JOURNAL OF BACTERIOLOGY, Apr. 1995, p. 1994–2001 Vol. 177, No. 8
0021-9193/95/$04.00+0
Copyright © 1995, American Society for Microbiology

DNA-Binding Domain of the RepE Initiator Protein of Mini-F Plasmid: Involvement of the Carboxyl-Terminal Region

FUJIHIKO MATSUNAGA,[1] YASUO KAWASAKI,[1]† MASAMICHI ISHIAI,[1] KEN NISHIKAWA,[2] TAKASHI YURA,[1]‡ AND CHIEKO WADA[1]*

Institute for Virus Research, Kyoto University, Kyoto 606-01,[1] and Protein Engineering Research Institute, Furuedai, Suita, Osaka 565,[2] Japan

Received 10 October 1994/Accepted 6 February 1995

The RepE initiator protein (251 residues) is essential for mini-F replication in *Escherichia coli* and exhibits two major functions: initiation of DNA replication from *ori2* and autogenous repression of *repE* transcription. Whereas the initiation is mediated by RepE monomers that bind to the *ori2* iterons (direct repeats), the autogenous repression is mediated by dimers that bind to the *repE* operator, which contains an inverted repeat sequence related to the iterons. We now report that the binding of RepE to these DNA sites is primarily determined by the C-terminal region of this protein. The mutant RepE proteins lacking either the N-terminal 33 (or more) residues or the C-terminal 7 (or more) residues were first shown to be defective in binding to both the *ori2* and the operator DNAs. However, direct screening and analysis of mutant RepEs which are specifically affected in binding to the *ori2* iterons revealed that the mutations (mostly amino acid substitutions) occur exclusively in the C-terminal region (residues 168 to 242). These mutant proteins exhibited reduced binding to *ori2* and no detectable binding to the operator. Thus, whereas truncation of either end of RepE can destroy the DNA-binding activities, the C-terminal region appears to represent a primary DNA-binding domain of RepE for both *ori2* and the operator. Analogous DNA-binding domains seem to be conserved among the initiator proteins of certain related plasmids.

The mini-F plasmid, like the parental F factor, is stably maintained in *Escherichia coli* with one or two copies per host chromosome. Replication of mini-F requires several host factors, including DnaA (14, 22, 31), HU (34, 38), and a subset of heat shock proteins (DnaK, DnaJ, and GrpE) (6, 18), besides the plasmid-encoded replication initiator protein, RepE (251 residues, 29 kDa) (see reference 21). A minimal mini-F consists of an origin of replication (*ori2*), *repE*, and a drug resistance gene such as *bla* and exhibits high copy numbers (10 to 15 copies per chromosome) (19) due to the lack of *incC*, which negatively modulates the copy number. The *ori2* region contains two DnaA boxes, an AT-rich region, and four direct repeats (iterons) of 19 bp, whereas the *repE* operator contains an inverted repeat whose half sequence (10 bp) resembles the *ori2* iterons (8-bp matches) (Fig. 1). The RepE protein, a sequence-specific DNA-binding protein, binds to the *ori2* iterons and to the operator (27, 36). The binding to these separate but related DNA sequences plays a key role in the regulation of mini-F replication: RepE acts as an initiator of DNA replication through binding to *ori2* and as an autogenous repressor of *repE* transcription through binding to the operator (see references 21, 30, and 37).

We recently reported that the two functions of RepE are carried out by structurally distinct forms of the protein. One of the RepE mutants (RepE54), previously selected for their ability to replicate in the *dnaJ* mutant host (16), produced hyperactive RepE that cannot form dimers (17), unlike the wild-type protein that is found mostly as dimers. The RepE54 monomers bind to *ori2* with a very high efficiency (ca. 500 times that of wild-type RepE) but hardly bind to the operator (17), consistent with the strikingly increased initiator activity and much reduced repressor activity observed in vivo (16). Furthermore, treatment of wild-type RepE dimers with protein denaturants markedly enhanced the *ori2* binding and reduced the operator binding, and the *ori2* binding was uniquely associated with monomers (17). These interesting findings and the lack of detailed information on the functional structures of RepE prompted us to determine the region(s) involved in the DNA binding. Whereas the initial studies of deletion mutants gave only limited information, direct screening and analysis of mutant RepEs specifically defective in iteron binding led us to find the critical involvement of the carboxyl-terminal region of about 75 amino acids. This region, which appears to be conserved among the related plasmids, probably represents a DNA-binding domain of RepE involved in binding to both the *ori2* iterons and the *repE* operator.

MATERIALS AND METHODS

Bacterial strains, phage, and plasmids. The *E. coli* strains, phage, and plasmids used are listed in Table 1. Strain KY1461 was constructed by introducing *thy*+ to HI2017 (15) by P1 transduction. Phage λ carrying the *araB*$_p$-DR$_2$-*lacZ* fusion gene was constructed by insertion of a pair of inversely oriented *ori2* iterons (DR) into the *Hin*dIII site of pMS434 and transfer of the resulting fusion into λpF13 by in vivo recombination (15). Plasmid pKV7202, producing RepE with a cluster of six histidines attached to the N-terminal end (His$_6$-RepE), was constructed by amplification of *repE* by PCR with pKV7190 and modified primers to create a *Bam*HI or *Hin*dIII site at the 5′ or 3′ end, respectively, and insertion of it into an expression vector, pQE9 (cut with *Bam*HI and *Hin*dIII) (see Fig. 2 and its legend). pKV7203 was constructed by PCR amplification of the *Eco*RI-*Sma*I segment of *repE* from pKV7202, deleting the histidine cluster, and subsequent substitution of it for the corresponding segment of pKV7202 (Fig. 2). The structures of pKV7202 and pKV7203 were confirmed by nucleotide sequencing. Synthesis of His$_6$-RepE or native RepE with these plasmids was induced by isopropyl-β-D-thiogalactopyranoside (IPTG); cells contained a compatible plasmid, pRep4, carrying *lacI*, whose product repressed *repE* transcription. pKV7204, like pKV718, was a defective plasmid carrying *ori2* whose replication depended on RepE provided in *trans* and was constructed by deletion of

* Corresponding author. Mailing address: Institute for Virus Research, Kyoto University, Shogoin-Kawaracho, Sakyo-ku, Kyoto 606-01, Japan. Phone: 81-75-751-4033. Fax: 81-75-751-3989.

† Present address: Institute for Microbial Diseases, Osaka University, Suita, Osaka 565, Japan.

