JOURNAL OF BACTERIOLOGY, Dec. 1985, p. 1004–1010
0021-9193/85/121004-07$02.00/0
Copyright © 1985, American Society for Microbiology
Vol. 164, No. 3

Escherichia coli rep Gene: Identification of the Promoter and N Terminus of the Rep Protein†

HANNA BIALKOWSKA-HOBRZANSKA,[1] CAROL A. GILCHRIST,[2] AND DAVID T. DENHARDT[2]*

Cancer Research Laboratory, University of Western Ontario, London, N6A 5B7,[2] and Department of Pathology, McMaster University, Hamilton, L8N 3Z5,[1] Ontario, Canada

Received 14 August 1984/Accepted 27 August 1985

The functional *Escherichia coli rep* gene, which encodes the M_r 67,000 Rep helicase, has been localized within a 2.55-kilobase sequence. Its regulatory region has been characterized by the use of *rep-lacZ* fusions. The direction of transcription of the *rep* gene is clockwise on the *E. coli* chromosome, as are the nearby *ilvC* and *rho* genes. The sequence of the *rep* control region was determined, and putative regulatory sequences were identified; no evidence for autoregulation of expression was obtained. Transcription of the gene was not enhanced during the SOS response. The location of the promoter and the beginning of the protein were confirmed by S1 nuclease mapping of the 5′ end of *rep* mRNA and determination of the NH_2-terminal sequence of the *rep* protein.

The *rep* gene of *Escherichia coli* codes for the Rep helicase, which is required by some phages (φX174, fd, P2) for their replication (35). Mutations in the *rep* locus do not seem to be lethal for the host cells, but they do reduce the rate of replication fork progression and, to compensate, increase the number of replication forks in the replicating *E. coli* chromosome (17). The finding that *rep uvrD* double mutants have reduced viability (32) or are not viable (36) suggests that at least one of the two helicases, Rep helicase or DNA helicase II, encoded in the *uvrD* gene (21, 25) is required at the replication fork. It is not so obvious why certain combinations of *rep* and *rho* mutations are lethal (13).

It is of interest that the loss of either helicase independently results phenotypically in at least partial DNA repair deficiency. The product of the *uvrD* gene is known to be involved in excision repair of pyrimidine dimers (27), mismatch repair (24), and genetic recombination (2). The Rep helicase has been implicated in the repair of DNA damage by the somewhat increased sensitivity of *rep* mutants to cross-linking agents (6) and UV (7) and X-ray (10) irradiation. Also, a mutation affecting minimal medium recovery and post replication repair after UV irradiation has been mapped in the vicinity of the *rep* gene (31).

We have previously reported the cloning and restriction analysis of the *rep* gene (4). In the present study we have localized and characterized the *rep* regulatory region.

MATERIALS AND METHODS

Bacterial strains, media, and chemicals. The bacterial strains used in this study are listed in Table 1. The media for growth were identical to those described by Miller (23). When needed, ampicillin, chloramphenicol, or tetracycline was added at a final concentration of 50, 25, or 10 μg/ml, respectively. To score for the Lac^+ phenotype, 5-bromo-4-chloro-3-indolyl-β-D-galactoside (Sigma Chemical Co.) was added to L agar at a final concentration of 40 μg/ml; hydrolysis of 5-bromo-4-chloro-3-indolyl-β-D-galactoside by β-galactosidase dyes the colonies or plaques blue.

β-Galactosidase assay. Cells were grown at 37°C in M9 medium containing 0.2% glucose and 0.2% Casamino Acids (Difco Laboratories) to ≈5 × 10^8 cells per ml and assayed for β-galactosidase by hydrolysis of *o*-nitrophenyl-β-galactosidase by the method of Miller (23). One unit produces 1 nmol of *o*-nitrophenol per min. Cell density was determined spectrophotometrically, and the specific activity was calculated as described previously (23).

Isolation and manipulation of DNA. Isolation and purification of plasmid DNA was performed as described previously (4). Digestion of DNA with restriction endonucleases was carried out under conditions specified by the supplier. All double digests were done sequentially, allowing the restriction enzyme requiring the lower salt concentration to react first; the appropriate adjustments were made to the reaction buffer, and the second restriction enzyme was added. Restriction fragments used in subcloning experiments were isolated from low-melting-point agarose gels by electroelution (39).

BAL 31 exonuclease digestion was performed as indicated in Maniatis et al. (20), except for a 3-h dialysis of linear plasmid DNA against BAL 31 buffer before exposure to BAL 31. The kinetics of BAL 31 digestion was determined with restriction fragments of appropriate size for calibration.

Recombinant plasmids were constructed in vitro by ligation of endonuclease-generated fragments with T4 DNA ligase (P-L Biochemicals, Inc.) at 15°C in 66 mM Tris chloride (pH 7.6)–6.6 mM $MgCl_2$–20 mM dithiothreitol–1 mM ATP for 18 to 22 h; DNA concentrations suggested by Dugaiczyk et al. (11) were used. DNA transformation was performed in a pH 5.8 buffer containing 100 mM RbCl–45 mM $MnCl_2$–10 mM $CaCl_2$–5 mM $MgCl_2$–15% sucrose (20).

Isolation of RNA. Cells were grown in L broth. At 5 × 10^8 cells per ml, NaN_3 was added to a final concentration of 0.1 M, and 100-ml portions were poured over frozen, crushed buffer (0.15 M NaCl, 50 mM Tris chloride [pH 8], 50 mM EDTA). The cells were harvested and suspended in 2 ml of 10 mM Tris chloride (pH 8)–10 mM EDTA–0.1 M NaCl. An equal volume of this buffer (≈100°C) containing 1% sodium dodecyl sulfate was added, and the mixture was kept at ≈100°C for 5 min. An equal volume of phenol equilibrated with 50 mM sodium acetate at 60°C (pH 5.5) was added after the heated lysate had been cooled down to 60°C. After two phenol extractions at 60°C, the aqueous phase was extracted

* Corresponding author.

