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

Luke Masson and Dan S.Ray*

Department of Biology and Molecular Biology Institute, University of California, Los Angeles, CA 90024, USA

Received 8 May 1986; Revised and Accepted 26 June 1986

ABSTRACT

E protein, the 29 kd product of the F plasmid repE gene, plays both positive and negative roles in the autoregulation of F replication. We have cloned and expressed the repE gene in an inducible ATG-fusion vector and have detected specific binding of E protein to the repE operator and to four 19-base pair direct repeats (incB) within the F plasmid replication origin ori2. Binding of E protein at the repE operator occurs with higher affinity than at ori2(incB) and gives almost complete protection to at least 30 base pairs, whereas binding of E protein to the direct repeats in the ori2 region shows an alternating pattern of enhanced and reduced sensitivity to DNAase cleavage consistent with a protein-induced folding of the DNA. These results provide direct biochemical support for a model of F plasmid replication in which the E protein serves both as an initiator of replication and as an autorepressor of its own synthesis.

INTRODUCTION

Replication of the 94.5 kb F plasmid is under stringent regulation in Escherichia coli, being present at only 1-2 copies per cell (for reviews, see refs 1 and 2). A mini-F plasmid containing only 2248 base pairs of F DNA appears identical to the parental F plasmid with respect to copy number, incompatibility, and stability (3). This plasmid possesses, in addition to an origin of replication (ori2), the repE gene which codes for a 29 kd initiator protein (E), and two incompatibility regions incB and incC. These two regions consist of nine 19-base pair (bp) direct repeat sequences present as clusters with opposite orientations located on either side of the repE gene (3,4). The four direct repeats in incB are contained within a 217 base pair sequence constituting ori2 (3). Several repE gene operator mutations have been localized to a 10-bp inverted repeat having an 8-bp sequence in common with each of the 19-bp direct repeats of incB (5).

Recently, models have been proposed in which the E protein plays a multifunctional role in F plasmid replication. Numerous studies have implicated a role for E protein as a trans-acting replication initiator

protein titratable by the 19-bp direct repeats (7,8). In addition, it is proposed that the E protein autoregulates its own synthesis by binding to the repE operator (5,6,7). However, the binding of E protein to either incB, incC or the repE operator has not yet been demonstrated. In fact, physical studies of the role of the E protein in F plasmid replication have been seriously hampered by its low intracellular concentration. The results of Komai et al., (9) show that E protein is not detectable in mini-cell extracts unless cloned on a high copy number vector next to a lac promoter. In this report, we describe the cloning and expression of the repE gene under the control of the Salmonella typhimurium araB promoter. In addition, we show the binding of E protein to both an inverted repeat sequence in the repE operator, and with a lower affinity to the incB direct repeats within ori2.

MATERIALS AND METHODS

Bacterial Strains and Plasmids

The following bacterial strains were used in this study: Escherichia coli MC1061 [araD131, Δ (ara-leu)7697, Δ lacX74, galU, galK, hsr-, hsm+, rpsL, F-] and JM109 [F', Δ lac-pro,thi-1, gyrA96,recA1, traD36]. JM109 (10) was used for both the propagation of M13 phages and for replicative form DNA production. The construction of phage and plasmids is outlined in figure 1. The pLMC3 plasmid is a derivative of pING1 containing the pifC gene of the F plasmid (unpublished results).

Enzymes

All restriction enzymes, T4 DNA ligase, and the Klenow fragment of DNA polymerase I were obtained from Bethesda Research Labs, New England Biolabs, and Boehringer Mannheim.

Buffers

TEMG buffer is 20mM Tris-HCl pH 7.5, 0.1mM EDTA, 10mM β -mercaptoethanol and 10% glycerol. TBE is 89mM Tris-base, 89mM boric acid and 2mM EDTA at pH 7.6. Agarose gel loading buffer contains 30% glycerol and 0.25% (W/V) of both xylene cyanol and bromophenol blue. Sodium dodecyl sulfate (SDS)-PAGE sample buffer consists of 62mM Tris-HCl (pH 6.8), 10% glycerol, 2% SDS, 5% β -mercaptoethanol and 0.003% bromophenol blue. Gel binding buffer is 20mM Tris-HCl (pH 7.5), 0.1 mM EDTA, 10mM β -mercaptoethanol and 25mM KCl.

Preparation of Plasmid and Phage DNA

Phage DNA was isolated as previously described (11). Restricted DNA fragments used in all cloning experiments were isolated electrophoretically

and purified from 1% agarose horizontal slab gels in TBE buffer containing 0.5 µg/ml ethidium bromide (12).

