Volume 9 Number 21 1981 Nucleic Acids Research

The structure and function of the regulatory elements of the Escherichia coli uvrB gene

Eva van den Berg*, Joke Zwetsloot*, Inge Noordermeer*, Hans Pannekoek*†, Ben Dekker+, Rein Dijkema+ and Hans van Ormondt+

*Department of Molecular Genetics, and +Laboratory of Physiological Chemistry, State University Leiden, 2333 AL Leiden, The Netherlands

Received 15 June 1981

ABSTRACT

The construction and properties of recombinant plasmids carrying the *Escherichia coli uvrB* gene, including its transcriptional- and translational regulatory elements, is reported. The DNA sequence of the region, which governs the expression of the *uvrB* gene, has been determined. Within this sequence two non-overlapping DNA segments match the model sequence for *Escherichia coli* promoters (1). The '-10 regions' and the '-35 regions' of the proposed *uvrB* promoters are, respectively, 5'TAAAAT (P1), 5'TATAAT (P2) and 5'TTGGCA (P1), 5'GTGATG (P2). The existence and the position of these promoters has been established by elimination of one promoter (P2), using molecular cloning procedures, by length measurements of *in vitro* synthesized 'run-off' transcripts and by protection of the *uvrB* regulatory region for S1 nuclease digestion using *in vivo* made RNA. Potential sites of interaction within the *uvrB* regulatory region with regulatory proteins, such as the LexA protein (2) and the UvrC protein (3) are discussed.

INTRODUCTION

The *Escherichia coli uvrA, uvrB* and *uvrC* genes, which are physically unlinked, encode proteins involved in the first reaction (*i.e.* incision) of the excision repair pathway. A mutation in any one of these genes impairs the introduction of single strand breaks into damaged DNA (4,5). It has been proposed that the UvrA, UvrB and UvrC proteins constitute an enzymatic complex (6). If, the three proteins are complexed then it is conceivable that the expression of the *uvr* genes is coordinated. Recently, data have been reported on the regulation of the *uvrA* and *uvrB* genes (7,8,2), which indicate that the expression of these genes can be induced by treatment of the cells with DNA damaging agents. Such treatments are known to activate the proteolytic function of the RecA protein, which protein in turn will inactivate several repressors causing the induction of various genes. The observations on the inducibility of the *uvrA* and *uvrB* genes are explained by assuming that the LexA protein acts as a repressor for these genes, whereas induction is carried out by the RecA protease which would cleave and inactivate the LexA protein.

The regulation of expression of the *uvrB* gene appears to be even more complex. We have recently found, in systems lacking the *lexA-recA* control circuit, that the *uvrB* gene is regulated by the UvrC protein (3). This protein reduced the biosynthesis of the UvrB protein and operates at the level of transcription. Consequently, different regulatory mechanisms affect the expression of the *uvrB* gene.

To better understand the action of regulatory proteins with the regulatory elements of the *uvrB* gene, it is necessary to elucidate its DNA sequence. In view of those perspectives, we report in this paper on the DNA sequence of the N-terminal part of the *uvrB* gene. The position and the structure of the *uvrB* promoter is described, while possibilities for sites of interaction of regulatory proteins with the *uvrB* regulatory region are discussed.

EXPERIMENTAL PROCEDURES

E.coli K12 bacterial strains and plasmids.

Isogenic strains XA106 and HP3435 (Δ*uvrB*) were described previously (9). The *bioD* mutant strains DD19 F^- *his* Sm^r *bioD*546 and N512 F^- *his* Sm^r *bioD*565 were obtained from A. Campbell (Stanford University, Cal.). Strain SR362 F^- *thyR metE leuB* Sm^r *recA*56 *uvrA*6 *uvrB*5 *phr* was obtained from K.C. Smith (Stanford University, Cal.). Plasmids pBR322 and pACYC177 were from, respectively, H.L. Heyneker (Genentech Inc. Cal.) and S.N. Cohen (Stanford University, Cal.).

Molecular cloning procedures.