† This is the paper no. 8 in a series on the *rep* mutation. The previous paper is reference 4.

TABLE 1. Bacterial strains

Bacterial strain	Relevant genotype	Reference or source
D94	*E. coli* HF4704 Thy(Ts) UvrA(?) Sup^- $\phi X174^s$ $P2^s$ λ^s	This laboratory
D92	D94 *rep-38*(Am)	This laboratory
MC1060	Δ(*lacIPOZYA*)*X74 galU galK rpsL hsdR*	M. J. Casadaban (8)
JM103	Δ(*lac-pro*) *rpsL endA sbcB15 hsdR4 supE* F'*traD36 proAB lacI*^{q}*Z*Δ*M15*	J. Messing

with $CHCl_3$-isoamyl alcohol (24:1, vol/vol), and RNA was precipitated with 2 volumes of ethanol at −20°C. The precipitate was collected by centrifugation and suspended in H_2O treated with 0.1% diethylpyrocarbonate (20).

Northern transfers. Total RNA (30 to 60 μg per lane) was fractionated by electrophoresis on 1.1% agarose gels in 10 mM sodium phosphate buffer (pH 7.0) with constant buffer recirculation. RNA samples were denatured in 1 M glyoxal–10 mM sodium phosphate (pH 7) at 50°C for 40 min. Glyoxylated RNA was transferred to nitrocellulose in 20X SSC (1X SSC is 0.15 M NaCl plus 0.015 M sodium citrate) within ≈20 h. Blots were baked in a vacuum oven for 2 h at 80°C to fix RNA and remove glyoxal. Hybridizations were done essentially as described by Thomas (37), except for the omission of dextran sulfate in the hybridization solutions.

DNA and protein sequencing. DNA sequences were determined by the dideoxy method with M13mp8 and M13mp9 as cloning vehicles and a synthetic 17-base primer from New England Biolabs (no. 1211). JM103 was used as a recipient in all transfection experiments. Two restriction fragments, a 0.95-kilobase (kb) *Eco*RI fragment and a 2.65-kb *Nru*I-*Bal*I fragment (see Fig. 2), were subcloned from pHBH30 (4) into the *Eco*RI and *Hin*dII sites of M13mp8 and M13mp9. The orientations of the inserted fragments were verified by restriction mapping. The clones of M13mp8 and M13mp9 with the 0.95-kb *Eco*RI fragment inserted in an orientation such that the interior *Eco*RI site of *rep* was located proximal to the primer, and M13mp8 with the 2.65-kb *Nru*I-*Bal*I fragment cloned in an orientation such that the *Nru*I site was located proximal to the primer, were chosen for sequencing. These clones were designated M13*rep*816, M13*rep*914, and M13*rep*827, respectively. The sequencing reactions were performed as described previously (29), with the exception that [α-^{35}S]dATP was used as the radioactive label (5). Sequence analysis was performed with the computer program of Larson and Messing (18). The N-terminal amino acids of the Rep protein were determined in an Applied Biosystems model 470A Gas Phase Sequenator. Approximately 100 pmol of Rep protein was subjected to 10 cycles of automated Edman degradation, and the phenylthiohydantoin derivatives of the sequenced amino acids were identified.

S1 mapping of transcripts. A 0.46-kb *Bam*HI-*Eco*RI fragment from M13*rep*827 containing the 5' flanking region of the *rep* gene was purified by electroelution from a 5% polyacrylamide gel. The DNA was 5' end labeled with ^{32}P, denatured, and then annealed with 125 μg of RNA isolated from D92(pHBH30) (4) as described above. S1 digestion at 37°C was performed as described by Ridgway et al. (26). The protected DNA was electrophoresed on an 8% sequencing gel alongside the products of the Maxam-Gilbert sequencing reactions (22) of the same fragment.

RESULTS

Determination of the minimal coding region of the *rep* gene. To define more precisely the termini of the region required for expression of the *rep* gene, we introduced deletions into the 3.2-kb *Xho*I-*Bal*I segment previously shown to encode the M_r 67,000 Rep protein (4). To remove sequences around