‡ Present address: HSP Research Institute, Kyoto 600, Japan.

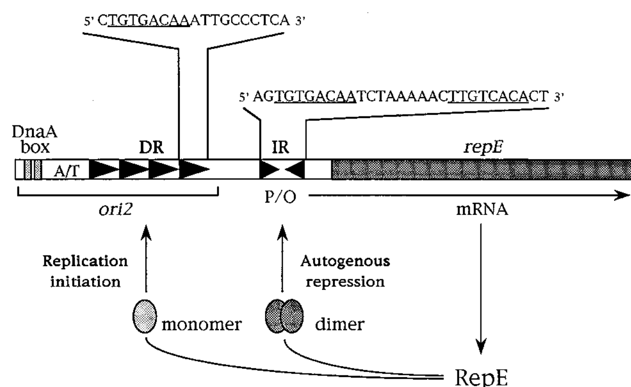

FIG. 1. The functions of the RepE initiator protein in the minimal mini-F plasmid. The RepE monomers bind to the four iterons (direct repeats [DRs]) of *ori2* to initiate replication, whereas the RepE dimers bind to the inverted repeat (IR) of the *repE* promoter-operator (P/O) to repress *repE* transcription. Parts of the repeated sequences are shown, and portions that are shared by the direct and inverted repeats are underlined.

the *Sma*I-*Eco*RV segment of *repE* from a mini-F plasmid carrying *cat* (unpublished result).

Construction of *repE* deletion mutants. Mutants producing RepEs that lack the N-terminal 33, 68, or 103 residues (ΔN33, ΔN68, or ΔN103) were constructed from pKV7202 by excision of a DNA segment flanked by two appropriate restriction sites. The mutants lacking the C-terminal end (602, 701, ΔC10, ΔC57, and ΔC148) were derived from pKV7190. RepE602 is a frameshift mutant described previously (16), and it lacks the C-terminal 6 residues and contains additional Ile and Pro. RepE701 and ΔC57 were constructed by insertion of an *Xba*I linker (CTCTAGAG) into the *Eco*RV or the *Pvu*II site, respectively. RepE701 lacks the C-terminal 7 residues but contains an additional Leu, whereas ΔC57 lacks the C-terminal 57 residues. ΔC10 was constructed by PCR amplification with a primer containing an amber codon, resulting in the loss of 10 terminal residues, and ΔC148 was obtained by excision of a segment flanked by *Nae*I and *Hpa*I sites.

Purification of RepE proteins. Wild-type RepE was purified as described previously (20, 23). Truncated RepEs (except for RepE602) were purified as His-tagged proteins with an Ni^{2+}-nitrilotriacetic acid (NTA) resin (QIA*express* system; QIAGEN). Strain M15, harboring pKV7202 (or its derivative) and pRep4, was grown at 30°C to log phase in 100 ml of L broth containing antibiotics. After induction with 1 mM IPTG for 4 h, cells were harvested and lysed for 1 h in 5 ml of buffer A (0.01 M Tris, 0.1 M NaH_2PO_4, 6 M guanidine hydrochloride, pH adjusted to 8.0) containing 10 mM β-mercaptoethanol and 1 mM phenylmethylsulfonyl fluoride (PMSF). Crude extract was loaded onto a 0.5-ml Ni^{2+}-NTA column, washed with 8 mM imidazole in buffer A, and eluted with 40 mM imidazole. The eluates containing RepE were dialyzed against 20 mM morpholineethanesulfonic acid (MES)-KOH (pH 6.0)–500 mM KCl–0.1 mM EDTA–10 mM β-mercaptoethanol–10% glycerol for 5 h and stored at −20°C with 50% glycerol. Purification was done at room temperature, and 0.5 to 1 mg of RepE (ca. 90% purity) was obtained, as determined by analysis with sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and staining with Coomassie brilliant blue.

Gel retardation assay for DNA-binding activities. The assay conditions with both purified RepE and crude extract were essentially as described elsewhere (20). The reaction mixture (20 μl) contained 20 mM Tris-HCl (pH 7.5), 40 mM NaCl, 40 mM KCl, 10 mM $MgCl_2$, 0.1 mM EDTA, 1 mM dithiothreitol, 0.1 mg of bovine serum albumin per ml, 10 μg of poly(dI-dC) per ml, 5 fmol of probe DNA end labeled with [γ-^{32}P]ATP, and RepE (or crude extract) added last. The *ori2* fragment (130 bp) containing iterons (but not DnaA boxes and an AT-rich region) or the operator fragment (180 bp) containing an inverted repeat was used as a DNA probe. The mixture was incubated at 30°C for 30 min, 2 μl of dye solution was added, and the mixture was electrophoresed with a 10% polyacrylamide gel. Gels were dried and analyzed with a Fujix bioimaging analyzer (BAS2000; Fuji, Tokyo, Japan).

Crude extracts were prepared from log-phase cells of KY1462 harboring pKV7203 (or its derivative) and pRep4, grown, and treated with 0.1 mM IPTG to induce RepE for 1.5 h in 5 ml of L broth at 30°C. Cells were collected and disrupted by sonication in 1 ml of buffer I (20 mM HEPES [*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid; pH 8.0], 100 mM NaCl, 2 mM EDTA, 7 mM β-mercaptoethanol, 1 mM PMSF, 10% glycerol [see reference 39]) and centrifuged (14,000 × *g* for 15 min), and the final supernatant was stored at −20°C. The approximate concentration of RepE protein in each extract was determined by immunoblotting with purified RepE as a reference (other proteins in crude extract little affected the band intensities of RepE; data not shown); all extracts contained RepE at a range of 0.09 to 0.18 mg/ml. Assays with crude extracts yielded only RepE-specific bands, and the results were comparable to those with purified RepEs with respect to both the band shift patterns and the apparent efficiencies of binding to either DNA site.

PCR mutagenesis of *repE*. Base substitution mutations of *repE* were obtained by PCR amplification (25) with *Taq* DNA polymerase (Wako Chemical, Osaka, Japan). The reaction mixture (100 μl) contained 10 mM Tris-HCl (pH 8.8), 50 mM KCl, 1.5 mM $MgCl_2$, 0.1 mM $MnCl_2$, 0.1% Triton X-100, 200 μM (each) dATP, dTTP, dGTP, and dCTP, and 4 U of *Taq* DNA polymerase. DNA amplification was carried out by 25 cycles of treatment at 94°C for 30 s, 50°C for 1 min, and 70°C for 2 min. The DNA fragments obtained were digested with appropriate restriction enzymes and ligated with the parental plasmid (pKV7203) cut with the same enzymes.