Digestion of plasmid DNA with restriction endonucleases was done as prescribed by the enzyme manufacturers. Ligation of DNA fragments containing cohesive ends and transformation of competent cells was carried out as described before (10). Initial characterization of a UV resistant phenotype of transformants of strain HP3435 (Δ*uvrB*) was done by streaking single colonies and irradiation of a segment of the plates with a UV dose of 150 erg/mm^2. A vector with a 'blunt' end and an EcoRI cohesive end was constructed as follows: 5 μg of pNP12 DNA (see Fig. 1) was digested to completion with PstI, subsequently extracted with phenol-chloroform and then precipitated with ethanol. The precipitated DNA was dissolved in sufficient water to give a concentration of about 100 μg/ml. This DNA preparation was incubated with S1 nuclease (50 units/ml) for 20 min at 22°C in 30 mM sodium-acetate (pH 4.6) 1 mM $ZnCl_2$, 50 mM NaCl and 10% (v/v) glycerol. The reaction was arrested by adding an equal volume of 0.1 M TRIS (pH 10.0) and the mixture was extracted with phenol-

chloroform. This DNA preparation was further purified on a Sephadex G-50 column and eluted with distilled water. The DNA was then digested with EcoRI, again phenol-chloroform extracted and finally, the two fragments were separated on a Malachit-green column (0.5 x 2 cm; Boehringer-Mannheim, BRD) by elution with a gradient of 0.5 M-1.2 M $NaClO_4$ in 10 mM sodium-phosphate (pH 6.0). The fractions containing the large fragment of pNP12 DNA were pooled, precipitated with ethanol and the excess of salt was removed by Sephadex G-50 chromatography. This preparation serves as a vector for cloning fragments with a 'blunt' end and an EcoRI end, according to the protocol described before (10). The preparation of 'blunt ended' HphI fragments of pNP12 DNA was done as outlined above. These fragments were subsequently digested with EcoRI and purified as outlined.

Plasmid-encoded protein synthesis in maxi-cells.

Maxi-cells of strain SR362 *recA uvrA uvrB phr*, containing various multicopy plasmids, were prepared essentially as described (3). Modifications of this procedure (communicated to us by W.D. Rupp, Yale University, New Haven) are the following: i) the bacteria were irradiated with a low UV dose of 10 erg/mm^2 and incubated for 1 h at 32°C, ii) D-cycloserine was administered to a final concentration of 200 µg/ml, followed by overnight incubation at 32°C to degrade chromosomal DNA. Labelling of maxi-cells with |^{35}S|-methionine, SDS-polyacrylamide gelelectrophoresis and fluorography was performed as described (3).

DNA sequencing.

The protocol for removal of 5' phosphate groups with alkaline phosphatase, labelling of 5' hydroxyl groups with γ-|^{32}P|-ATP and T4-polynucleotide kinase and the chemical modification and degradation of labelled DNA has been described (11).

In vitro transcription using restriction fragments as templates.

Plasmid pNP12 DNA was digested to completion with HaeIII and subjected to electrophoresis on a 5% polyacrylamide slabgel. The HaeIII fragment (about 290 bp.), containing the *uvrB* promoter, was eluted from gelslices as outlined (11). Incubations were carried out for 10 min at 37°C. The reaction mixture (20 µl) contained: 25 mM TRIS-HCl (pH 8.0), 10 mM $MgCl_2$, 50 mM KCl, 0.1 mM dithiothreitol, 5% (v/v) glycerol, 0.16 mM each of ATP, GTP and CTP, 7.5 - 10 µM α-|^{32}P|-UTP (specific radioactivity 200 Ci/mmole), 10 µg/ml DNA and 12.5 µg/ml RNA polymerase. The reaction was arrested by adding 5 volumes of cold 25 mM TRIS-HCl (pH 8.0), 10 mM $MgCl_2$, 50 mM KCl, 0.1 mM dithiotreitol

and subsequently extracted with phenol-chloroform. After precipitation with ethanol, the RNA was dissolved in 70% (v/v) formamide, 10% (v/v) glycerol, containing dye markers bromophenol blue and xylenecyanol FF. Samples were loaded on a 8% polyacrylamide, 7 M urea slabgel and subjected to electrophoresis, followed by autoradiography.