Construction of pIT2 and pREPE5 Plasmids

In order to express the E gene product in vivo, an inducible ATG-fusion expression vector (pIT2) was created (Fig. 1A). The 1444-bp EcoRI/BalI fragment containing the tetracycline resistance gene of pBR322 was substituted for the EcoRI/DraI fragment which contains the ampicillin resistance gene in the pING1 expression vector. This latter vector contains the inducible Salmonella typhimurium araB promoter, the first 450 base pairs of the araB coding sequence and the araC gene (13). In addition, the SmaI/DdeI fragment (45.35-46.01 Fkb on the 94.5 kb F map) from the repE gene (Fig. 1B) was cloned into the unique SmaI site of pING1 rendering it out of frame with the araB' sequence. Oligonucleotide-directed mutagenesis (14) was employed to create a NcoI site at the ATG start codon of the truncated araB gene (araB') generating the ATG-fusion vector pIT2. During the creation of the NcoI restriction site, the first nucleotide upstream of the ATG codon was changed from a G to a C residue. This did not appear to have an effect on the inducible araB promoter as demonstrated below. A NcoI site was also created at the ATG start codon of the repE gene using M13ΔElac-f2 which contains the PstI fragment of F plasmid (44.1-45.88 Fkb). The NcoI/PstI fragment from the resulting mutagenized clone M13ΔElac-NF2 was cloned into the large NcoI/PstI fragment of pIT2 creating the final repE gene plasmid construct, pREPE5.

Preparation of Protein Extracts

Cells (200 mls) were grown overnight at 37°C in mineral base medium (0.7% K₂HPO₄, 0.3% KH₂PO₄, 0.1% (NH₄)₂SO₄, 0.01% MgSO₄, 0.4% casamino acids (Difco) and 40 µg/ml thiamine-HCl) containing either ampicillin (100 µg/ml) or tetracycline-HCl (25 µg/ml). The cells were then inoculated into fresh medium and grown to an O.D.₅₉₅ = 0.2 and induced by the addition of L-arabinose (Cal-Biochem) to 1% (W/V) final concentration. Whole cell extracts were prepared from cultures of E. coli MC1061(pREPE5) or E. coli MC1061(pLMC3) either uninduced or after induction with 1% arabinose for four hours at 37°C in mineral base medium. Cells were harvested in TEMG buffer, containing 0.4M KCl. The cells were lysed by passage through a French pressure cell and then centrifuged for two hours at 150,000 x g. Nucleic acids were removed by passing the supernatant through a DEAE-sephacel column (10 ml bed volume) in the same buffer. The KCl concentration of the flow-through fraction was reduced to 50mM by dialysis. This

fraction was used for the gel retardation assays without further purification. Partial purification of the E protein for use in footprint studies was achieved by DEAE-sephacel chromatography. The nucleic acid-free extract was passed over a 1.5 x 15 cm column in TEMG buffer containing 50mM KCl and the bound proteins (which includes E protein) were eluted with a linear gradient of 50mM to 500mM KCl in TEMG buffer. Fractions containing E protein binding activity, as shown by the gel retardation assay (described below), eluted at approximately 400mM KCl. The pooled E protein fractions represent approximately a 100-fold purification of E protein.

Polyacrylamide Gel Retardation Assays

Binding of the E protein to DNA was detected by the polyacrylamide gel protein-DNA binding assay of Fried and Crothers, (15) as modified by Carthew et al., (16). Protein extracts were incubated with 10 ng of end-labeled, double-stranded DNA in the presence of 400 ng of poly(dI-dC).poly(dI-dC) (Pharmacia) in a final volume of 15 μ l. The binding reactions were performed in gel binding buffer at 37°C for 30 min. After the addition of 2 μ l agarose gel loading buffer, the binding mixtures were electrophoresed through low ionic strength 6% (W/V) polyacrylamide gels (acrylamide:bisacrylamide 30:1) in TBE. Following electrophoresis, the gels were dried and autoradiographed at -70°C with Quanta III intensifying screens.

DNAase I Footprinting

DNAase I footprints were obtained by the method of Galas and Schmitz, (17) with the following modifications. Partially purified E protein (DEAE-Sephacel pooled fractions) was incubated with 10 ng of the appropriate end-labeled DNA fragment, 600 ng of poly(dI-dC).poly(dI-dC) and 5mM MgCl₂ in a standard binding assay. After 25 min of incubation at 37°C, one ng of DNAase I was added and the incubation continued for an additional 10 min. The reaction was stopped by placing the mixture on ice and adjusting the volume and the EDTA concentrations to 100 μ l and 12mM respectively. The mixtures were extracted once with phenol:chloroform:isoamyl alcohol (24:24:1) followed by chloroform:isoamyl alcohol (24:1) and then ethanol precipitated. The dried precipitates were resuspended in 80% formamide in TBE and incubated at 95°C for 5 min. Samples were subsequently electrophoresed on 10% (W/V) polyacrylamide - 7M urea sequencing gels (acrylamide:bisacrylamide 20:1) in TBE. Sequence positions were determined by electrophoresis of the DNAase I cleaved fragments adjacent to A+G sequencing reactions of the same fragment (18).