Other methods. Immunoblotting was done as described previously (19), with antiserum made against purified RepE protein. DNA manipulations, SDS-PAGE, and nucleotide sequencing were performed by the standard procedures (35).

RESULTS

Deletion of the N- or C-terminal end of RepE affects DNA binding. Although putative helix-turn-helix DNA-binding motifs are found in the internal (residues 63 to 82) (28) and C-terminal (residues 221 to 242) regions of RepE (3), their

TABLE 1. Bacterial strains, phage, and plasmids

Strain, phage, or plasmid	Relevant genotype	Reference or source
Strains		
KY1461	F^- Δ(*ara-leu*)*7697* Δ(*lac-pro*) *thi trpA38*(Oc)	This work
KY1462	KY1461 (λφ*araB$_p$*-DR$_2$-*lacZ*) pRep4	This work
M15	F^- *lac ara gal mtl*	2
Phage		
λpF13	*imm*^{21} φ*araB$_p$*-*lacZ*	15
λpF13-*repE$_p$*-*lacZ*	*imm*^{21} φ*repE$_p$*-*lacZ*	19
λ*araB$_p$*-DR$_2$-*lacZ*	*imm*^{21} φ*araB$_p$*-DR$_2$-*lacZ*	This work
Plasmids		
pKV7190	*ori*(pMB1) *trpR cat* φ*trp$_p$*-*repE*	19
pMS434	*ori*(ColE1) *bla* φ*araB$_p$*-*lacZ*	15
pQE9	*ori*(pMB1) *bla*	QIAGEN
pRep4	*ori*(p15A) *neo lacI*	QIAGEN
pKV7202	pQE9 φ(His$_6$-*repE*)	This work
pKV7203	pQE9 *repE*	This work
pKV718	*ori2* (mini-F) *bla*	19
pKV7204	*ori2* (mini-F) *cat*	This work

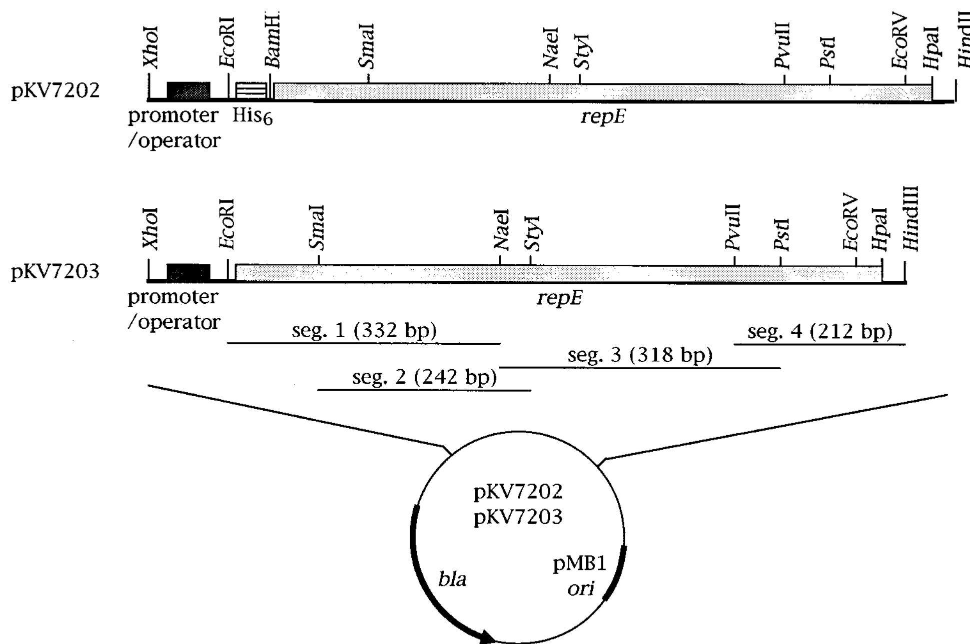

FIG. 2. Structure and partial restriction maps of *repE* expression plasmids. Plasmids pKV7202 and pKV7203 (derived from pQE9) consist of the pMB1 origin, *repE*, and *bla*. Whereas pKV7202 produces the N-terminal histidine-tagged RepE (His_6-RepE), pKV7203 produces native RepE. A modified phage T5 promoter (read by *E. coli* RNA polymerase) and two tandemly arranged *lac* operators control the *repE* transcription. The histidine-tagged RepE contained the following 16 residues at its N-terminal end: (M)RGSHHHHHHGSIEGR, fused at the EGR end to the second codon of RepE. The four (overlapping) segments (1 to 4) of *repE* used for PCR mutagenesis and isolation of DNA binding-defective mutants (Table 2) are indicated below.

possible involvement in DNA binding had not been exploited. We constructed a series of N- and C-terminal deletion derivatives of RepE and tested the initiator and repressor activities in vivo and the DNA-binding activities with purified proteins in vitro. To facilitate protein purification, RepE with a histidine cluster attached to the N-terminal end was constructed. The resulting His_6-RepE retained both initiator and repressor activities in vivo and DNA-binding activities in vitro that were essentially identical to those of RepE without added histidines (data not shown), although possible effects on the deletion mutants have not been excluded.

When the N-terminal 33, 68, or 103 residues were removed by deletion, the truncated proteins obtained virtually lacked the initiator and repressor activities in vivo and DNA-binding activities both for the *ori2* iterons and for the operator (Fig. 3). On the other hand, RepE lacking the N-terminal 17 residues was previously shown to be fully active in DNA binding though inactive in the initiation function (23). Provided that these results can be directly compared with each other in spite of differences in construction and in other conditions employed, a region spanning at least residues 17 to 33 appeared to be involved in DNA binding directly or indirectly.

Deletion of the C-terminal 6 residues (RepE602) affected both the activities in vivo and the DNA-binding activities in vitro only slightly, in agreement with our previous results (16). Further extension of deletion to 10 or more residues (ΔC10, ΔC57, and ΔC148) markedly affected the activities in vivo and the binding to both DNAs in vitro. When the C-terminal 7 residues were deleted (RepE701), the in vivo activities were lost almost completely (the protein level was comparable to that of the wild type) but the DNA-binding activities in vitro were little affected. That this apparent discrepancy is due to unstable binding of RepE701 to DNA was shown by competition experiments. The RepE701-iteron complex was competed for by excess unlabeled DNA at much higher rates than was the wild-type RepE-iteron complex, whereas the RepE701-operator complex was competed for only at a slightly higher rate than was the wild-type control (Fig. 4). The efficiency of binding of RepE701 was significantly reduced particularly in the binding to the operator. These results taken together indicated that a region very close to the C-terminal end of RepE is critically involved in binding to the DNAs. It seemed evident, however, that such analyses of deletion mutants alone cannot provide information sufficient to confine a region or regions specifically involved in binding to either *ori2* or the operator DNA.