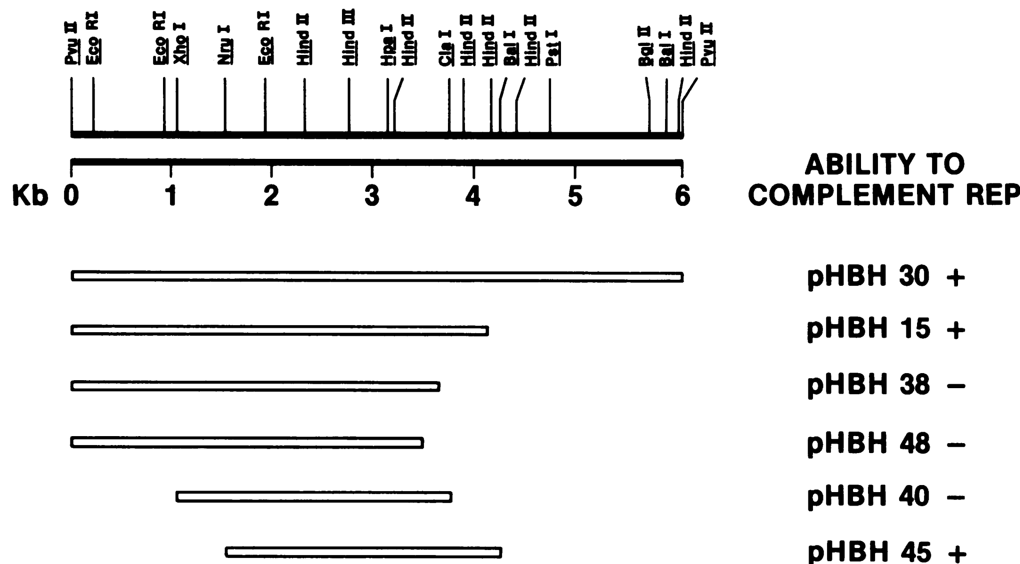

FIG. 1. Restriction endonuclease map and Rep phenotype of *rep* plasmids used in this work. The map of the 6-kb *Pvu*II fragment of pHBH30 has been previously reported (4). The *Cla*I, *Hin*dII, and *Nru*I sites were mapped in this study. The relative position of the *Hpa*I and *Hin*dII sites, separated by less than 50 bp, has not been established. The ability to complement *rep-38* was determined by the efficiency of plating of ϕX174 on Ap^r transformants. The (+) sign indicates that the efficiency of plating of ϕX174 was at least 90% of that on D94 (Rep$^+$). The (−) sign indicates that the efficiency of plating on the given strain was less that 5×10^{-9}. The open line represents the *E. coli* sequences contained within the pBR322 vector.

the *Bal*I site, the plasmid pHBH30 (4) was linearized with *Bgl*II (a unique *Bgl*II site is located 1.5 kb to the right of the *Bal*I site) and treated with sufficient BAL 31 exonuclease to remove 1.5 to 2.2 kb of DNA from each end. The resulting DNA was treated with the Klenow fragment of DNA polymerase I to increase the proportion of blunt-ended termini and ligated in the presence of a 10-base-pair (bp) *Kpn*I linker. The products were then used to transform D92 (*rep-38*).

Ampicillin-resistant (Ap^r) transformants were analyzed for both their ability to propagate φX174 phage (as a criterion for the presence of a functional *rep* gene) and the restriction properties of the resident plasmid. A deletion extending 100 bp to the left of the *Bal*I site at 4.25 kb (in pHBH15) left the *rep* function intact (Fig. 1). In contrast, plasmids with deletions extending either to the *Cla*I site (in pHBH40) or beyond the *Cla*I site to the left (in pHBH38 and pHBH48) were unable to complement *rep-38* mutations. Preliminary sequence data suggest that the *rep* gene extends about 200 bp to the right of the *Cla*I site. Thus the *Cla*I site lies just within the region required for expression of the *rep* gene.

To delineate further the 5′ boundary of *rep*, we constructed plasmid pHBH45 by joining the 2.65-kb *Nru*I-*Bal*I fragment of pHBH30 to the *Pvu*II site of pBR322. pHBH45 complemented the *rep-38* mutation (Fig. 1), indicating that

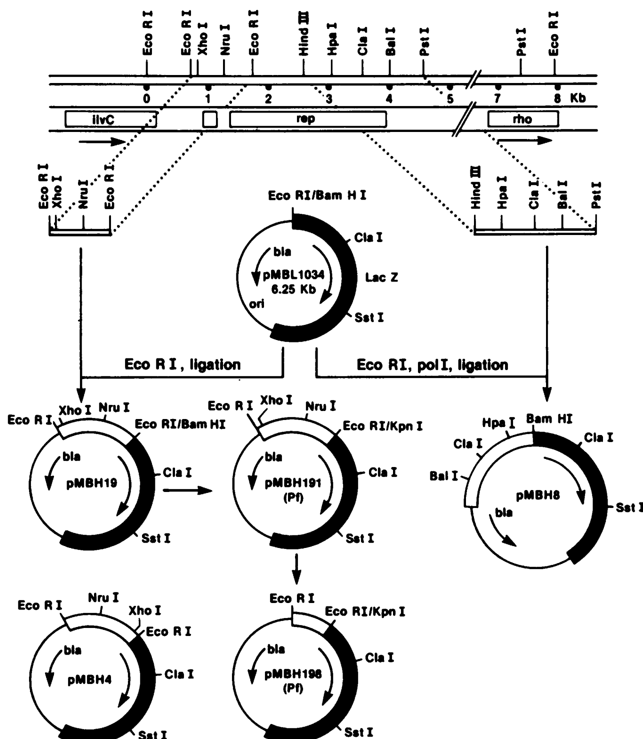

FIG. 2. Structure of *rep-lacZ* fusion pMBH plasmids. Their construction is described in the text. The top lines represent a portion of the restriction map of the *rep* gene region of the *E. coli* genome (4). The zero point is arbitrarily chosen at the approximate end of the *ilvC* gene, and the directions of transcription of *ilvC* and *rho* are indicated. Only restriction sites relevant to this work are shown. The open line denotes *E. coli* chromosomal DNA, the solid line represents *lacZ* DNA, and the thin line denotes the sequence from pBR322. The open box between *ilvC* and *rep* indicates a gene product of unknown function (Atlung, personal communication). Arrows indicate the direction of transcription. Pf stands for protein fusion.

