Identification of the uvrC gene product
(plasmid/maxicell/transposon Tn1000/DNA repair/truncated peptide)

AZIZ SANCAR*, BARRY M. KACINSKI†, D. LAWRIE MOTT‡, AND W. DEAN RUPP*†

Departments of *Therapeutic Radiology and †Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, Connecticut 06510

Communicated by Edward A. Adelberg, June 1, 1981

ABSTRACT We have constructed a multicopy plasmid that carries the uvrC gene of Escherichia coli. By inserting the transposon Tn1000 (previously designated y6) into this plasmid, we obtained many derivatives that fail to complement uvrC34. The proteins synthesized by the original plasmid and the uvrC::Tn1000 derivatives were labeled in maxicells and analyzed on gels, demonstrating that a protein of M, 70,000 encoded by the original uvrC+ plasmid was absent from the mutated nocomplementing derivatives; this protein is presumed to be the uvrC gene product. We found that this protein of M, 70,000 binds to DNA and have partially purified the uvrC gene product by DNA-cellulose chromatography. Because some of the uvrC::Tn1000 derivatives produce truncated polypeptides, the orientation of expression and the location of the promoter were determined by correlating the sizes of the truncated polypeptides with the sites of insertion of Tn1000.

The excision repair pathway is one of the major mechanisms for repairing damaged DNA in Escherichia coli. In this mode of repair, the bases that are damaged by UV radiation or chemicals are removed by a combined action of endo- and exonucleases and the resulting gap is filled by DNA polymerases and sealed by ligase (for review, see ref. 1). The uvrA, uvrB, and uvrC genes control the early steps of excision repair in vitro, and all three gene products are required for the endonucleolytic incision of UV-irradiated DNA in vitro (2, 3). By using this in vitro complementation assay, Seeberg and his colleagues have begun to separate the various components and have reported a DNA-binding activity of the partially purified uvrA+ protein (4–6). However, the functions of the individual gene products in the incision reaction and the site of incision relative to the damaged base have not been determined because the purified proteins have not been available.

In an effort to identify, characterize, and overproduce the uvr gene products, we have developed an alternative approach that combines several methods. The first step is to clone a gene on a multicopy plasmid. The next step is to use the maxicell procedure to label those proteins encoded by the plasmid carrying the cloned gene. As genes adjacent to the selected gene may also encode proteins, it is necessary to identify unambiguously the specific protein from the gene in question. This has been accomplished by inactivating the cloned gene by insertion of the transposon Tn1000 (previously designated y6). The protein lost by this insertional mutation is thus the product encoded by the gene in question. Furthermore, in many cases, these insertional mutations are unstable and can be excised by the transposon, and the excised transposon can thus be used to determine the orientation of the gene by comparison of their sizes with the insertion sites mapped by using restriction endonucleases. Maxicell preparations are useful because they can be used to obtain radiochemically pure preparations of active enzymes that can then be used as tracers for further purification of the proteins.

By using this approach, we have identified and partially purified the gene products of all three uvr genes. We found (7) that the uvrA gene codes for a protein of M, 114,000 that binds to DNA and that the uvrB gene codes for a protein of M, 84,000 (8) that has no DNA binding activity by itself (unpublished results). In this communication, we show that the uvrC gene codes for a protein of M, 70,000 that binds to DNA. The DNA-binding property of the uvrC protein has enabled us to obtain radiochemically labeled uvrC protein in one chromatographic step with little contamination by other radioactive proteins.

MATERIALS AND METHODS

Bacterial Strains. The E. coli K-12 derivatives used were MG1063 (recA1/F·), CSR603 (recA1 uvrA6 phr-1, DR153 (recA3 ΔuvrB), DR1984 (recA1 uvrC34), and DM1415 (recA1 spr-1). All of these strains have been described (9–11).

Cloning of the uvrC Gene. Approximately 4 μg of plC13-12 DNA and 2 μg of pBR322 DNA were digested with Pst I restriction endonuclease, and the mixture was ligated with T4 DNA ligase. Competent DR1984 (recA1 uvrC34) cells were transformed with the ligated DNA, and 0.1-ml samples were plated on Luria agar plates containing tetracycline (10 μg/ml). The plates were irradiated at a UV fluence of 1.0 J/m2 from a germicidal lamp and incubated at 37°C for 18 hr. There were ~10 surviving colonies per plate, and 6 out of 10 colonies tested showed a UV survival similar to that of recA+ uvrC+ cells. One of the resistant colonies, DR1984/pDR3003, was chosen for further studies. It contained a 7.8-kilobase (kb) plasmid, pDR3003, that transformed strain DR1984 to a UV-resistant tetracycline-resistant phenotype at a frequency of ~106 transformants per μg of DNA. This plasmid, pDR3003, was thus presumed to carry the uvrC gene.