Direct screening for mutant RepEs with reduced DNA-binding activities. We then attempted to isolate amino acid substitution mutants that are specifically affected in DNA binding by directly screening for those defective in binding to the *ori2* iterons. Accordingly, a system in which the DNA-binding activity of RepE could be monitored by *lacZ* expression was devised: *ori2* iterons were placed in between the *araB* promoter and *lacZ* on λ prophage, and RepE protein was provided from pKV7203 in *trans* (Fig. 5). Under these conditions, RepE is expected to interfere with *lacZ* expression only when it binds to the iterons. Thus, the level of *lacZ* expression was adjusted by choice of an appropriate iteron arrangement and arabinose concentration: two sets of inversely oriented iterons were used, because this proved effective in discrimination between white colonies that produce wild-type RepE and red colonies that produce a DNA-binding-defective RepE (ΔC57 or ΔC10) on MacConkey lactose agar containing 0.05% arabinose.

PCR mutagenesis was employed, because random base substitutions were expected at high frequencies in the presence of Mn^{2+}, which reduces the fidelity of *Taq* DNA polymerase in a concentration-dependent manner (25). The four mutually overlapping segments of *repE* (Fig. 2) were separately ampli-

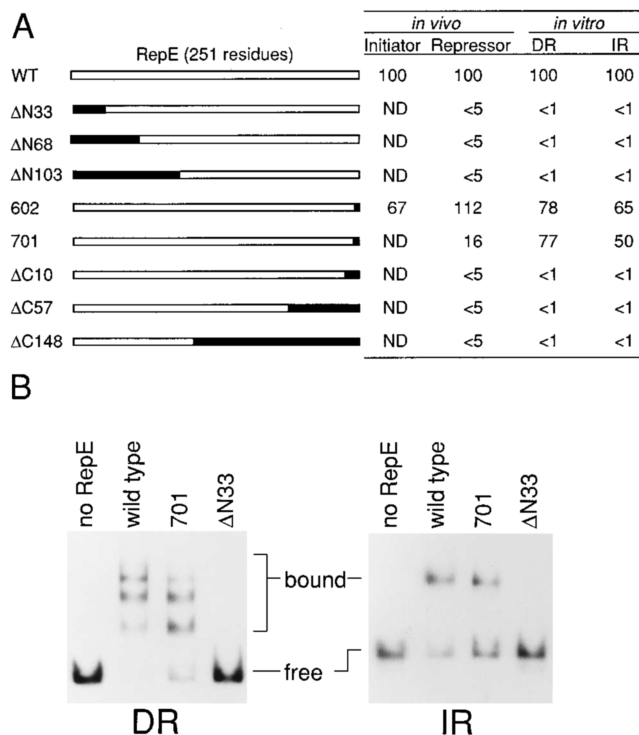

FIG. 3. (A) The positions and activities of N- and C-terminal deletion derivatives of RepE. The initiator activity was determined by measurement of the copy number of an *ori2* plasmid (pKV718 or pKV7204) which can replicate only when RepE is supplied in *trans* (19). The repressor activity was determined by measurement of β-galactosidase activity expressed from a *repE*$_p$-*lacZ* operon fusion whose transcription can be repressed by RepE supplied in *trans* (19). The in vivo activities were determined with strains that produced RepE either with (ΔN33, ΔN68, ΔN103, and ΔC148) or without (602, 701, ΔC10, and ΔC57) the histidine tags attached to the N-terminal end. DNA-binding activities were examined with mutant RepEs (except for RepE602) that were purified as histidine-tagged proteins in gel retardation assays (see panel B); the concentrations of RepE required for 50% binding (20) were calculated and presented as percentages of that required for the wild type (WT). Open and closed bars indicate the presence or absence of the RepE segment, respectively. ND, not detected (no transformants of an *ori2* plasmid were obtained). (B) DNA binding of purified His$_6$-RepEs. Examples of autoradiographs of DNA bands in a gel retardation assay with wild-type or mutant RepEs (701 and ΔN33) are shown. A 3-pmol sample of RepE was used under the conditions described in Materials and Methods. DR, direct repeat; IR, inverted repeat.

fied under the conditions described in Materials and Methods. The ends of the fragments obtained were digested with appropriate restriction enzymes and were ligated with the parental plasmid (pKV7203) cut with the same enzymes. The resulting plasmids were used to transform strain KY1462, and red colonies that appeared after 18 h at 30°C were picked from 1,000 transformants for each segment: 15, 8, 36, and 74 colonies were obtained for segments 1, 2, 3, and 4, respectively.

Mutations affecting DNA binding are localized in a C-terminal region. The clones obtained as described above containing potential *repE* mutations were first analyzed by immunoblotting with antiserum raised against RepE to obtain rough estimates of the amount and size of RepE produced. Many clones produced little or no RepE or produced truncated RepE and were not studied further. The remaining 28 clones mostly contained mutagenized segments from C-terminal ends: 0, 1, 5, and 22 clones were obtained for segments 1, 2, 3, and 4, respectively. These mutant clones were examined for nucleotide sequence alterations by dideoxy methods and for DNA-binding activities in vitro with crude extracts. As expected, most of the clones tested produced RepEs that were deficient in binding to the *ori2* iterons and to the *repE* operator, as illustrated by the representative results (Fig. 6).

