-2′-guanosine triphosphate: enhanced resolution in M13 dideoxy sequencing. BioTechniques **4:**428–432.
5. **Bechhofer, D. H., and D. H. Figurski.** 1983. Map location and nucleotide sequence of *korA*, a key regulatory gene of promiscuous plasmid RK2. Nucleic Acids Res. **11:**7453–7469.
6. **Belogurov, A. A., E. P. Delver, and O. V. Rodzevich.** 1992. IncN plasmid pKM101 and IncI1 plasmid ColIb-P9 encode homologous antirestriction proteins in their leading regions. J. Bacteriol. **174:**5079–5085.
7. **Belogurov, A. A., E. P. Delver, and O. V. Rodzevich.** 1993. Plasmid pKM101 encodes two nonhomologous antirestriction proteins (ArdA and ArdB) whose expression is controlled by homologous regulatory sequences. J. Bacteriol. **175:**4843–4850.
8. **Chattoraj, D. K., R. J. Mason, and S. H. Wickner.** 1988. Mini-P1 plasmid replication: the autoregulation-sequestration paradox. Cell **52:**551–557.
9. **Cohen, S. N., A. C. Y. Chang, and L. Hsu.** 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. Proc. Natl. Acad. Sci. USA **69:**2110–2114.
10. **Datta, N., R. W. Hedges, E. J. Shaw, R. B. Sykes, and R. H. Richmond.** 1971. Properties of an R-factor from *Pseudomonas aeruginosa*. J. Bacteriol. **108:**1244–1249.
11. **Delver, E. P., V. U. Kotova, G. B. Zavilgelsky, and A. A. Belogurov.** 1991. Nucleotide sequence of the gene (*ard*) encoding the antirestriction protein of plasmid ColIb-P9. J. Bacteriol. **173:**5887–5892.
12. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–399.
13. **Dunn, J. J., M. Elzinga, K.-K. Mark, and F. W. Studier.** 1981. Amino acid sequence of the gene 0.3 protein of bacteriophage T7 and nucleotide sequence of its mRNA. J. Biol. Chem. **256:**2579–2585.
14. **Eckdahl, T. T., and J. N. Anderson.** 1990. Conserved DNA structures in origins of replication. Nucleic Acids Res. **18:**1609–1612.
15. **Figurski, D. H., R. F. Pohlman, D. H. Bechhofer, A. S. Prince, and C. A. Kelton.** 1982. Broad host range plasmid RK2 encodes multiple *kil* genes potentially lethal to *Escherichia coli* host cells. Proc. Natl. Acad. Sci. USA **79:**1935–1939.
16. **Figurski, D. H., C. Young, H. C. Schreiner, R. F. Pohlman, D. H. Bechhofer, A. S. Prince, and T. F. D'Amico.** 1985. Genetic interactions of broad host-range plasmid RK2: evidence for a complex replication regulon, p. 227–241. *In* D. R. Helinski, S. N. Cohen, D. B. Clewell, D. A. Jackson and A. Hollaender (ed.), Plasmids in bacteria. Plenum Publishing Corporation, New York.
17. **Goncharoff, P., S. Saadi, C.-H. Chang, L. H. Saltman, and D. H. Figurski.** 1991. Structural, molecular, and genetic analysis of the *kilA* operon of broad-host-range plasmid RK2. J. Bacteriol. **173:**3463–3477.
18. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166:**557–580.
19. **Harley, C. B., and R. P. Reynolds.** 1987. Analysis of *E. coli* promoter sequences. Nucleic Acids Res. **15:**2343–2361.
20. **Harrington, R. E.** 1992. DNA curving and bending in protein-DNA recognition. Mol. Microbiol. **6:**2549–2555.
21. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *E. coli* promoter DNA sequences. Nucleic Acids Res. **11:**2237–2255.
22. **Hedges, R. W., and M. Matthew.** 1979. Acquisition by *Escherichia coli* of plasmid-borne β-lactamases normally confined to *Pseudomonas* spp. Plasmid **2:**269–278.
23. **Hershfield, V., H. W. Boyer, C. Yanofsky, M. A. Lovett, and D. R. Helinski.** 1974. Plasmid ColE1 as a molecular vehicle for cloning and amplification of DNA. Proc. Natl. Acad. Sci. USA **71:**3455–3459.