In Vivo Protein Labeling

E. coli MC1061(pREPE5) and *E. coli* MC1061(pIT2) were grown overnight at 37°C in mineral base medium containing tetracycline-HCl (25 µg/ml). On the following day, the cells were diluted 1:50 into M9 medium (19) containing 0.5% fructose, 25 µg/ml tetracycline-HCl and all the common amino acids (50 µg/ml each) with the exception of methionine. The cells were then grown to an O.D.₅₉₅ = 0.2, induced with 1% arabinose and incubated for an additional two hours at 37°C. 15 µCi of [³⁵S]-methionine was incubated with 0.6 ml of cells for one min at 37°C after which 0.15 ml of cold 50% (W/V) trichloroacetic acid was added and the precipitate harvested by a brief centrifugation. The precipitate was washed twice with cold acetone, air-dried and resuspended in SDS-PAGE sample buffer, boiled for 5 min and electrophoresed through a 15% (W/V) SDS-polyacrylamide gel.

RESULTS

Fusion of the repE Gene to the Inducible araB Promoter

In order to overproduce the E protein in *E. coli*, we have created an ATG-fusion expression vector (pIT2) through the use of oligonucleotide-directed mutagenesis. The mutagenesis was designed to create an *Nco*I restriction site at the ATG translational start codon of the *araB'* gene in the high expression vector pING1. Possession of this site makes this pBR322-based plasmid attractive for the expression of non-fusion proteins from genes having either a *Nco*I or a *Afl*III restriction site at their ATG start codons. For our purposes, a *Nco*I site was created by oligonucleotide-directed mutagenesis at the *repE* gene ATG start codon thus allowing the subsequent cloning of the *repE* gene into pIT2. The resulting pREPE5 plasmid contains the complete *repE* gene in place of the *araB'* gene and is consequently under control of the inducible *araB* promoter (Fig. 1C). The tetracycline resistance gene was cloned into pIT2 to replace the β-lactamase gene since both β-lactamase and the E protein share similar molecular weights.

Expression of the repE Gene Product

After transformation of the final plasmid constructs, pIT2 and pREPE5, into *E. coli* MC1061, synthesis of either the ARA B' protein or the 29 kd E protein was detected by pulse-labeling of cells with [³⁵S]-methionine. As shown in figure 2, cells containing the pIT2 plasmid synthesized the 17 kd ARA B' protein in the presence (lane B) but not the absence (lane A) of arabinose. This demonstrates that the oligonucleotide-directed mutagenesis

employed to create the NcoI restriction site at the araB' ATG translational start site had no deleterious effects on the expression of the araB' gene. In addition, the appearance of a 34 kd protein, corresponding to the size of the tetracycline-resistance protein occurs in the presence of arabinose (lane B). This protein and a 29 kd protein are both induced in E. coli MC1061(pREPE5) cells in the presence (lane D) but not the absence (lane C) of arabinose. The identities of the 17 kd and the 29 kd proteins were verified by immunoprecipitation with antisera raised against an ARA B'-E fusion protein (data not shown). Total cellular proteins, from uninduced (lane E) or induced (lane F) E. coli MC1061(pREPE5), were also examined by SDS-PAGE followed by staining with Coomassie blue. The stained gel reveals both a 34 and a 30 kd protein band after four hours of induction by arabinose. Incubation of the bacteria for longer than four hours in the presence of arabinose did not increase the level of either protein. It should be noted here that the E protein migrates, during SDS-PAGE, at a slightly higher molecular weight (approx. 30-31 kd) when examining total cell protein. However, after partial purification, this protein migrates as a 29 kd protein (data not shown).

It is interesting to note that although the repE gene is expressed under the control of the araB regulatory sequences, the E protein is synthesized at a reduced level as compared to the ARA B' protein expressed by pIT2 or other proteins expressed from this promoter (13).