Isolation of uvrC::Tn1000 Mutants. Tn1000 was inserted into uvrC by mating MG1063/pDR3003 [tetracycline resistant (TetR) streptomycin sensitive (StrR)] with DR1984 [tetracycline sensitive (TetS) streptomycin resistant (StrR)], selecting for TetR StrR transconjugants, and then screening for UV-sensitive colonies. The rationale and mechanics of the selection procedure have been described (7, 8, 10).

Maxicells. The proteins encoded by the uvrC+ plasmid and its derivatives were identified in maxicells (12). Strain CSR603 was used as the host strain in experiments in which analytical gels were used for analysis, and DM1415 (recA1 spr-1) was used as the host strain for larger scale preparations of uvrC protein. The plasmid-bearing cells were irradiated at a UV fluence of 25 J/m2. D-cycloserine (200 μg/ml) was added to the culture 1 hr after irradiation, and the cells were incubated at 37°C for 12–14 hr. The culture was then labeled with [35S]methionine, and the labeled proteins were analyzed by electrophoresis on NaDodSO4/polyacrylamide gels [acylamide/N,N′-diallyltar-
tardiamide (10:1]) followed by autoradiography of the dried gels.

Cells carrying the uvrC plasmid pDR3003 have an extreme tendency to aggregate in liquid cultures. This aggregation is not dependent on a functional uvrC protein as it is observed in all the uvrC::Tn1000 derivatives of pDR3003. Cell aggregation leads to nonuniform irradiation during the maxicell procedure and can lead to a high background in the labeling of plasmid proteins. This problem is somewhat ameliorated by growing the cells in a medium containing 0.1 M NaCl and irradiating at a cell density of ~5 × 10^10/ml.

DNA-Cellulose Chromatography. The uvrC protein was partially purified by DNA-cellulose chromatography. Denatured DNA-cellulose (P-L Biochemicals) was mixed with an equal volume of buffer A [50 mM 3-(N-morpholino) sulfonic acid adjusted to pH 7.5 with KOH/100 mM KCl/10 mM 2-mercaptoethanol/1 mM EDTA/25% (wt/vol) glycerol], and the resultant slurry was packed into a Pharmacia K-9 column to give a 15-ml bed volume. The column was equilibrated by washing with 3 bed volumes of buffer A. Maxicells from 500 ml of a culture of DM1415/pDR3003 labeled with [35S]methionine were lysed as described (3). The cell-free extract (total volume, 5 ml) was dialyzed against buffer A and applied to the column. The column was washed with 50 ml of buffer A, and elution was carried out with a gradient of KCl in buffer A (0.1-1.0 M). Fractions of 1.5 ml were collected, and 0.1 ml of each fraction was assayed in a liquid scintillation counter.

RESULTS

Cloning of the uvrC Gene. The Clarke-Carbon plasmid pLC13-12 is known to carry the flaD gene (13, 14) and has been reported by van Sluis and Pannekoek (15) to carry the uvrC gene, in accordance with the close proximity of the map positions of these two genes (16). We confirmed that this plasmid complements a defect in uvrC. LC13-12 was mated with DR1984, and it was observed that all colicine E1-resistant Str^R transconjugants were also UV resistant (data not shown). The uvrC gene was subcloned from pLC13-12 into pBR322 to generate pDR3003 (tet^ uvrC^). The effect of this plasmid on the survival of DR1984 (recA1 uvrC34) is shown in Fig. 1. The plasmid restores the UV survival of DR1984 to the level of a recA strain carrying a wild-type uvrC gene. pDR3003 did not have any effect on the survival of CSB903 (recA1 uvrA6 phr-1) or DR153 (recA3 ΔuvrB) (data not shown), indicating that the complementation was specific for the uvrC gene. We thus conclude that pDR3003 carries the uvrC gene and not some nonspecific suppressor.