24. **Ingram, L. C., M. H. Richmond, and R. B. Sykes.** 1973. Molecular characterization of the R factors implicated in the carbenicillin resistance of a sequence of *Pseudomonas aeruginosa* strains isolated from burns. Antimicrob. Agents Chemother. **3:**279–288.
25. **Jagura-Burdzy, G., F. Khanim, C. A. Smith, and C. M. Thomas.** 1992. Crosstalk between plasmid vegetative replication and conjugative transfer: repression of the *trfA* operon by *trbA* of broad host range plasmid RK2. Nucleic Acids Res. **20:**3939–3944.
26. **Jovanovic, O. S., and D. H. Figurski.** Unpublished results.
27. **Kittell, B. L., and D. R. Helinski.** 1993. Plasmid incompatibility and replication control, p. 223–242. *In* D. B. Clewell (ed.), Bacterial conjugation. Plenum Press, New York.
28. **Kornacki, J. A., R. S. Burlage, and D. H. Figurski.** 1990. The *kil-kor* regulon of broad-host-range plasmid RK2: nucleotide sequence, polypeptide product, and expression of regulatory gene *korC*. J. Bacteriol. **172:**3040–3050.
29. **Kornacki, J. A., C.-H. Chang, and D. H. Figurski.** 1993. The *kil-kor* regulon of promiscuous plasmid RK2: structure, products, and regulation of two operons that constitute the *kilE* locus. J. Bacteriol. **175:**5078–5090.
30. **Larsen, M. H., J. N. Anderson, and D. H. Figurski.** Unpublished results.
31. **Larsen, M. H., and D. H. Figurski.** Unpublished results.
32. **Lessl, M., D. Balzer, R. Lurz, V. L. Waters, D. Guiney, and E. Lanka.** 1992. Dissection of IncP conjugative plasmid transfer: definition of the transfer region Tra2 by mobilization of the Tra1 region in *trans*. J. Bacteriol. **174:**2493–2500.
33. **Marck, C.** 1988. 'DNA Strider': a 'C' program for the fast analysis

of DNA and protein sequences on the Apple Macintosh family of computers. Nucleic Acids Res. **16:**1829–1836.

34. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labelled DNA with base-specific chemical cleavages. Methods Enzymol. **65:**499–560.
35. **McEachern, M. J., M. Filutowicz, S. Yang, A. Greener, P. Mukhopadhyay, and D. R. Helinski.** 1986. Elements involved in the copy number regulation of the antibiotic resistance plasmid R6K. Banbury Rep. **24:**195–208.
36. **Miron, A., S. Mukherjee, and D. Bastia.** 1992. Activation of distant replication origins in vivo by DNA looping as revealed by a novel mutant form of an initiator protein defective in cooperativity at a distance. EMBO J. **11:**1205–1216.
37. **Motallebi-Veshareh, M., D. Balzer, E. Lanka, G. Jagura-Burdzy, and C. M. Thomas.** 1992. Conjugative transfer functions of broad-host-range plasmid RK2 are coregulated with vegetative replication. Mol. Microbiol. **6:**907–920.
38. **Mukherjee, S., H. Erickson, and D. Bastia.** 1988. Enhancer-origin interaction in plasmid R6K involves a DNA loop mediated by initiator protein. Cell **52:**375–383.
39. **Mulligan, M. E., D. K. Hawley, R. Entriken, and W. R. McClure.** 1984. *Escherichia coli* promoter sequences predict *in vitro* RNA polymerase selectivity. Nucleic Acids Res. **12:**789–800.
40. **Olsen, R. H., and P. Shipley.** 1973. Host range and properties of the *Pseudomonas aeruginosa* R factor R1822. J. Bacteriol. **113:**772–780.
41. **Pal, S. K., and D. K. Chattoraj.** 1988. P1 plasmid replication: initiator sequestration is inadequate to explain control by initiator-binding sites. J. Bacteriol. **170:**3554–3560.
42. **Pansegrau, W., E. Lanka, P. T. Barth, D. H. Figurski, D. G. Guiney, D. Haas, D. R. Helinski, H. Schwab, V. A. Stanisich, and C. M. Thomas.** 1994. Complete nucleotide sequence of Birmingham IncPα plasmids: compilation and comparative analysis. J. Mol. Biol. **239:**623–663.
43. **Perri, S., and D. R. Helinski.** 1993. DNA sequence requirements for interaction of the RK2 replication initiation protein with plasmid origin repeats. J. Biol. Chem. **268:**3662–3669.