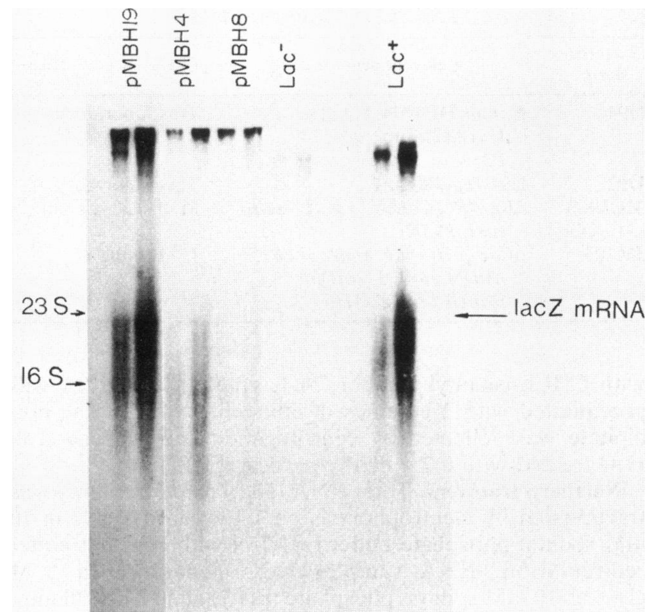

FIG. 3. Transcription of *lacZ* DNA directed by flanking regions of the *rep* gene. Total cellular RNA was extracted from D94 (Lac$^+$), MC1060 (Lac$^-$), or MC1060 carrying the indicated plasmid, electrophoresed on an agarose gel, and transferred to nitrocellulose as described in Materials and Methods. Each RNA preparation was loaded in adjoining lanes in 30 μg (left) and 60 μg (right) portions per slot. The Northern blot was hybridized to a ^{32}P-labeled *Bam*HI-*Sst*I fragment of the *lacZ* gene.

the *Nru*I site lies outside the *rep* gene boundary. These data place the *rep* boundaries within a 2.55-kb fragment of DNA with the *Nru*I (at 1.6 kb) and *Hin*dII (at 4.15 kb) sites outside the *rep* coding region.

Location of the regulatory signals of the *rep* gene. We used the *lac* fusion vector pMBL1034 (33) to locate the promoter of the *rep* gene. This vector is a derivative of pBR322 that contains a *lacZ* gene lacking the promoter, ribosome-binding site, and first eight codons. This *lacZ* allele can be regulated by promoters placed in the correct orientation and reading frame. The left part (0.95-kb *Eco*RI fragment) and the right part (2-kb *Hin*dIII-*Pst*I fragment) of the *rep* gene were fused with the *lacZ* gene. One of these fusions should result in the coupling of the *rep* message to the *lacZ* message. The position and the orientation of the 0.95-kb *Eco*RI and 2-kb *Hin*dIII-*Pst*I fragments were determined (Fig. 2) by restriction analyses. If the *rep* gene were transcribed clockwise on the standard *E. coli* map (as the *ilvC* and *rho* genes are), then the 0.95-kb *Eco*RI fragment would contain the *rep* promoter region. Alternatively, the *Hin*dIII-*Pst*I fragment would have the regulatory signals if the *rep* gene were transcribed counterclockwise.

None of the plasmids pMBH19, pMBH4, or pMBH8 was able to restore the Lac$^+$ phenotype when introduced into *E. coli* MC1060. Because this was likely the result of out-of-frame fusions of the coding sequences, we investigated which fusions were capable of directing the synthesis of the *lacZ* mRNA. RNA was isolated from strains D94 (Lac$^+$), MC1060 (Lac$^-$), MC1060(pMBH19), MC1060(pMBH8), and MC1060(pMBH4), fractionated on agarose gels, transferred onto nitrocellulose, and hybridized to a ^{32}P-labeled *Bam*HI-*Sst*I fragment of the *LacZ* gene. Analysis of the results presented in Fig. 3 suggested that the 0.95-kb *Eco*RI fragment in either orientation (pMBH4, pMBH19) promoted

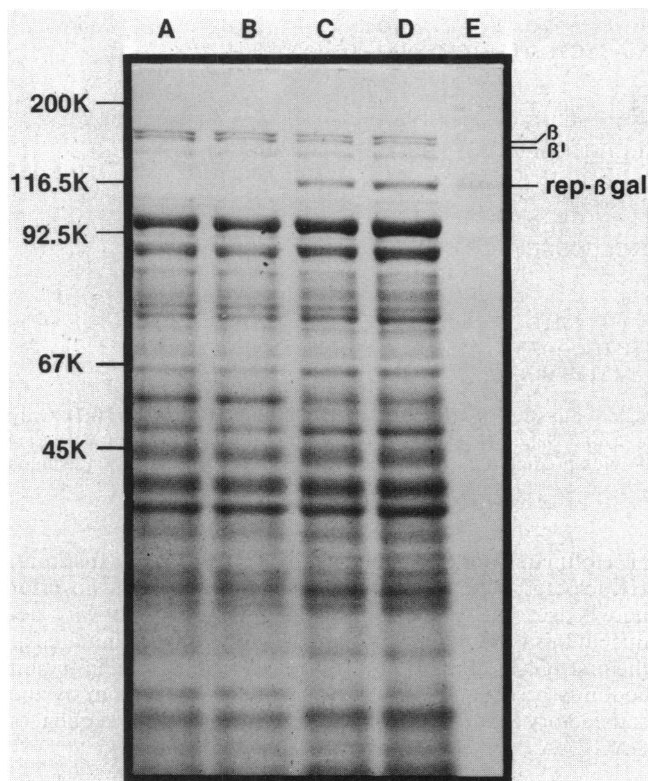

FIG. 4. Identification of the *rep-lacZ* hybrid protein. The cells were grown in M9 medium to 5×10^8 cells per ml and then centrifuged, concentrated 10-fold in cold M9 buffer, and sonicated for 2 min at 4°C (22, 32). The supernatant was treated with DNase I at a final concentration of 10 μg/ml, followed by precipitation with acetone. The protein pellet was dissolved in Laemmli buffer, boiled for 5 min, and subjected to electrophoresis on a 6% (wt/vol) sodium dodecyl sulfate-polyacrylamide gel (16). The Coomassie blue-stained gel is shown. Lanes: A, strain MC1060(pMBL1034); B, strain MC1060(pMBH19); C, strain MC1060(pMBH191); D, strain MC1060(pMBH198); E, purified β-galactosidase of M_r 116,500. The positions of Bio-Rad molecular weight standards (myosin, M_r 200,000; phosphorylase B, M_r 92,500; bovine serum albumin, M_r 67,000; ovalbumin, M_r 45,000) and the positions of the β and β' chains of RNA polymerase are also indicated.