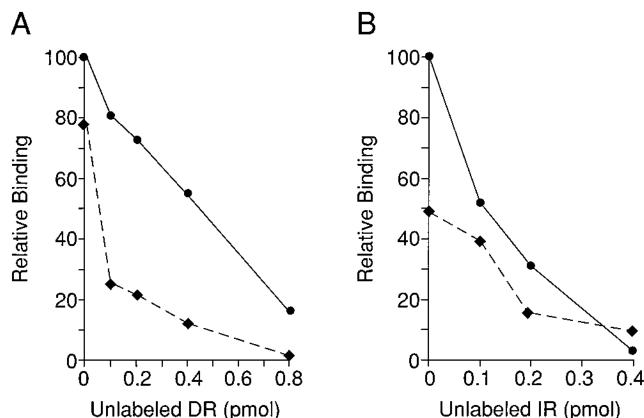

FIG. 4. Competition for binding with purified wild-type RepE or RepE701 protein by excess unlabeled DNA. The reaction conditions were as described for Fig. 3B, except that various amounts of unlabeled direct repeat (DR) or inverted repeat (IR) DNA were added simultaneously with the RepE protein. The ratios of radioactive RepE-bound DR (A) or IR DNA (B) to total DNA (*y* axes) were analyzed with a bioimaging analyzer and normalized to that of the wild type without competitor. ●, wild-type RepE; ◆, RepE701.

All amino acid changes (deduced from altered nucleotide sequences) that were associated with deficiencies in DNA binding are summarized in Table 2. These changes were found within the C-terminal region (segments 3 and 4); the majority (at least 10 independent mutants) contained a single amino acid change, four mutants contained simultaneous changes in two amino acids, and one (frameshift) mutant contained alterations in 29 consecutive amino acids and lost 4 terminal residues. Most significantly, all the missense mutations causing an amino acid change(s) were strictly confined to, but scattered within, a region spanning residues 168 to 242, close to the C-terminal end (Fig. 7).

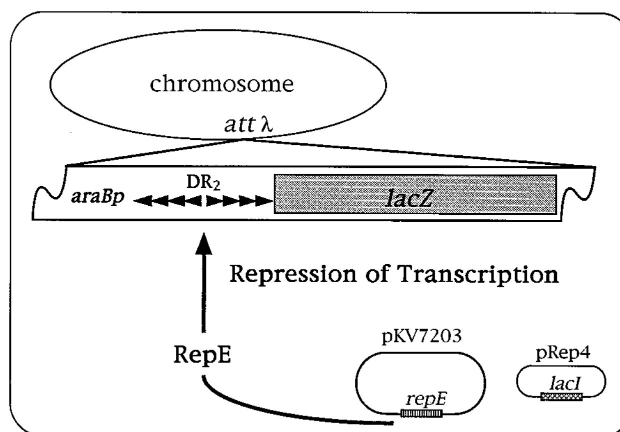

FIG. 5. The experimental system used for screening RepE mutants defective in DNA binding. A sample of pKV7203 containing a mutagenized *repE* segment was used to transform strain KY1462, which is lysogenic for λpF13 (φ*araB*$_p$-DR$_2$-*lacZ*) and harbors pRep4 (producing the *lac* repressor). The *repE* transcription on pKV7203, when induced by IPTG or lactose, provides RepE protein that can bind to the *ori2* iterons and interferes with *lacZ* transcription from the *araB* promoter in the presence of a limited amount of arabinose. In contrast, a mutant RepE that cannot bind to the iterons is expected not to interfere with transcription under the same conditions.

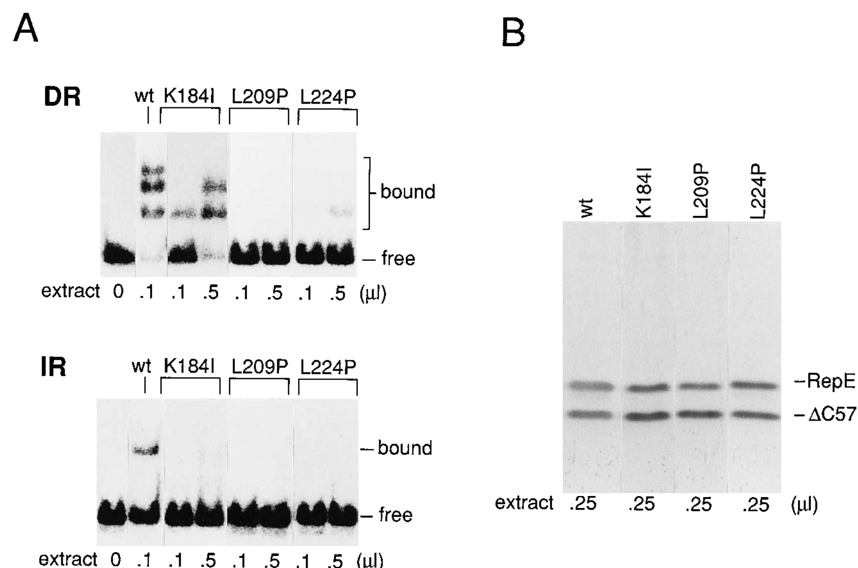

FIG. 6. (A) Representative results of gel retardation assays with crude extracts with direct repeat (DR) or inverted repeat (IR) DNA as a probe. Experiments were done as described in Materials and Methods with the indicated amounts of each extract. (B) Immunoblotting of crude extracts for RepE. Each lane contained 0.25 μl of wild-type (wt) or mutant extract and 31.6 ng of purified RepEΔC57 as a reference; the RepE concentrations were approximately 0.13, 0.10, 0.16, and 0.18 mg/ml of extract for wt, K1841, L209P, and L224P, respectively. The positions of RepE in crude extracts and purified ΔC57 (Fig. 3A) are indicated.

As to the DNA-binding activities, many mutant RepEs exhibited much reduced but significant binding to *ori2* iterons but no detectable binding to the operator (Table 2), indicating that the mutations affected the operator binding more severely than the iteron binding. The question of whether this was due to subsidiary effects of mutations on dimerization which would specifically affect the operator binding remains to be investigated. On the other hand, some of the single amino acid changes (S to P at position 168 [S168P], I188N, and L209P) caused apparently complete loss of binding to both DNAs, suggesting that these latter residues are among those that are most critical for DNA binding. The mutations that are not listed in Table 2 included seven recurrent deletions lacking residues 192 to 197 (spanning the *Pvu*II site) that are totally defective in DNA binding. On the basis of these results and those from the deletion analyses (Fig. 3), we conclude that the