44. **Pinkney, M., R. Diaz, E. Lanka, and C. M. Thomas.** 1988. Replication of mini RK2 plasmid in extracts of *Escherichia coli* requires plasmid-encoded protein TrfA and host-encoded proteins DnaA, B, G DNA gyrase and DNA polymerase III. J. Mol. Biol. **203:**927–938.
45. **Prentki, P., and H. Krisch.** 1984. *In vitro* insertional mutagenesis with a selectable DNA fragment. Gene **29:**303–313.
46. **Read, T. D., A. T. Thomas, and B. M. Wilkins.** 1992. Evasion of type I and type II DNA restriction systems by IncI1 plasmid ColIb-P9 during transfer by bacterial conjugation. Mol. Microbiol. **6:**1933–1941.
47. **Saiki, R. K., D. H. Gelfand, S. Stoffel, S. I. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich.** 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239:**487–491.
48. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
49. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
50. **Sherratt, D.** 1989. Tn*3* and related transposable elements: site-specific recombination and transposition, p. 163–184. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
51. **Shine, J., and L. Dalgarno.** 1974. The 3′ terminal sequence of *E. coli* 16S ribosomal RNA: complementarity to nonsense triplet and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71:**1343–1346.
52. **Shine, J., and L. Dalgarno.** 1975. Determination of cistron specificity in bacterial ribosomes. Nature (London) **254:**30–34.
53. **Stalker, D. M., C. M. Thomas, and D. R. Helinski.** 1981. Nucleotide sequence of the region of the origin of replication of the broad host range plasmid RK2. Mol. Gen. Genet. **181:**8–12.
54. **Studier, F. W., and B. A. Moffatt.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189:**113–130.
55. **Thomas, C. M., and D. R. Helinski.** 1989. Vegetative replication and stable inheritance of IncP plasmids, p. 1–25. *In* C. M. Thomas (ed.), Promiscuous plasmids of Gram-negative bacteria. Academic Press Ltd., San Diego, Calif.
56. **Thomas, C. M., J. P. Ibbotson, N. Y. Wang, C. A. Smith, R. Tipping, and N. M. Loader.** 1988. Gene regulation on broad host range plasmid RK2: identification of three novel operons whose transcription is repressed by both KorA and KorC. Nucleic Acids Res. **16:**5345–5359.
57. **Thomas, C. M., D. M. Stalker, and D. R. Helinski.** 1981. Replication and incompatibility properties of segments of the origin region of replication of the broad host range plasmid RK2. Mol. Gen. Genet. **181:**1–7.
58. **Thomson, V. J., O. S. Jovanovic, R. F. Pohlman, C.-H. Chang, and D. H. Figurski.** 1993. Structure, function, and regulation of the *kilB* locus of promiscuous plasmid RK2. J. Bacteriol. **175:**2423–2435.
59. **VanWye, J. D., E. C. Bonson, and J. N. Anderson.** 1991. Species-specific patterns of DNA bending and sequence. Nucleic Acids Res. **19:**5253–5261.
60. **Villarroel, R., R. W. Hedges, R. Maenhaut, J. Leemans, G. Engler, M. Van Montagu, and J. Schell.** 1983. Heteroduplex analysis of P-plasmid evolution: the role of insertion and deletion of transposable elements. Mol. Gen. Genet. **189:**390–399.
61. **Waters, S. H., P. Rogowsky, J. Grinsted, J. Altenbuchner, and R. Schmitt.** 1983. The tetracycline resistance determinants of RP1 and Tn*1721*: nucleotide sequence analysis. Nucleic Acids Res. **11:**6089–6105.
62. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.
63. **York, D., and M. Filutowicz.** 1993. Autoregulation-deficient mutant of the plasmid R6K-encoded π protein distinguishes between palindromic and non-palindromic binding sites. J. Biol. Chem. **268:**21854–21861.
64. **Young, C., D. H. Bechhofer, and D. H. Figurski.** 1984. Gene regulation in plasmid RK2: positive control by *korA* in the expression of *korC*. J. Bacteriol. **157:**247–252.
65. **Young, C., R. S. Burlage, and D. H. Figurski.** 1987. Control of the *kilA* gene of the broad-host-range plasmid RK2: involvement of *korA*, *korB*, and a new gene, *korE*. J. Bacteriol. **169:**1315–1320.