S1 nuclease protection of DNA containing the *uvrB* promoter region by *in vivo* made RNA

25 ml cultures of strain HP3435 (Δ*uvrB*) or of strain HP3435 containing plasmid pNP12 were grown in L-broth till a density of 6 x 10^8 cells per ml and nucleic acids were isolated as described (12). This preparation was treated with 60 µg/ml pancreatic DNase for 30 min at 37°C, extracted with phenol and chloroform and precipitated with ethanol. As a control the DNase digestion to remove competitor pNP12 DNA during DNA-RNA hybridization was omitted for part of the nucleic acid preparation. Plasmid pNP12 DNA was uniformily labelled by 'nick translation' using DNA polymerase I and α-|^{32}P|-dATP as described (13). Labelled pNP12 DNA was digested with HaeIII, subjected to electrophoresis on a 5% polyacrylamide slabgel and the fragment (290 bp.) containing the *uvrB* regulatory region was isolated (11). DNA-RNA hybridization of |^{32}P|-HaeIII DNA and *in vivo* made RNA, followed by incubation with S1 nuclease was done essentially as described (14), except that hybridization was performed at 45°C and the S1 nuclease digestion was for 30 min at 19°C with 10 - 20 units of enzyme per ml. Single stranded |^{32}P|-labelled DNA was analyzed on a 8% polyacrylamide, 7 M urea slabgel.

Enzymes and radiochemicals

Restriction enzymes were from the following manufacturers: HaeIII (BRL Inc. Rockville, MD), HphI and TaqI (Biolabs, Beverley, MA), EcoRI (Miles Lab., Elkhart, Ind.). PstI and AluI were donated by H.L. Heyneker (Genentech Inc., San Francisco, Cal.). DNA polymerase I was from Boehringer Mannheim (BRD). S1 nuclease from *Aspergillus oryzae* was purchased from P-L Biochemicals Inc., Milwaukee, Wis. *Escherichia coli* RNA polymerase was purified by Polymin P precipitation and elution with NaCl (15), followed by chromatography on DEAE-cellulose (16) and finally by affinity chromatography on Heparin-Sepharose (17). This preparation is of high specific activity and at least 90% pure as judged by electrophoresis on a SDS-polyacrylamide gel. α-|^{32}P|-UTP and α-|^{32}P|-dATP were obtained from the Radiochemical Centre (Amersham, England). Enzymes and radiochemicals employed for DNA sequencing have been described (18).

RESULTS AND DISCUSSION

Previous studies have shown that cloning of the *E.coli uvrB* locus, located on an EcoRI fragment derived from the transducing phage $\lambda b2att^2$, on multicopy plasmids results in the expression of the *uvrB* gene under control of a promoter on the vector part (19). Although those *uvrB* recombinant plasmids lack the genuine *uvrB* promoter and an N-terminal segment of the gene, they fully complement the UV resistance in an UvrB deletion strain. Apparently, the N-terminal part of the *uvrB* gene product is not essential for the enzymatic activity required for UV resistance. This conclusion is strengthened by our finding that non-polar mutations in the N-terminal part of the *uvrB* gene do not affect the UV resistance displayed by UvrB deletion strains harbouring mutated *uvrB* plasmids (20).

The direction of transcription of the *uvrB* gene has been established both on the chromosome (*i.e.* from *bio* to *chlA*) and on an *uvrB* plasmid, denoted pNP5 (10, 19), which consists of an EcoRI fragment inserted into plasmid pMB9 (21)(Fig. 1). The direction of transcription and the observation that plasmid pNP5 lacks the *uvrB* promoter allowed us to conclude that the promoter is located upstream of the EcoRI site on the chromosome (corresponding with the EcoRI site at 0% on plasmid pNP5 in Fig. 1). To localize the regulatory elements of the *uvrB* gene more precisily we cloned more extended fragments than the EcoRI fragment.

Rupp et al. (22) have shown that the colony bank plasmid pLC25-23 (23) carries the biotin operon and the *uvrB* gene. Since the cloned fragment on plasmid pNP5 does not contain sites for the enzymes PstI and BamHI, we have chosen these restriction endonucleases to digest pLC25-23 DNA in order to reclone the *uvrB* gene on either pBR322 (24) or on pACYC177 (25). Plasmid pNP10 (9.6 kb) is composed of a PstI fragment (5.2 kb) inserted into the unique PstI site in the β-lactamase gene (Ap^r) of plasmid pBR322, while pNP12 (9.5 kb) results from the integration of a PstI-BamHI fragment (4.8 kb) into plasmid pACYC177 (Fig. 1). Both plasmids fully restore an UV resistant phenotype of UvrB deletions strains (illustrated for pNP12 in Fig. 3).