Binding of E Protein to the repE Operator and to ori2(incB)

The detection of protein-DNA interactions in our assays is based upon the slower electrophoretic mobility of end-labeled DNA fragments, in low ionic strength polyacrylamide gels, caused by the binding of protein to DNA

Figure 1. Fusion of the repE Gene to the araB Promoter. A. M13ΔE1ac-f2, which contains the 1.78 kb PstI fragment of F plasmid (44.1-45.88 Fkb) (heavy line), was used as a single-stranded template for the creation of an NcoI site at the translational start codon of the repE gene (shaded box) by oligonucleotide-directed mutagenesis. The NcoI-PstI fragment of M13ΔE1ac-Nf2 was subcloned into the agarose gel-purified, araB'-deleted large fragment of pIT2 restricted with NcoI and PstI giving rise to the final construct, pREPE5.

B. Genetic loci and restriction map of mini-F plasmid sequences. Restriction sites involved in cloning are shown with coordinates in kb from the F plasmid physical map. Small arrows correspond to the number, position, and direction of the direct repeats found in incB and incC. Large arrow shows the direction of transcription of the repE gene (dark box).

C. Nucleotide sequences of araB (small type) and F plasmid (large type) regions involved in pREPE5 construction. Arrows denote the translational start sites and direction of translation of both the araB and repE genes.

Figure 2. Arabinose-Induced Expression of ARA B' and E Proteins. *E. coli* MC1061 cells containing either the pIT2 or the pREPE5 plasmids were grown and induced by L-arabinose as described in Materials and Methods. Samples were removed either after two hours for $[^{35}\text{S}]$ -methionine labeling of nascent proteins (lanes A-D), or after four hours for direct analysis of unlabeled whole cell proteins (lanes E and F). All samples were analyzed by SDS-PAGE (15% W/V). Lanes A and B represent uninduced and induced *E. coli* MC1061(pIT2) cell proteins respectively. Lanes C and D represent uninduced and induced *E. coli* MC1061(pREPE5) cell proteins. Lanes E and F represent uninduced and induced *E. coli* MC1061(pREPE5) total cell proteins stained with Coomassie Brilliant Blue. Lane G contains Bio-Rad molecular weight size markers (phosphorylase b, M_r 92,500; bovine serum albumin, M_r 68,000; ovalbumin, M_r 45,000; carbonic anhydrase, M_r 31,000; soybean trypsin inhibitor, M_r 20,000; lysozyme, M_r 14,000).

(15,16). To discriminate between specific and non-specific DNA binding, the alternating copolymer poly(dI-dC).poly(dI-dC) was included as competitor DNA in the reaction mixtures. Cellular protein extracts were examined for their ability to bind specifically to fragments of DNA implicated in the regulation of F plasmid replication. Through the use of repE-lac translational fusions, a putative repE operator site has been

Figure 3. Specific Binding of the E Protein to DNA Containing the *repE* Operator. Reaction mixtures containing crude cell extracts were incubated with 10 ng of 3' end-labeled *AluI/AvaI* (190 bp) fragment for 30 min at 37°C and electrophoresed through a low ionic strength, 6% (W/V) polyacrylamide gel. The gel was dried and autoradiographed as described in Materials and Methods. Lane A represents no extract; lane B, induced *E. coli* MC1061 (pLMC3) extract (1.36 µg); Lane C, uninduced *E. coli* MC1061(pREPE5) extract (1.36 µg). Lanes D-I contain 0.085, 0.17, 0.34, 0.68, 1.02, and 1.36 µg of extract from induced *E. coli* MC1061(pREPE5). Lane J contains 1.36 µg of extract from induced *E. coli* MC1061(pREPE5) and 10 ng of unlabeled *AluI/AvaI* fragment. All reactions contain 500 ng of poly(dI-dC).poly-(dI-dC). A 70-bp end-labeled fragment of M13 RF DNA (lowest band) was included as a non-specific internal control. The position of the unbound *AluI/AvaI* fragment is indicated by an arrow.

localized to an inverted repeat overlapping the *repE* promoter (5,6). We have examined the binding of E protein to the 190-bp *AluI/AvaI* fragment (Fig. 1B) containing the *repE* operator as well as to the 160-bp *BglII/AluI ori2(incB)* fragment which contains the four 19-bp direct repeats. (The five direct repeats comprising the *incC* incompatibility region were not considered in this work). As shown in figure 3, extracts from either induced *E. coli* MC1061 cells possessing a modified pING1 vector (pLMC3) or uninduced *E. coli* MC1061 cells containing the pREPE5 plasmid do not bind the *AluI/AvaI* fragment (lanes B and C). However, extracts of induced cells carrying the pREPE5 plasmid can specifically retard the migration of this fragment. The relative amount of the slowest migrating band increases with