TABLE 2. DNA-binding-defective RepE mutants with amino acid alterations

Mutant RepE[a]	Amino acid (codon) change	Segment no.[b]	DNA-binding activity (%)[c]	
			DR	IR
S168P	S(UCC)→P(CCC)	3	<1	<1
K1841	K(AAA)→I(AUA)	3	21	<1
I188N	I(AUC)→N(AAC)	3	<1	<1
Q193L	Q(CAG)→L(CUG)	4	21	<1
M201T[d]	M(AUG)→T(ACG)			
L194P	L(CUG)→P(CCG)	3	5	<1
L194P	L(CUG)→P(CCG)	4	1	<1
S197R	S(AGU)→R(AGA)	4	7	<1
M201T	M(AUG)→T(ACG)	4	<1	<1
S225P[d]	S(UCA)→P(CCA)			
R207P	R(CGC)→P(CCC)	4	<1	<1
F208L[d]	F(UUC)→L(CUC)			
L209P	L(CUG)→P(CCG)	4	<1	<1
N217D	N(AAC)→D(GAC)	4	5	<1
T236S[d]	T(ACU)→S(UCU)			
L224P	L(CUC)→P(CCC)	4	3	<1
F240S	F(UUU)→S(UCU)	4	3	<1
F242S	F(UUC)→S(UCC)	4	9	<1
FS219–247[e]	Frameshift (219–247)	4	<1	<1

[a] Mutant RepEs are designated by a codon number (1 for AUG) flanked by the one-letter amino acid codes of the wild type (left) and the mutant (right).
[b] The numbers have the same references as those shown in Fig. 2.
[c] The DNA-binding activity was determined with several dilutions of crude extract and corrected for the amount of RepE in the extract used (determined by immunoblotting). Values are presented as percentages of the activity of the wild-type RepE. Averages of two experiments are shown. DR, direct repeat; IR, inverted repeat.
[d] Found simultaneously with another mutant (shown immediately above) in the same clone; the DNA-binding activities of the double mutants are presented.
[e] The predicted amino acid sequences of the C-terminal ends (from residue 219) are as follows: wild type, . . . RTPMRLSYIEKKKGRQTTHIVFSFRDITSMTTG (251 residues); mutant, . . . KLQCASHTLRKRKAARRLISYFPSAISLP (247 residues).

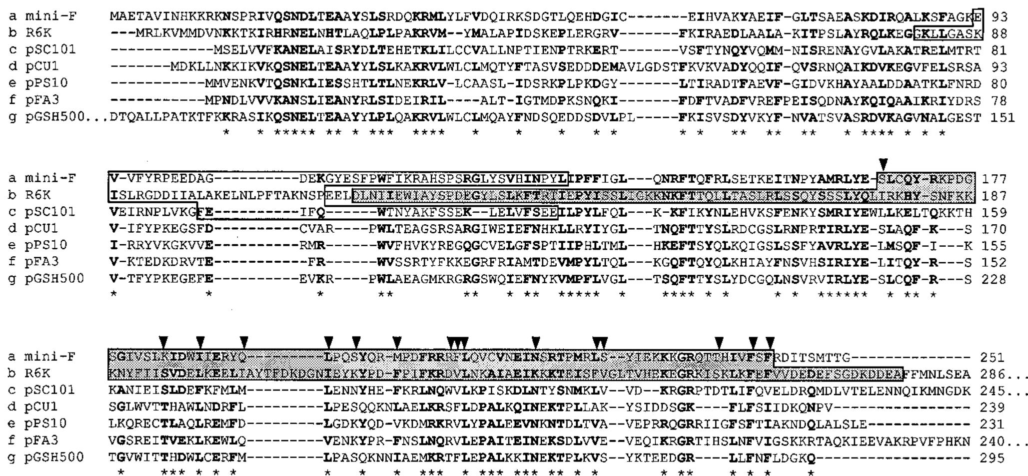

FIG. 7. Amino acid sequence alignment of RepE with other initiator proteins. Six initiator proteins of bacterial plasmids that seem to be closely related to RepE were aligned. Asterisks indicate where identical or similar amino acids (show in boldface type) were found in at least five of the seven proteins shown. Shaded and open boxes indicate the regions that are supposed to be involved in DNA binding and copy number control, respectively. Arrowheads indicate positions of missense *repE* mutations affecting DNA binding (Table 2). The RepA proteins of P1 plasmid and Rts1 plasmid were not included because they showed relatively little sequence similarity with those listed here. a, RepE (mini-F, *E. coli* [32]); b, π (R6K, *E. coli* [8]); c, RepA (pSC101, *E. coli* [1]); d, ORF239 (pCU1, *E. coli* [24]); e, RepA (pPS10, *Pseudomonas syringae* [33]); f, 39K basic protein (pFA3, *Neisseria gonorrhoeae* [11]); g, RepB (pGSH500, *Klebsiella pneumoniae* [4]). Note that terminal sequences of some of the proteins are not shown.

C-terminal region of RepE spanning residues 168 to 242 is of primary importance in binding not only to *ori2* iterons but also to the *repE* operator, although possible involvement of another region(s) was not excluded.

DISCUSSION

The present finding that the C-terminal region (residues 168 to 242) of RepE is critical for its binding both to the *ori2* iterons and to the *repE* operator suggests strongly that this region is primarily involved in binding to the 8-bp sequence (TGTGACAA) shared by these DNA sites. An analogous DNA sequence shared by direct and inverted repeats had previously been shown to be involved in binding of initiator protein π of plasmid R6K (9, 29). On the other hand, the overall structures of the two DNA sites are dissimilar and permit binding of distinct forms of RepE: the *ori2* iterons of four tandem repeats bind monomers, whereas the palindromic operator of *repE* binds dimers (17). Furthermore, binding of RepE monomers to *ori2* introduces a local bending followed by an unwinding of the adjacent AT-rich region (17a), whereas binding of dimers to the operator interferes with binding or functioning of RNA polymerase and represses *repE* transcription. Thus, the portion of 19-bp iterons unique to *ori2* (not shared by the operator), which is well conserved among the origins of related plasmids (7), may be important specifically for binding or a subsequent function(s) of RepE monomers in initiating DNA replication.