transcription of *lacZ*. This implies that this fragment has at least two transcriptional start signals, one for the *rep* gene transcribed from the *Nru*I site toward the proximal *Eco*RI site and the other oppositely oriented towards a gene located to the left of the *Nru*I site, possibly encoding an M_r 11, 500 protein (Atlung, personal communication).

Plasmid pMBH8 harboring the right portion of the *rep* gene (*Hin*dIII-*Pst*I fragment) fused to the *lacZ* gene did not produce runoff *lacZ* mRNA; this supports the argument that the *rep* promoter is in the right part of the 0.95-kb *Eco*RI fragment. We obtained the same results in three separate Northern hybridization experiments. The hybridization seen at the top of the gel (Fig. 3) is due to the chromosomal DNA (less than 1%) present in the RNA samples. When the total RNA sample was treated with DNase, this signal disappeared (data not shown). Dot-blot hybridization data with RNA isolated by the guanidinium-CsCl method (20) to decrease DNA contamination confirmed these conclusions.

To construct *rep-lacZ* translational fusions that could produce a hybrid protein consisting of an enzymatically active segment of β-galactosidase joined to the amino terminus of the *rep* gene product, the coding region of the *rep* gene was aligned in phase with the β-galactosidase gene. pHBH19 was cleaved at the unique *Bam*HI site and digested with BAL 31 to remove up to 20 nucleotides per end. In alternative approaches, *Bam*HI cohesive ends were converted to blunt ends with the Klenow fragment of DNA polymerase I and then ligated with and without a 10-bp *Kpn*I linker. Ligated DNA was used to transform MC1060; the resulting transformants were selected on L-agar plates containing 5-bromo-4-chloro-3-indolyl-β-D-galactoside and ampicillin. All of these approaches yielded deep blue Lac$^+$ clones among the Ap^r transformants. Restriction analyses of plasmid DNAs isolated from 10 Ap^r Lac$^+$ transformants revealed that the *rep-lacZ* junction had been altered. These functional *rep-lacZ* fusions support the conclusion that the direction of transcription of *rep* is clockwise on the pMBH191 plasmid and on the standard *E. coli* map.

The hybrid protein encoded by the *rep-lacZ* fused gene would be expected to have an M_r greater than 116,500 (the size of native β-galactosidase) if more than eight amino acids had been added to β-galactosidase. Analyses of the extracts from strain MC1060 carrying pMBH191 (or pMBH198; see Fig. 2) revealed a new protein whose size was almost the same as that of native β-galactosidase (Fig. 4). This protein was absent in extracts from strain MC1060 carrying either pMBL1034 or pMBH19. These data indicate that with an appropriate adjustment of the reading frame a new protein can be made as the consequence of the insertion of the 0.95-kb *Eco*RI fragment into the *Eco*RI site of pMBL1034; this protein was presumed to be the product of the *rep-lacZ* gene fusion. The size of the *rep-lacZ* hybrid protein is that expected from the addition to β-galatosidase of 10 amino acids, the number predicted from the DNA sequence (Fig. 5).

To substantiate our belief that the *rep* regulatory signals were located within the 0.46-kb *Nru*I-*Eco*RI fragment immediately proximal to the 5' end of the *rep* gene, we removed the DNA to the left of the *Nru*I site (Fig. 2). After cleavage of pMBH191 with *Nru*I and *Xho*I, the *Xho*I cohesive end was repaired with the Klenow fragment of DNA polymerase I (the *Nru*I cut was blunt), and the linear DNA was then recircularized with T4 DNA ligase to form pMBH198 (Fig. 2). Restriction analysis of the plasmid DNA isolated from deep blue Ap^r transformants confirmed the presence of an *Xho*I-*Nru*I deletion within the 0.95-kb *Eco*RI fragment joined to the *lacZ* gene. The level of β-galactosidase activity in MC1060(pMBH198) was found to be identical to that in MC1060(pMBH191), i.e., 10,300 ± 300 U per mg of protein. These experiments demonstrated that the *rep* promoter, ribosome-binding site, and amino-terminal part of the gene were located between the *Nru*I and *Eco*RI sites.

To ascertain whether the Rep protein negatively regulated expression of the *rep* gene by acting on sequences in this region, the β-galactosidase level of cells carrying pMBH198 was measured in the presence of a compatible plasmid known to express Rep protein. There was no diminution of β-galactosidase activity directed by the *rep-lac* fusion plasmid (Table 2).

Experiments designed to determine whether the transcription of the *rep* gene was enhanced by DNA damage revealed no such effect. There was very little increase in β-galactosidase expression after treatment with either UV or mitomycin c, and derepression was not observed under conditions of *lexA* deficiency (data not shown). We conclude therefore that the *rep* gene is not among the set of genes, which includes *uvrD*, induced during the SOS response.