Properties of plasmids pNP10 and pNP12

The orientation of the *uvrB* gene on plasmids pNP10 and pNP12 can be deduced from the data acquired with plasmid pNP5 (10, 19). The *uvrB*-containing segment on the cloned EcoRI fragment of pNP5 is located between 75.4% and 100% on the physical map and comprises approximately 1,920 bp. The distance between the PstI and the EcoRI site on the inserted fragment in pNP10 and pNP12 is 1,250 bp. The PstI site must be located upstream of the EcoRI site

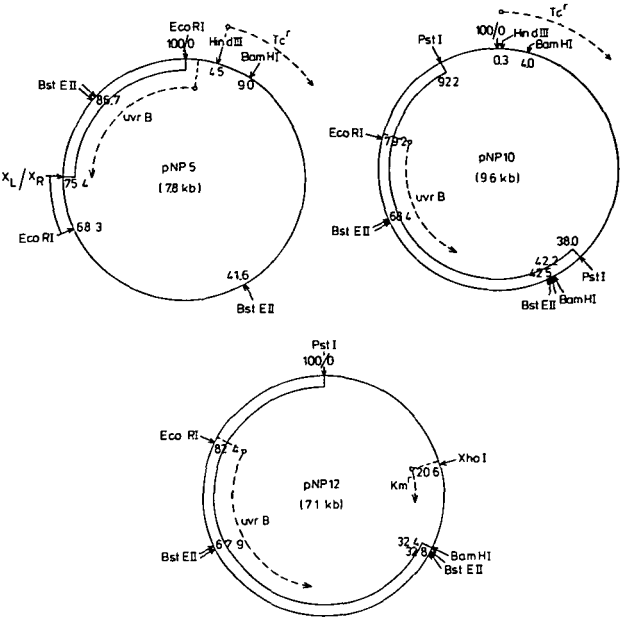

Figure 1. Diagrammatic representation of the structure of plasmids pNP5, pNP10 and pNP12. Plasmid pNP5 (Tc^r) consists of vector pMB9 (Tc^r; 5.3 kb, ref. 21) and an EcoRI fragment (2.5 kb) harbouring the *uvrB* gene. This EcoRI fragment has been derived from an *uvrB* transducing phage $\lambda b2att^2$ (10). The segment between 75.4% and 100% is of bacterial origin, whereas the region from 68.3 to 75.4% is of phage λ origin. Expression of the *uvrB* gene on pNP5 is dependent on a pMB9 promoter located on the 354 bp. EcoRI-HindIII fragment of pMB9 (19, 28). Plasmid pNP10 (Ap^s Tc^r) consists of vector pBR322 (Ap^r Tc^r; 4.4 kb, ref. 24) and an *uvrB*-containing PstI fragment (5.2 kb) derived from the colony-bank plasmid pLC25-23 (23). This PstI fragment has been inserted into the unique PstI site within the β-lactamase gene of pBR322. Plasmid pNP12 (Ap^s Km^r; 7.1 kb) is composed of vector pACYC177 (Ap^r Km^r; 3.7 kb, ref. 25) and a PstI-BamHI fragment (4.8 kb), derived from pLC25-23, harbouring the *uvrB* gene. This PstI-BamHI fragment has replaced the original PstI-BamHI segment of the vector. The position of restriction sites for various restriction endonucleases, which are relevant for the determination of the localization and the orientation of the *uvrB* gene, is indicated. The dotted lines represent the direction of transcription of the *uvrB* gene and of genes specifying resistance to antibiotics.

within the *uvrB* gene, since no PstI site is present within the *uvrB* segment of pNP5. This conclusion is confirmed by the position of the BstEII sites within the *uvrB* gene relative to the EcoRI sites on plasmids pNP10 and pNP12 (results not shown). Unless the *uvrB* gene is preceded by another gene belonging to the same operon it is conceivable that the *uvrB* promoter is

Table 1. Marker rescue of *bioD* mutations by plasmid pNP12

Strain	Plasmid	Recombination frequency
DD19 *bioD*546	pACYC177	$<10^{-6}$
DD19 *bioD*546	pNP12	2×10^{-3}
N512 *bioD*565	pACYC177	$<10^{-6}$
N512 *bioD*565	pNP12	4×10^{-4}