Figure 4. Specific Binding of the E Protein to DNA Containing the *ori2* (*incB*) 19-bp Direct Repeats. Reaction mixtures containing crude cell extracts were incubated with 10 ng of end-labeled *Bgl*II/*Alu*I (160 bp) fragment for 30 min at 37°C and electrophoresed through a low ionic strength, 6% (W/V) polyacrylamide gel. The gel was dried and autoradiographed as described in Materials and Methods. Lane A represents no extract; lane B, induced *E. coli* MC1061(pLMC3) extract (1.36 µg); Lane C, uninduced *E. coli* MC1061(pREPE5) extract (1.36 µg). Lanes D-I contain 0.085, 0.17, 0.34, 0.68, 1.02, and 1.36 µg of extract from induced *E. coli* MC1061(pREPE5). Lane J contains 1.36 µg of extract from induced MC1061(pREPE5) and 10 ng of unlabeled *Bgl*II/*Alu*I fragments. All reactions contain 500 ng of poly(dI-dC).poly(dI-dC). A 70-bp end-labeled fragment of M13 RF DNA (lowest band) was included as a non-specific internal control. The position of the unbound *Bgl*II/*Alu*I fragment is indicated by an arrow.

increasing protein concentration (lanes D-J). This band can be specifically competed for by the addition of unlabeled *Alu*I/*Ava*I fragment (lane J). As an internal control, an end-labeled, nonspecific 70-bp DNA fragment was included (bottom band). When the *Bgl*II/*Alu*I fragment containing the four direct repeats is substituted for the *Alu*I/*Ava*I fragment, specific binding is also observed (Fig. 4). One major difference between the two fragments is the appearance of multiple retarded bands when the direct repeat fragment is used. When the protein concentration is increased even further than shown in figure 4, two additional bands begin to appear which migrate even slower than the two shifted bands shown in this figure (data not shown). These results are consistent with the

presence of multiple binding sites. A comparison between figures 3 and 4 indicates that the E protein has a higher affinity for the repE operator than for the ori2(incB) fragment. The appearance of a slower migrating species first occurs at a protein level of 0.17 μ g for the operator fragment (Fig. 3, lane E) as compared to 0.68 μ g for the direct repeat fragment (Fig. 4, lane G).

DNAase I Footprinting

To specifically localize E protein binding sites, E protein-DNA complexes were allowed to form during the standard binding reaction using partially purified E protein (DEAE-Sephacel fraction, see Materials and Methods) and were then subjected to a brief DNAase I treatment followed by polyacrylamide gel electrophoresis under denaturing conditions. The repE operator fragment, 3' end-labeled at the AvaI site, displayed a large DNAase I protected region of at least 30 nucleotides in length with a single hypersensitive C residue (Fig. 5) within one of the 10 base-pair inverted repeat sequences of the repE operator. Binding of E protein to this site is consistent with the results of Rokeach *et al.*, (5) who mapped repE operator mutations to this inverted repeat segment.

The footprint of the BglII/SmaI fragment, containing both the incB direct repeats and the repE operator inverted repeat, revealed the protection of a 5-bp sequence 5'-TGTC A-3'. This sequence is found within the 8-bp sequence 5'-TTGTCACA-3' common to both the inverted repeat of the repE operator and to each of the 19-bp direct repeats of incB (Figs. 6 and 7). However, there are two separate 5'-TGTC A-3' sequences located between direct repeats 2 and 3 as well as near the ATG start codon of the repE gene (Fig. 7), which remain sensitive to DNAase I cleavage. Since both of these 5-bp sequences do not form part of the larger 8-bp sequences found within each of the direct and inverted repeats, this would suggest that adjacent nucleotides, although sensitive to DNAase I cleavage, assist in E protein-DNA interactions. Protection of specific sites within the incB repeats requires substantially higher levels of E protein than for protection of the repE operator. In addition to the protected 5'-TGTC A-3' sequence within the incB repeats, protection of non-homologous sequences three to four nucleotides long are also observed at the 5' boundary of each of the incB repeats 1,2, and 3 and at a site located between repeats 2 and 3. The spacing between protected sites varies between 9 and 13 base pairs indicating that the protected sites are separated by about a turn of the helix (10.5 base pairs) on the average. Enhanced cleavage at some sites is

also observed interspersed between the protected sites. An especially strong enhancement of the DNAase I cleavage of several nucleotides occurs within repeat number 4 (Fig. 6).

The higher affinity of E protein for the repE operator is apparent in both the gel retardation assays and the footprint of the 350-bp BglII/SmaI fragment. It should be noted that the protection of the 19-bp repeats occurred at the same concentration of E protein regardless of whether or not the repE operator was present on the same fragment (data not shown).