Identification of the promoter and N-terminal protein se-

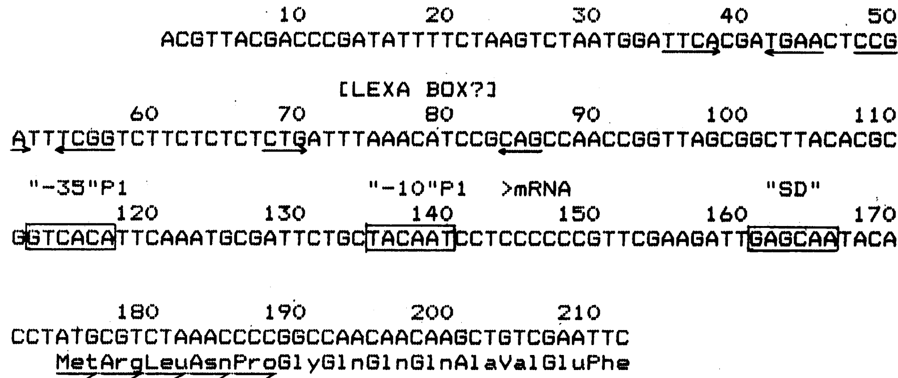

FIG. 5. Nucleotide sequence of the *rep* control region. DNA sequencing of both strands was performed by the dideoxy method (27). Only the nucleotide sequence of the coding strand is given. The −35, −10, and Shine-Dalgarno sequences are enclosed by boxes. Some regions of dyad symmetry are underlined with facing arrows. The first 13 amino acids predicted by the nucleotide sequence are shown. Residues identified by protein sequencing are identified by →.

quence. The nucleotide sequence of the 0.46-kb *Nru*I-*Eco*RI fragment was determined on both strands by sequencing the M13*rep*816, M13*rep*914, and M13*rep*827 clones described in Materials and Methods. Figure 5 displays the DNA sequence of the 212-bp segment of DNA at the *Eco*RI end of the *Nru*I-*Eco*RI fragment encompassing the transcription and translation initiation regions of the *rep* gene. In this segment there are two potential ATG initiation codons (positions 123 and 174) in the same open reading frame. Analysis of the sequence from −60 to +40 around each ATG with the W101 weighting function of Stormo et al. (34) indicated that the AUG at position 174 was more likely to be the start of a legitimate gene. The other two reading frames contain nonsense codons and lack good initiation signals.

Various promoterlike sequences can be identified upstream of the ATG at position 174 based on comparison of the nucleotide sequence of this DNA segment with the canonical promoter sequence (14). The one labeled P1 possesses a very good −10 region (TACAAT) separated by 17 bp from the −35 sequence GTCACA, which matches satisfactorily the consensus −35 TTGACA sequence. S1 mapping of the 5′ end of the *rep* transcript (Fig. 6) revealed that the predominant initiation site was the C at position 144; some transcript termini mapped to each side of this C, but not anywhere else within the 500-nucleotide fragment used. This indicates that P1 is the major promotor for the *rep* gene.

The sequence of the amino terminus of the Rep protein was determined by Geoff Flynn (Queens University, Kingston, Canada) with a sample of Rep protein provided by H. Hoffmann-Berling and A. Monem (Max Planck Institute, Heidelberg, Federal Republic of Germany). With no prior knowledge, the sequence Met-Arg-Leu-Asn-Pro was deduced; this is the predicted sequence for a protein initiated at the first ATG following the P1 promoter. The AUG initiation codon is preceded at an appropriate point upstream by the satisfactory Shine-Dalgarno ribosome-binding site sequence, GAGCAA.

TABLE 2. Expression of *rep-lacZ* is not regulated by Rep protein[a]

lac fusion plasmid	plasmid in *trans*	β-Lactamase (U/mg of protein)	β-Galactosidase (U/mg of protein)
pMBH198	None	21.2	10,500
pMBH198	pACYC184	20.6	11,300
pMBH198	pHBH139	21.8	10,900

[a] Strain MC1060 carrying the indicated plasmids was grown at 37°C in M9 minimal medium supplemented with ampicillin and chloramphenicol to log phase and assayed for enzyme activities as described previously (4) and in Materials and Methods. pHBH139 plasmid is a derivative of pACYC184 (9) carrying the *Nru*I-*Bal*I fragment (rep^+) cloned within the tetracycline resistance gene and known to overproduce Rep protein. Plasmid pMBH198 is described in the legend to Fig. 2.

DISCUSSION

In our previous report we described the isolation of the *rep* gene on a 3.2-kb *Xho*I-*Bal*I fragment of *E. coli* DNA (4). In

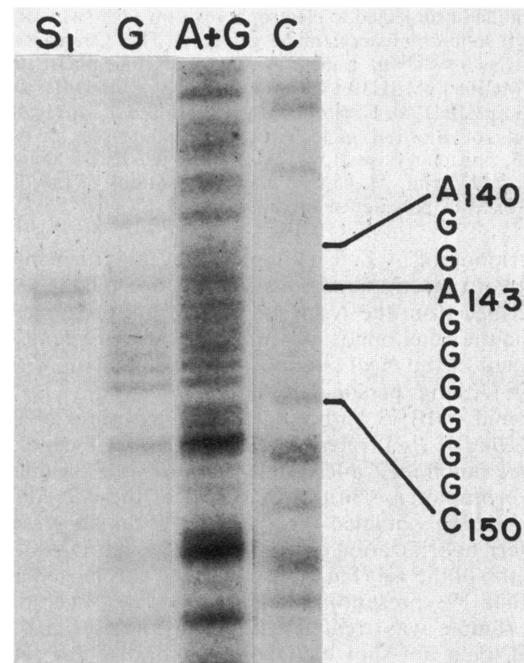

FIG. 6. S1 mapping of the 5′ end of Rep mRNA. The RNA was hybridized to a ^{32}P-5′ end-labeled restriction fragment, subjected to S1 digestion, and electrophoresed on a 8% polyacrylamide sequencing gel as described in Materials and Methods. Since one nucleotide is destroyed in the reaction, these data indicate that the major start site is at base 144.

this work we have further confined the *rep* gene to a 2.55-kb region of DNA. The nearest restriction sites outside the *rep* coding region are *NruI* at 1.6 kb and *HindII* at 4.15 kb (Fig. 1). Bäumel et al. (3) reported the cloning of the *rep* gene on a *HindIII* fragment derived from pLC44-7, the same plasmid we started with. We are unable to reconcile their results with ours.