Although a putative helix-turn-helix motif near the C-terminal end of RepE (residues 222 to 243 [3]) was previously noted, it was hardly significant when an improved detection program (5) was used. Furthermore, the missense mutations affecting DNA binding were scattered in the C-terminal region (residues 168 to 242) rather than localized between residues 222 and 243 (Table 2 and Fig. 7). Besides, a computer search revealed no evidence of homology in the C-terminal region (residues 168 to 242) with other known DNA-binding motifs. Thus, the DNA-binding domain of RepE disclosed here might exhibit a feature(s) not previously found with other DNA-binding proteins. On the other hand, a more typical helix-turn-helix motif is found in the internal region (residues 64 to 83) of RepE (28). Indeed, substitution of a single amino acid (A68R), highly conserved among the known helix-turn-helix motifs, was found to markedly reduce the initiator and repressor functions in vivo and the DNA-binding activities in vitro (unpublished results). It is possible that this mutation affected DNA binding indirectly, because the helix-turn-helix region, like the neighboring central region (residues 93 to 135; see below) (16), may be involved in interaction with the C-terminal region in exerting the initiator and repressor functions. Alternatively, this region could directly participate in the DNA binding. The present analysis of deletion mutants as well as the previous results of Kline et al. (23) suggested that the N-terminal region spanning from 17 to 33 residues cannot be deleted without loss of DNA-binding activities (Fig. 3B and reference 23). This region corresponds to residues that are highly conserved among several initiator proteins (Fig. 7, residues 18 to 39 of RepE) and overlaps with a putative leucine zipper motif (residues 21 to 55) discussed below. However, no amino acid substitution mutants defective in DNA binding could be isolated from the N-terminal region in this study (Table 2). In addition, the N-terminal one-third of the R6K initiator protein, which is structurally similar to RepE (Fig. 7; see below), is dispensable for DNA binding (10). Thus, the involvement of this region in DNA binding may be indirect, through participation, for example, in stable maintenance of higher-order structures. Further studies are required to settle this point.

The pSC101 initiator protein (RepA, 316 residues), like mini-F RepE, has been reported to bind to directly repeated sequences of its origin monomeric form and to an inverted repeat of its operator in a dimeric form (26). Moreover, a deletion of the C-terminal 105 (but not 89) residues markedly affects the activities of DNA binding to both target sequences (26). This agrees very well with the present findings with RepE (Fig. 3), because the C-terminal truncation of 89 or 105 residues for RepA may be equivalent to that of 4 or 20 residues, respectively, for RepE, according to the sequence alignment shown in Fig. 7. Although the mode of plasmid R6K replica-

tion may be more complex, the dual function of the π initiator protein (305 residues) is similar though not identical to that of RepE of mini-F: π protein dimers are supposed to bind both the origin, γ (direct repeat), and the operator (inverted repeat) for initiation and repression, respectively. Again, the DNA-binding domain of π is reported to reside within the C-terminal region (residues 117 to 278) (10, 13), consistent with our findings with RepE (Fig. 7).

Among the other functional domains of RepE are those for dimerization, for interaction with heat shock proteins such as DnaK and DnaJ, and for control of plasmid copy number. It is interesting to note that the RepE54 mutation (R118P) that severely affects dimerization is located within the central region (residues 93 to 135), where a number of "copy-up" mutations due to an increase in initiator activity and/or a decrease in repressor activity are known to be localized (16, 19). On the other hand, a putative leucine zipper motif generally associated with protein-protein interaction has been found in the N-terminal region of several plasmid initiator proteins (residues 21 to 55 of RepE) (12). Elucidation of regions specifically involved in these distinct or interrelated functions requires further investigation.

Although the initiator proteins of bacterial plasmids do not seem to represent a group of proteins highly homologous in their primary structure, the sequence alignment presented in Fig. 7 revealed several features that had not been anticipated previously. First, a number of well-conserved residues can be found throughout the entire sequence. Second, the region primarily involved in copy number control is localized in the central segment of the initiator proteins of mini-F, R6K, and pSC101. At least these proteins, and possibly others, may well assume analogous higher-order structures that are specifically required for the initiator or the repressor functions. In any event, the C-terminal region of RepE identified in this study and analogous regions of other initiator proteins would play a crucial role in their binding to cognate DNAs and in subsequent functions.

ACKNOWLEDGMENTS

We are grateful to T. Nagata for helpful interest and discussion, Y. Miwa and H. Kagoshima for advice in protein purification and PCR mutagenesis, and M. Mihara for laboratory supplies.

This work was supported in part by grants from the Ministry of Education, Science and Culture, Japan.

REFERENCES

1. **Armstrong, K. A., R. Acosta, E. Lender, Y. Machida, M. Pancotto, M. McCormick, H. Ohtsubo, and E. Ohtsubo.** 1984. A 37×10^3 molecular weight plasmid-encoded protein is required for replication and copy number control in the plasmid pSC101 and its temperature-sensitive derivative pFS1. J. Mol. Biol. **175:**331–347.
2. **Beckwith, J.** 1963. Restoration of operon activity by suppressors. Biochim. Biophys. Acta **76:**162–164.
3. **Bex, F., P. Piérard, A. Desmyter, P. Drèze, M. Colet, and M. Couturier.** 1986. Mini-F E protein: the carboxy-terminal end is essential for *E* gene repression and mini-F copy number control. J. Mol. Biol. **189:**293–303.
4. **Da Silva-Tatley, F. M., and L. M. Steyn.** 1993. Characterization of a replicon of the moderately promiscuous plasmid, pGSH500, with features of both the mini-replicon of pCU1 and the *ori-2* of F. Mol. Microbiol. **7:**805–823.
5. **Dodd, I. B., and J. B. Egan.** 1990. Improved detection of helix-turn-helix DNA-binding motifs in protein sequences. Nucleic Acids Res. **18:**5019–5026.
6. **Ezaki, B., T. Ogura, H. Mori, H. Niki, and S. Hiraga.** 1989. Involvement of DnaK protein in mini-F plasmid replication: temperature-sensitive *seg* mutations are located in the *dnaK* gene. Mol. Gen. Genet. **218:**183–189.
7. **Filutowicz, M., S. Dellis, I. Levchenko, M. Urh, F. Wu, and D. York.** 1994. Regulation of replication of an iteron-containing DNA molecule. Prog. Nucleic Acid Res. Mol. Biol. **48:**239–273.
8. **Germino, J., and D. Bastia.** 1982. Primary structure of the replication initiation protein of plasmid R6K. Proc. Natl. Acad. Sci. USA **79:**5475–5479.
9. **Germino, J., and D. Bastia.** 1983. Interaction of the plasmid R6K-encoded replication initiator protein with its binding sites on DNA. Cell **34:**125–134.
10. **Germino, J., and D. Bastia.** 1984. Rapid purification of a cloned gene product by genetic fusion and site-specific proteolysis. Proc. Natl. Acad. Sci. USA **81:**4692–4696.
11. **Gilbride, K. A., and J. L. Brunton.** 1990. Identification and characterization of a new replication region in the *Neisseria gonorrhoeae* β-lactamase plasmid pFA3. J. Bacteriol. **172:**2439–2446.
12. **Giraldo, R., C. Nieto, M.-E. Fernández-Tresguerres, and R. Díaz.** 1989. Bacterial zipper. Nature (London) **342:**866.
13. **Greener, A., M. S. Filutowicz, M. J. McEachern, and D. R. Helinski.** 1990. N-terminal truncated forms of the bifunctional π initiation protein express negative activity on plasmid R6K replication. Mol. Gen. Genet. **224:**24–32.
14. **Hansen, E. B., and M. B. Yarmolinsky.** 1986. Host participation in plasmid maintenance: dependence upon *dnaA* of replicons derived from P1 and F. Proc. Natl. Acad. Sci. USA **83:**4423–4427.
15. **Hirano, M., K. Shigesada, and M. Imai.** 1987. Construction and characterization of plasmid and Lambda phage vector systems for study of transcriptional control in *Escherichia coli*. Gene **57:**89–99.
16. **Ishiai, M., C. Wada, Y. Kawasaki, and T. Yura.** 1992. Mini-F plasmid mutants able to replicate in *Escherichia coli* deficient in the DnaJ heat shock protein. J. Bacteriol. **174:**5597–5603.
17. **Ishiai, M., C. Wada, Y. Kawasaki, and T. Yura.** 1994. Replication initiator protein RepE of mini-F plasmid: functional differentiation between monomers (initiator) and dimers (autogenous repressor). Proc. Natl. Acad. Sci. USA **91:**3839–3843.