DISCUSSION

Binding of a replication initiator protein to repeated sequences in or near an origin of replication is a common feature in the regulation of the replication of many plasmids. For example, replication of the plasmid R6K depends on a plasmid-encoded initiator protein (20,21) which binds to tandem repeats of a 22-bp sequence. Binding to these direct repeats is essential for replication. In addition, the R6K initiator protein autoregulates its synthesis at the transcriptional level by binding to its own promoter (22,23). Similar regulatory mechanisms appear to be utilized by the plasmids pSC101 (24,25), P1 (26) and Rts1 (27).

Control of F plasmid replication is of considerable interest because of its stringent regulation at a very low copy number of approximately one per chromosomal equivalent (28). Control of F plasmid replication has been shown to be negative (29,30) and to involve directly repeated sequences (3). It has recently been shown that the repE gene is essential for F replication (4,8,31) and that its synthesis is likely autoregulated (6,7). Trawick and Kline, (7) have proposed a model of F regulation involving both a titration of E protein by binding to multiple repeated sites and autoregulation of transcription of the repE gene through binding of E protein to the repE operator. Different forms of the E protein are hypothesized to be involved in these two interactions. Titration of one

Figure 5. DNAase I Footprint Analysis of the repE Operator Fragment. Lane A contains an A+G Maxam and Gilbert sequencing ladder (18) of the AluI/AvaI fragment 3' end-labeled at the AvaI site. Bands representing individual nucleotides in the A+G ladder are shifted down one nucleotide relative to the corresponding DNAase I cleaved fragments. The AluI/AvaI fragment was subjected to DNAase I digestion in the absence of added E protein (Lane B) or after incubation with 300 ng of partially purified E protein (DEAE-Sephacel fraction) Lane C. Arrows mark the position of the 10-bp inverted repeats and the solid lines denote the boundaries of the DNAase I protected region. The dashed line marks the hypersensitive base. The bottom of the gel represents the 3' end of the fragment.

form is proposed to be mediated by binding to 19-bp direct repeat sequences within the incB and incC loci flanking the repE gene. An alternate form of E protein is proposed to bind to a 10-bp inverted repeat sequence identified as the operator for the repE gene (5).

Our studies of the binding of E protein to specific fragments of the F plasmid provide direct biochemical evidence for the binding of the E protein to the repE operator and to ori-2(incB). Binding of the E protein to the inverted repeat of the repE operator was found to occur at a much lower concentration than for the 19-bp direct repeats within ori2. Moreover, the patterns of DNAase I protection are entirely different at these sites in spite of the fact that they share a common eight nucleotide sequence 5'TGTCACA-3'. A stretch of at least 30 nucleotides spanning the operator region, which includes the inverted repeats, is protected from DNAase I cleavage being interrupted only by a single enhanced cleavage site at a C residue. The protection pattern of the repE operator appears asymmetric with respect to the inverted repeats with the protection extending further beyond one repeat than the other.

Examination of the protection pattern within ori2(incB) revealed that the protected bases occur in clusters of 3-5 nucleotides and appear to be of two types. One type is specific and is found within each of the direct repeats centering around the common sequence 5'-TGTC A-3'. This sequence is contained within the 8-bp sequence shared between the incB repeats and the inverted repeats. Resistance to DNAase I digestion at these sites is presumably due to the direct physical interaction of the E protein with the DNA. The other type of DNAase I protection is non-sequence specific. These protected sites also occur in clusters and are interspersed between the specific sites. Moreover, they appear to be protected as a consequence of their distance from the specific sites. Each non-specific protected cluster occurs approximately one turn of the DNA helix from a specific binding site. In addition, sites of enhanced cleavage are located about

Figure 6. DNAase I Footprint Analysis of the ori2(incB) Fragment. Lanes A and J contain an A+G Maxam and Gilbert sequencing ladder of the BglII/SmaI fragment 3' end-labeled at the BglII site. Bands representing individual nucleotides in the A+G ladder are shifted down one nucleotide relative to the corresponding DNAase I cleaved fragments. This fragment was subjected to DNAase I digestion after incubation with no E protein (lane B), 0.3 μ g (lane C), 0.6 μ g (lane D), 1.2 μ g (lane E), 2.4 μ g (lane F), 3.6 μ g (lane G), 4.8 μ g (lane H) and 8.4 μ g (lane I) of partially pure E protein (DEAE-Sephacel fraction). The nucleotide sequence of the four 19-bp direct repeats and their location as well as the location of the inverted repeats (I.R.) are shown. The bottom of the gel represents the 3' end of the fragment.