The *rep* gene maps between *ilvC* (approximate M_r, 55,000) and *rho* (approximate M_r, 53,000) (Fig. 2) at about 84.5 map units on the *E. coli* chromosome. Between *ilvC* and *rep* there is about 0.8 kb of sequence that encodes an M_r 11,500 protein of unknown function (Atlung, personal communication). On the other side of *rep*, about 200 bp upstream of *rho*, the gene (*trxA*) encoding thioredoxin (M_r 12,000) has recently been mapped (15). The protein product of this gene, known to be an essential component of T7 DNA polymerase, has also been shown to be required for filamentous phage production (19, 28). Comparison of our mapping data with those of Russel and Model (28) place *trxA* about 2.5 kp downstream of the C terminus of the *rep* gene. We have detected by in vitro transcription-translation an M_r 54,000 protein of unknown function encoded in this region (4); this agrees with an observation made by Atlung (personal communication).

The construction and analysis of gene fusions between *rep* and *lacZ* revealed that the direction of *rep* gene transcription was clockwise on the standard *E. coli* map. The production of a hybrid protein consisting of the amino-terminal sequence of the *rep* protein joined to an enzymatically active β-galactosidase moiety placed the *rep* regulatory elements within a 0.46-kb *NruI-EcoRI* fragment. The DNA sequence supported this assignment.

The 5′ initiation point of the *rep* mRNA was localized at nucleotide C at position 144 (Fig. 5). Although transcription in *E. coli* more commonly starts with a purine residue, C is used in some *E. coli* promoters (14). The promoter contains a very good Pribnow hexamer (TACAAT) having five out of six nucleotides identical with the −10 consensus sequence (TATAAT) and a −35 sequence (GTCACA) that matches four nucleotides of the −35 consensus sequence (TTGACA) (14). The −35 and −10 sequences were found to be separated from each other by 17 bp, the preferred spacing. The five amino-terminal residues of the *rep* protein were identified and found to correspond to those deduced from the nucleotide sequence. These data unequivocally place the translation initiation site in vivo at nucleotide 174.

Analysis of the sequence upstream of *rep* revealed a possible LexA-binding site (CTGATTTAAACATCCGCAG) located downstream of a potential −10 sequence (TCTTCT) at nucleotides 58 to 63. However, an unusual feature of this hypothetical SOS box is the 13 bp (instead of 10 bp) separation of the CTG-CAG palindrome (38). Although we have observed that the *rep* promoter is slightly inducible by DNA-damaging agents, our genetic data indicate that the LexA protein does not repress the expression of the *rep* gene (data not shown). We did not find detectable derepression of synthesis of the *rep-lacZ* fusion protein in a LexA⁻(Def) strain, a genetic background in which a set of damage-inducible genes was found to be fully derepressed (38). This distinguishes the *rep* gene from the *uvrD* gene, which is controlled in part by the *lexA* repressor (1, 12).

ACKNOWLEDGMENTS

This research was supported by grants from the National Cancer Institute, the Medical Research Council, General Motors Diesel Division and Silverwood Dairies Ltd. C.A.G. was supported by a University of Western Ontario Visa Student Bursary.

We thank H. Hoffmann-Berling and A. Monem for purified *rep* protein, T. G. Flynn for determining the amino-terminal sequence, M. J. Casadaban and J. H. Miller for strains, G. Chaconas for constructive suggestions, S. Anderson and J. Miller for assistance with the sequencing, J. Colasanti for pertinent assistance, and L. Bonis and D. Marsh for preparation of the manuscript.