17a. **Kawasaki, Y., F. Matsunaga, and C. Wada.** Unpublished data.

18. **Kawasaki, Y., C. Wada, and T. Yura.** 1990. Roles of *Escherichia coli* heat shock proteins DnaK, DnaJ and GrpE in mini-F plasmid replication. Mol. Gen. Genet. **220:**277–282.
19. **Kawasaki, Y., C. Wada, and T. Yura.** 1991. Mini-F plasmid mutants able to replicate in the absence of σ^{32}: mutations in the *repE* coding region producing hyperactive initiator protein. J. Bacteriol. **173:**1064–1072.
20. **Kawasaki, Y., C. Wada, and T. Yura.** 1992. Binding of RepE initiator protein to mini-F DNA origin (*ori2*): enhancing effects of *repE* mutations and DnaJ heat shock protein. J. Biol. Chem. **267:**11520–11524.
21. **Kline, B. C.** 1985. A review of mini-F plasmid maintenance. Plasmid **14:**1–16.
22. **Kline, B. C., T. Kogoma, J. E. Tam, and M. S. Shields.** 1986. Requirement of the *Escherichia coli dnaA* gene product for plasmid F maintenance. J. Bacteriol. **168:**440–443.
23. **Kline, B. C., G. S. Sandhu, B. W. Eckloff, and R. A. Aleff.** 1992. Site-specific proteolysis of mini-F plasmid replication protein RepE destroys initiator function and generates an incompatibility substance. J. Bacteriol. **174:**3004–3010.
24. **Krishnan, B. R., P. R. Fobert, U. Seitzer, and V. N. Iyer.** 1990. Mutations within the replicon of the IncN plasmid pCU1 that affect its *Escherichia coli polA*-independence but not its autonomous replication ability. Gene **91:**1–7.
25. **Leung, D. W., E. Chen, and D. V. Goeddel.** 1989. A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. Technique **1:**11–15.
26. **Manen, D., L.-C. Upegui-Gonzalez, and L. Caro.** 1992. Monomers and dimers of the RepA protein in plasmid pSC101 replication: domains in RepA. Proc. Natl. Acad. Sci. USA **89:**8923–8927.
27. **Masson, L., and D. S. Ray.** 1986. Mechanism of autonomous control of the *Escherichia coli* F plasmid: different complexes of the initiator/repressor protein are bound to its operator and to an F plasmid replication origin. Nucleic Acids Res. **14:**5693–5711.
28. **Masson, L., and D. S. Ray.** 1988. Mechanism of autonomous control of the *Escherichia coli* F plasmid: purification and characterization of the *repE* gene product. Nucleic Acids Res. **16:**413–424.
29. **McEachern, M. J., M. Filutowicz, and D. R. Helinski.** 1985. Mutations in direct repeat sequences and in a conserved sequence adjacent to the repeats result in a defective replication origin in plasmid R6K. Proc. Natl. Acad. Sci. USA **82:**1480–1484.
30. **Muraiso, K., T. Tokino, T. Murotsu, and K. Matsubara.** 1987. Replication of mini-F plasmid in vitro promoted by purified E protein. Mol. Gen. Genet. **206:**519–521.
31. **Murakami, Y., H. Ohmori, T. Yura, and T. Nagata.** 1987. Requirement of the *Escherichia coli dnaA* gene function for *ori-2*-dependent mini-F plasmid replication. J. Bacteriol. **169:**1724–1730.
32. **Murotsu, T., K. Matsubara, H. Sugisaki, and M. Takanami.** 1981. Nine unique repeating sequences in a region essential for replication and incompatibility of the mini-F plasmid. Gene **15:**257–271.
33. **Nieto, C., R. Giraldo, E. Fernández-Tresguerres, and R. Díaz.** 1992. Genetic and functional analysis of the basic replicon of pPS10, a plasmid specific for *Pseudomonas* isolated from *Pseudomonas syringae* pathovar *savastanoi*. J. Mol. Biol. **223:**415–426.

34. **Ogura, T., H. Niki, Y. Kano, F. Imamoto, and S. Hiraga.** 1990. Maintenance of plasmids in HU and IHF mutants of *Escherichia coli*. Mol. Gen. Genet. **220:**197–203.
35. **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.
36. **Tokino, T., T. Murotsu, and K. Matsubara.** 1986. Purification and properties of the mini-F plasmid-encoded E protein needed for autonomous replication control of the plasmid. Proc. Natl. Acad. Sci. USA **83:**4109–4113.
37. **Wada, C., M. Imai, and T. Yura.** 1987. Host control of plasmid replication: requirement for the σ factor σ^{32} in transcription of mini-F replication initiator gene. Proc. Natl. Acad. Sci. USA **84:**8849–8853.
38. **Wada, M., K. Kohno, F. Imamoto, and Y. Kano.** 1988. Participation of *hup* gene product in *ori2*-dependent replication of fertility plasmid F. Gene **70:**393–397.
39. **Zerbib, D., M. Jakowec, P. Prentki, D. J. Galas, and M. Chandler.** 1987. Expression of proteins essential for IS*1* transposition: specific binding of InsA to the ends of IS1. EMBO J. **6:**3163–3169.