Figure 7. Summary of the DNAase I Protected and Hypersensitive Regions. Protected nucleotides are represented by (*) and hypersensitive sites by (-). Both fragments used to footprint the *repE* operator or the four 19-bp direct repeats were 3' end-labeled. The 19-bp direct repeats and the 10-bp inverted repeats are outlined (boxes) with the 5'-TGTC A-3' protected sites bracketed.

4-6 base pairs from each specific binding site. This pattern of alternating protected and enhanced cleavage sites places all of the protected sites on one side of the DNA and the enhanced cleavage sites on the opposite side of the DNA and is strikingly similar to the pattern of DNA protection observed for DNA sequences located between pairs of operators by the λ repressor protein as described by Hochschild and Ptashne, (32). They observed that the DNA between operator binding sites separated by five or six turns of the helix became alternately hypersensitive and resistant to DNAase I cleavage at half turn intervals as repressor molecules bind cooperatively to the two operator sites. Their results as well as those of Drew and Travers (33), attribute alternating clusters of enhanced and diminished DNAase I sensitivity to bending of the DNA. Sites on the inside of a bend are viewed as less accessible to DNAase I while sites on the outside of a bend are more accessible. Whether the non-specific sites of reduced sensitivity to DNAase I are due to a cooperative binding of E protein to sites adjacent to the specific binding sites or to a bending of the DNA between pairs of specific sites separated by two turns of the helix might be distinguished by an examination of E protein protection of these sequences from specific chemical modifications. It should be noted that binding of the R6K initiator protein to direct repeat sequences has been reported to induce bending at one of the R6K plasmid replication origins (34). Initiator protein-induced bending of replication origins may represent a necessary folding of the DNA to create an appropriate structure for directing initiation. An intrinsic bend in the replication origin of SV40 has recently been found to be a determinant for the binding of the large T antigen, a protein required for initiation at the SV40 origin (35).

The differences in the binding patterns and binding affinities between the *repE* operator and *ori2(incB)* likely reflect different protein-DNA

structures in these two regions. The actual physical form of E protein either in solution or during assembly on the DNA is unknown, however, some inferences may be drawn from the pattern of fragment retention in the gel retardation experiments. The appearance of greater than one slower migrating band seen during gel retardation of the incB fragment may represent intermediate structures suggesting a stepwise binding of multiple proteins. Alternatively, each band may represent the binding of a single monomer of E protein at different sites in incB giving rise to complexes of slightly different mobility. Footprint analysis of each of the retarded species should distinguish between these possibilities.

It is of interest to note that a molecule of dnaA protein, having a molecular weight of 52 kd, protects a 9-bp consensus sequence (36). Since the common sequence specifically protected by E protein is only four to five base pairs long, a monomer of the 29 kd protein would appear to be sufficient for this extent of protection. Although the inverted repeats of the repE operator contains two copies of the common sequence, the absence of intermediate bands during gel retardation analysis of the repE operator fragment and the almost complete protection of a sequence of more than 30 nucleotides suggests the binding of an oligomeric form of E protein. Alternatively, monomers of the E protein may bind to the repE operator with a very high cooperativity. Physical studies of the purified E protein and its complexes with both the repE operator and incB and incC repeats should provide additional insight into the structures of these DNA-protein complexes.

From both gel retardation experiments and DNAase I footprinting, it appears that the E protein has a significantly higher affinity for the repE operator than it does for the ori2(incB) repeats. If initiation at ori2 requires saturation of the incB binding sites by E protein then some basal level of E protein synthesis by the repressed repE gene may be necessary for the eventual triggering of initiation. Upon replication of the plasmid, the total number of binding sites would double leading to a redistribution of E protein, repression of the new copy of the repE gene, and a reduction of the amount of E protein bound at ori2 below the threshold level required for initiation. Further replication of the plasmid would not occur until the threshold level of E protein complexed at ori2 had been achieved again. Loss of the plasmid by cell division prior to achieving this threshold level would be prevented by the ccdB gene which functions to suppress cell division until after F has been replicated (37).

ACKNOWLEDGEMENTS

We are grateful to Dr. R. Imber for providing the M13ΔElac-f2 and the ARA B'-E fusion protein antiserum, and to Jane Hines for the construction of a plasmid carrying the F plasmid PstI-DdeI fragment (45.88-46.01 Fkb).

This work was supported by a grant from the National Institutes of Health (AI10752). Equipment for these studies was provided in part by a grant from the California Biotechnology Research and Education Program. Oligonucleotide synthesis was supported in part by a grant (CA16042) from the National Cancer Institute to the Jonsson Comprehensive Cancer Center. L.M. was the recipient of postdoctoral fellowship from the Medical Research Council of Canada.