Identification of the uvrC Protein. pDR3003 (tet^ uvrC^) produces several proteins in maxicells (Fig. 2); the M_r 35,000 protein corresponds to the tet protein of the vector. To determine which of the other proteins is the uvrC gene product, we obtained Tn1000 insertions in the uvrC gene. Of the 191 independent pDR3003::Tn1000 isolates tested, we found that 44 were UV sensitive and presumed these to be insertions into the uvrC gene. When the proteins encoded by these plasmids were analyzed in maxicells, we found that they did not synthesize the M_r 70,000 protein. Fig. 2 also shows the maxicell proteins of one of these plasmids, pDR3093 (tet^ uvrC^) and pBR322 (tet^ uvrC^). The M_r 70,000 protein has disappeared and a smaller polypeptide (M_r 65,000) is now observed. We conclude that the M_r 70,000 protein is the uvrC gene product and that the M_r 65,000 protein is a truncated derivative of the uvrC protein produced as a consequence of the interruption of the uvrC gene by Tn1000.

![](image1.png)

**Fig. 1.** UV survival of DR1984 (recA1 uvrC34) and its plasmid-carrying derivatives. ○, DR1984; □, pDR3003 in DR1984; △, pDR3024 in DR1984. Strains were grown to stationary phase in Luria broth and dilutions were plated on Luria agar plates that were irradiated with UV light from a germicidal lamp at a rate of 0.05 J/m^2/sec. The surviving colonies were counted 24 hr after irradiation.

![](image2.png)

**Fig. 2.** Synthesis of uvr proteins in maxicell: autoradiogram of a 14% polyacrylamide gel. Lanes: 1, pBR322 (tet^ blu^); 2, pDR2000 (tet^ blu^ uvrA^); 3, pDR1494 (tet^ blu^ uvrB+); 4, pDR3003 (tet^ blu^ uvrC^); 5, pDR3093 (tet^ blu^ uvrC3093::Tn1000). M_r markers used to determine the size of the uvrC protein were E. coli RNA polymerase β subunits (160,000), uvrA (144,000), uvrB (84,000), bovine serum albumin (67,000), recA protein (35,000), carbonic anhydrase (29,000), and E. coli single-strand DNA binding protein (19,000).
Polypeptides produce by the insertion on (tety pDR8003; graphed. CSR603 Tn1O0O.)tm, polypeptides detectable.

**FIG. 3.** Proteins synthesized by uvrC::Tn1000 derivatives of pDR3003 (tet+ uvrC*). 35S-Labeled maxicell proteins were separated on a 14% polyacrylamide gel that was then dried and autoradiographed. CSR903 was the host strain. Lanes: 1, pBR322; 2, pDR3024; 3, pDR3003; 4, pDR3006; 5, pDR3003; 6, pDR3006; 7, pDR3017. M, markers were the same as used in Fig. 2. The protein band seen at M, 38,000 in lanes 4 and 7 is a background cellular protein that is labeled in suboptimal maxicell preparations.

**Physical Map of uvrC.** A number of the uvrC::Tn1000 plasmids produce truncated peptides in maxicells (Fig. 3). These truncated polypeptides are presumed to be caused by the termination codons introduced into uvrC by Tn1000 (7) and therefore their sizes reflect the distances between the insertion sites and the beginning of the uvrC structural gene. The sites of Tn1000 insertion in these uvrC::Tn1000 plasmids were determined by restriction mapping (data not shown). By comparing the insertion sites with the sizes of the truncated polypeptides, the start point of the structural gene was determined to be at 7.0 kb of the pDR3003 map (Fig. 4).

One of the pDR3003::Tn1000 plasmids produced UV sensitivity intermediate between that of pDR3003 and that of the other uvrC::Tn1000 insertion derivatives when present in a uvrC34 strain (Fig. 1). Although this plasmid, pDR3024, does not produce the 70,000 protein in maxicells, rather than producing a truncated uvrC derivative, it encodes a M, 71,000 protein slightly larger than the normal uvrC protein (Fig. 3). The mapped insertion site of Tn1000 in this plasmid is near the end of the uvrC gene corresponding to the COOH terminus of the uvrC protein (Fig. 4), suggesting the following interpretation. (i) Tn1000 is inserted in the uvrC gene at a position near, but preceding, the normal termination codon for the uvrC protein. (ii) Transcription from the uvrC promoter will proceed into Tn1000 and, depending on the position of the first in-phase termination codon, translation of this hybrid uvrC::Tn1000 mRNA can produce a hybrid polypeptide slightly longer than the original uvrC protein. As this hybrid protein contains most of the amino acids of the uvrC protein, it may well retain partial function, accounting for the intermediate UV sensitivity of a strain carrying this plasmid.