LITERATURE CITED

1. **Arthur, H. M., and P. B. Eastlake.** 1983. Transcriptional control of the *uvrD* gene of *Escherichia coli*. Gene **25**:309–316.
2. **Arthur, H. M., and R. G. Lloyd.** 1980. Hyper-recombination in *uvrD* mutants of *Escherichia coli* K-12. Mol. Gen. Genet. **180**:185–191.
3. **Bäumel, I., T. F. Meyer, and K. Geider.** 1984. Functional aspects of *Escherichia coli rep* helicase in unwinding and replication of DNA. Eur. J. Biochem. **138**:247–251.
4. **Bialkowska-Hobrzanska, H., and D. T. Denhardt.** 1984. The *rep* mutation. VII. Cloning and analysis of the functional *rep* gene of *Escherichia coli* K-12. Gene **28**:93–102.
5. **Biggin, M. D., T. J. Gibson, and G. F. Hong.** 1983. Buffer gradient gels and ³⁵S-label as an aid to rapid DNA sequence determination. Proc. Natl. Acad. Sci. USA **80**:3963–3965.
6. **Bridges, B. A.** 1984. Further characterization of repair of 8-methoxypsoralen crosslinks in UV-excision-defective *Escherichia coli*. Mutat. Res. **132**:153–160.
7. **Calendar, R., B. Lindqvist, G. Sironi, and A. J. Clark.** 1970. Characterization of *rep* mutants and their interactions with P2 phage. Virology **40**:72–83.
8. **Casadaban, M. J., A. Martinez-Arias, S. K. Shapira, and J. Chou.** 1983. β-galactosidase gene fusions for analyzing gene expression in *Escherichia coli* and yeast. Methods Enzymol. **100**:293–308.
9. **Chang, A. C. Y., and S. N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. **134**:1141–1156.
10. **Denhardt, D. T., M. Iwaya, and L. L. Larison.** 1972. The *rep* mutation. II. Its effect on *Escherichia coli* and on the replication of bacteriophage φX174. Virology **49**:486–496.
11. **Dugaiczyk, K. A., H. Winfield, and H. M. Goodman.** 1975. Ligation of *EcoRI* endonuclease-generated DNA fragments into linear and circular structures. J. Mol. Biol. **96**:171–184.
12. **Easton, A. M., and S. R. Kushner.** 1983. Transcription of the *uvrD* gene of *Escherichia coli* is controlled by the *lexA* repressor and by attenuation. Nucleic Acids Res. **11**:8625–8640.
13. **Fassler, J. S., I. Tessman, and E. S. Tessman.** 1985. Lethality of the double mutations *rho rep* and *rho ssb* in *Escherichia coli*. J. Bacteriol. **161**:609–614.
14. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11**:2237–2255.
15. **Höög, J.-O., H. Von Bahr-Lindström, S. Josephson, B. J. Wallace, S. R. Kushner, H. Jörnvall, and A. Holmgren.** 1984. Nucleotide sequence of the thioredoxin gene from *Escherichia coli*. Biosci. Rep. **4**:917–923.
16. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
17. **Lane, H. E. D., and D. T. Denhardt.** 1975. The *rep* mutation. IV. Slower movement of replication forks in *E. coli rep* strains. J. Mol. Biol. **97**:99–112.
18. **Larson, R., and J. Messing.** 1982. Apple II software for M13 shotgun DNA sequencing. Nucleic Acids Res. **10**:39–49.
19. **Lim, C.-J., B. Haller, and J. A. Fuchs.** 1985. Thioredoxin is the bacterial protein encoded by *fip* that is required for filamentous bacteriophage f1 assembly. J. Bacteriol. **161**:799–802.
20. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

21. **Maples, V. F., and S. R. Kushner.** 1982. DNA repair in *Escherichia coli:* identification of the *uvrD* gene product. Proc. Natl. Acad. Sci. USA **79:**5616–5620.
22. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labelled DNA with base-specific chemical cleavages. Methods Enzymol. **65:**499–560.
23. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **Nevers, P., and H. C. Spatz.** 1975. *Escherichia coli* mutants *uvrD* and *uvrE* deficient in gene conversion of λ heteroduplexes. Mol. Gen. Genet. **139:**233–248.
25. **Oeda, K., T. Horiuchi, and M. Sekiguchi.** 1982. The *uvrD* gene of *E. coli* encodes a DNA-dependent ATPase. Nature (London) **298:**98–100.
26. **Ridgway, A. A., R. A. Swift, H.-J. Kung, and D. J. Fujita.** 1985. In vitro transcription analysis of the viral promoter involved in c-*myc* activation in chicken B lymphomas: Detection and mapping of two RNA initiation sites within the reticuloendotheliosis virus long terminal repeat. J. Virol. **54:**161–170.
27. **Rothman, R. H., and A. J. Clark.** 1977. Defective excision and postreplication repair of uv-damaged DNA in a *recL* mutant strain of *E. coli* K-12. Mol. Gen. Genet. **155:**267–277.
28. **Russel, M., and P. Model.** 1984. Characterization of the cloned *fip* gene and its product. J. Bacteriol. **157:**526–532.
29. **Sanger, F., A. R. Coulson, B. G. Barrell, A. J. H. Smith, and B. A. Roe.** 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. **143:**161–178.
30. **Scott, J. F., and A. Kornberg.** 1978. Purification of the *rep* protein of *Escherichia coli.* J. Biol. Chem. **253:**3292–3297.
31. **Sharma, R. C., N. J. Sargentini, and K. C. Smith.** 1983. New mutation (*mmrA1*) in *Escherichia coli* K-12 that affects minimal medium recovery and postreplication repair after UV irradiation. J. Bacteriol. **154:**743–747.
32. **Siegel, E. C.** 1973. Ultraviolet-sensitive mutator strain of *Escherichia coli* K-12. J. Bacteriol. **113:**145–160.
33. **Silhavy, T. Y., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
34. **Stormo, G. D., T. D. Schneider, L. Gold, and A. Ehrenfeucht.** Use of the 'Perceptron' algorithm to distinguish translational initiation sites in *E. coli.* Nucleic Acids Res. **10:**2997–3011.
35. **Takahashi, S., C. Hours, M. Iwaya, H. E. D. Lane, and D. T. Denhardt.** 1978. The *Escherichia coli rep* gene, p. 393–400. *In* D. T. Denhardt, D. H. Dressler, and D. S. Ray (ed.), The single-stranded DNA phages. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
36. **Taucher-Scholz, G., M. Abdel-Monem, and H. Hoffmann-Berling.** 1983. Functions of DNA helicases in *Escherichia coli.* UCLA Symp. Mol. Cell. Biol. **10:**1–12.
37. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA **77:**5201–5205.
38. **Walker, G. C.** 1984. Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli.* Microbiol. Rev. **48:**60–93.
39. **Yang, R. C., J. Liss, and R. Wu.** 1979. Elution of DNA from agarose gels after electrophoresis. Methods Enzymol. **68:**176–182.