*To whom correspondence should be addressed

REFERENCES

1. Lane, H.E.D. (1981) Plasmid 5, 100-126.
2. Kline, B.C. (1985) Plasmid 14, 1-16.
3. Murotsu, T., Matsubara, K., Sugisaki, H., and Takanami, M. (1981) Gene 15, 257-271.
4. Maki, S., Miki, T., and Horiuchi, T. (1984) Mol. Gen. Genet. 194, 337-339.
5. Rokeach, L.A., Sogaard-Andersen, L., and Molin, S. (1985) J. Bacteriol. 164, 1262-1270.
6. Sogaard-Andersen, L., Rokeach, L.A., and Molin, S. (1984) EMBO J. 3, 257-262.
7. Trawick, J.D., and Kline, B.C. (1985) Plasmid 13, 59-69.
8. Watson, L.A., Phua, S.H., Berquist, P.L., and Lane, H.E.D. (1982) Gene 19, 173-178.
9. Komai, N., Nishizawa, T., Hayakawa, Y., Murotsu, T., and Matsubara, K. (1982) Mol. Gen. Genet. 186, 193-203.
10. Yanisch-Perron, Veira, J., and Messing, J. (1985) Gene 33, 103-119.
11. Hines, J.C., and Ray, D.S. (1980) Gene 11, 207-218.
12. Yang, R.C.-A., Lio, J., and Wu, R. (1979) Meth. Enzymol. 68, 176-182.
13. Johnston, S., Lee, J.-H., and Ray, D.S. (1985) Gene 34, 137-146.
14. Zoller, M.J., and Smith, M. (1983) Meth. Enzymol. 100, 468-500.
15. Fried, M., and Crothers, D.M. (1981) Nucl. Acids Res. 9, 6505-6525.
16. Carthew, R.W., Chodosh, L.A., and Sharp, P.A. (1985) Cell 43, 439-448.
17. Galas, D.J., and Schmitz, A. (1978) Nucl. Acids Res. 5, 3157-3170.
18. Maxam, A., and Gilbert, W. (1980) Meth. Enzymol. 65, 499-525.
19. Miller, J.H. (ed.) (1972). In Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
20. Germino, J., and Bastia, D. (1982) Proc. Natl. Acad. Sci. USA 79, 5475-5479.
21. Stalker, D.M., Kolter, R., and Helinski, D. (1982) J. Mol. Biol. 161, 33-43.
22. Filutowicz, M., Davis, G., Greener, A., and Helinski, D. (1985) Nucl. Acids. Res. 13, 103-114.

23. Kelley, W., and Bastia, D. (1985) *Proc. Natl. Acad. Sci. USA* 82, 2574-2578.
24. Armstrong, K.A., Acosta, R., Ledner, E., Machida, Y., Pancotto, M., McCormick, M., and Ohtsubo, H. (1984) *J. Mol. Biol.* 175, 331-347.
25. Linder, P., Churchward, G., Guixian, X., Yu, Yi-Yi, and Caro, L. (1985) *J. Mol. Biol.* 181, 383-391.
26. Abeles, A.L., Snyder, K.M., and Chattoraj, D.K. (1984) *J. Mol. Biol.* 173, 307-324.
27. Kamio, Y., Tabuchi, A., Itoh, Y., Katagiri, H. and Terawaki, Y. (1984) *J. Bacteriol.* 158, 307-312.
28. Collins, J., and Pritchard, R.H. (1973) *J. Mol. Biol.* 78, 143-155.
29. Tsutsui, H., and Matsubara, K. (1981) *J. Bacteriol.* 147, 509-516.
30. Tsutsui, H., Fujima, A., Murotsu, T., and Matsubara, K. (1983) *J. Bacteriol.* 155, 337-344.
31. Tolun, A. and Helinski, D.R. (1982) *Mol. Gen. Genet.* 186, 372-377.
32. Hochschild, A., and Ptashne, M. (1986) *Cell* 44, 681-687.
33. Drew, H.R., and Travers, A.A. (1985) *J. Mol. Biol.* 186, 773-790.
34. Mukherjee, S., Patel, I., and Bastia, D. (1985) *Cell* 43, 189-197.
35. Ryder, K., Silver, S., DeLucia, A.L., Fanning, E., and Tegtmeyer, P. (1986) *Cell* 44, 719-725.
36. Fuller, R.S., Funnell, B.E., and Kornberg, A. (1984) *Cell* 38, 889-900.
37. Ogura, T., and Hiraga, S. (1983) *Cell* 32, 351-360.