**Radiochemical Purification of the uvrC Protein.** The availability of radiochemically labeled uvrC protein in maxicell preparations enabled us to study its interaction with the other uvr proteins, as well as with DNA. We thus found that uvrC protein binds to single-stranded DNA. This observation led us to devise a simple procedure for obtaining 35S-labeled uvrC protein substantially free from other radioactive proteins. A maxicell extract is applied to a DNA-cellulose column, and the proteins are eluted with a KCl gradient; a radioactive peak elutes at ~350 mM KCl (Fig. 5A). When the proteins in this peak are analyzed by polyacrylamide gel electrophoresis and autoradiography, the only apparent radioactive protein band corresponds to uvrC protein (Fig. 5B). These properties of the uvrC protein are similar to those of the uvrA protein (7) but contrast to those of the uvrB protein, which does not bind to DNA under these conditions (unpublished results).

**DISCUSSION.**

Our data suggest that the uvrC protein is a DNA-binding protein of M, 70,000. This conclusion is based on the evidence that, in maxicells of strains carrying the uvrC plasmid, a protein of M, 70,000 is produced and that this protein disappears in all uvrC::Tn1000 derivatives that fail to complement a uvrC mu-

**FIG. 4.** Locating the uvrC promoter. (A) Sites of insertion of Tn1000 in the uvrC plasmid pDR3003. Clockwise insertions of Tn1000 (with regard to the pDR3003 restriction map) are indicated by arrows inside the circle and counterclockwise insertions are indicated by arrows outside the circle. (When proceeding around the map clockwise, the two EcoRI sites in Tn1000 precede the BamHI site in Tn1000 for the clockwise insertion of Tn1000.) pBR322 DNA; the location and orientation of the uvrC gene is indicated by a heavy arrow on the plasmid map. The locations and orientations of the insertions were determined by digesting the plasmids with BamHI, EcoRI, and Psi I (data not shown). (B) Size of truncated uvrC polypeptides as a function of Tn1000 insertion sites. The sizes of the truncated polypeptides produced by uvrC::Tn1000 plasmids (from Fig. 3) are plotted as a function of the insertion sites shown in A. The plasmids pDR3027, pDR3017, and pDR3029 do not produce stable truncated peptides detectable in maxicells (Fig. 3; data not shown).
Radirochemical purification of the uvrC protein. (A) Chromatography on single-strand DNA-cellulose of maxicell proteins in the cell-free extract from DM1415/pDR3003. Acid-precipitable radioactivity is plotted as a function of KCl concentration in the eluate. (B) Polyacrylamide gel analysis of the peak fraction of the eluate from the DNA cellulose column. A 10% polyacrylamide gel was dried and autoradiographed for 7 days. Lanes: 1, 50 μl of cell-free extract that was applied to the column; 2, 75 μl of fraction 33; 3, 25 μl of fraction 33.

The role of the uvrC gene was controversial. Although there is clearly a requirement for the uvrC gene product in cell-free incision experiments, uvrC mutants accumulate single-strand breaks in vivo, albeit at a low rate, a phenomenon not observed with either uvrA or uvrB mutants (2, 17). One explanation, suggested by Seeberg et al. (2), is that the cells may have an analog of uvrC protein that is partially functional and allows abnormal incision. However, other explanations are just as plausible. For example, if uvrC is an essential gene, then all previously isolated uvrC mutations would retain partial activity, explaining why they might be detected by the more sensitive in vitro assay but not in the cell-free complementation assay. To distinguish among these various possibilities, it is essential to conduct cell-free experiments with purified uvr proteins. The identification of the uvrC protein reported here, together with its radirochemical purification, should help in achieving this goal.

Now that the three uvr genes, uvrA, uvrB, and uvrC, have been cloned on multicopy plasmids and the gene products have been identified, rapid progress can be expected in understanding the precise steps of nucleotide excision repair catalyzed by these proteins.

Yoakum et al. (18, 19) also have reported the cloning of the uvrC gene and identification of the gene product.

We thank Drs. C. A. van Sluis and H. Pannekoek for generously providing information prior to publication regarding restriction sites in the region of the uvrC gene. This work was supported by National Institutes of Health Grants CA 05319 and CA 16038 and American Cancer Society Grant PDT-80C. B. M. K. was a fellow of the U.S. Public Health Service Medical Scientist Training Program (Grant GM 07205).