Strains DD19 *bioD*546 and N512 *bioD*565 were transformed either with pACYC177 DNA (25) or with pNP12 DNA. Kanamycin resistant transformants are unable to grow on minimal medium plates (30) without biotin, indicating that complementation of the *bioD* mutations does not occur. Recombinants were scored after plating dilutions on minimal medium plates, containing 50 µg per ml kanamycin, but lacking biotin. Control plates of diluted bacteria were minimal medium containing kanamycin and biotin.

located on the Pst-EcoRI fragments (1,250 bp.) of pNP10 and pNP12.

On the *E.coli* gene map the *uvrB* gene has been positioned adjacent to the right (C-terminal) side of the *bioD* gene. Both the *bioD* gene and the *uvrB* gene are transcribed clockwise on the chromosome, whereas the direction of transcription on plasmid pNP12 is depicted as anti-clockwise (Fig. 1). Unless the distance between the *bioD* gene and the *uvrB* gene exceeds 1,250 bp., the PstI-EcoRI fragment must harbour at least a part of the *bioD* gene and also the *uvrB* promoter. These assumptions were confirmed by marker rescue experiments for chromosomal *bioD* mutations and plasmid pNP12 (Table 1). Recombination between chromosomal DNA of either strain DD19 (*bioD*546) or strain N512 (*bioD*565) with pNP12 yields recombinants which grow in the absence of biotin. Bio^+ recombinants are not found when these strains harbour plasmid pACYC177. Complementation of the *bioD* mutations by plasmid pNP12 (or pNP10) is not observed, indicating that either the *bioD* is not expressed on these plasmids or they contain only part of the *bioD* gene. We conclude that the PstI-EcoRI fragment (1,250 bp.) harbours at least a part of the *bioD* gene and therefore must contain the *uvrB* promoter.

Several arguments can be advanced which indicate that the expression of the cloned *uvrB* gene is mediated by the *uvrB* promoter and not by a promoter located on one of the vectors. Firstly, the direction of transcription of the *uvrB* gene on pNP10 is parallel to that of the β-lactamase (Ap^r) gene of pBR322 (26) in which gene the PstI fragment has been integrated. The introduction of a polar mutation, through the integration of transposon Tn5 (Km^r;

ref. 27), between the promoter of the β-lactamase gene and the *uvrB* promoter does not alter the UV resistant phenotype of UvrB deletion strains transformed with pNP10::Tn5 DNA. This observation indicates that the adjacent promoter of the β-lactamase gene is not required for expression of the *uvrB* gene. Secondly, the *uvrB* gene on plasmid pNP12, located on a PstI-BamHI fragment, has been cloned on the vector pACYC177 (25) in a region which does not contain promoters. Hence, it is likely that the expression of the *uvrB* gene on pNP12 is mediated by the *uvrB* promoter, which provides for a level of UV resistance in UvrB deletion strains identical to that of a wild type uvrB$^+$ strain (see Fig. 3).

Localization of the *uvrB* promoter

In this paragraph we aim to determine more precisily the localization of the *uvrB* promoter within the PstI-EcoRI fragment (1,250 bp.) of plasmid pNP12 by subcloning DNA segments of this fragment. A search for restriction endonucleases, which cleave the PstI-EcoRI fragment, revealed that HaeIII has a single site at approximately 220 bp. from the EcoRI site (schematically given in Fig. 2). Removal of the PstI-HaeIII fragment (about 1,000 bp.) from pNP12 was accomplished, employing the following procedure. pNP12 DNA was linearized with the enzyme PstI and subsequently treated with the single strand specific nuclease S1 from *Aspergillus oryzae* (29) to remove protruding ends. This DNA was then digested with EcoRI and the large fragment, containing a cohesive EcoRI end and a 'blunt' end was purified and used as a vector. Plasmid pNP12 was also digested with both HaeIII and EcoRI, yielding 'blunt' ended HaeIII fragments and two fragments with one HaeIII end and one EcoRI cohesive end. Ligation of the constructed vector and the mixture of HaeIII- and HaeIII-EcoRI fragments, followed by transformation of strain HP3435 (Δ*uvrB*) and

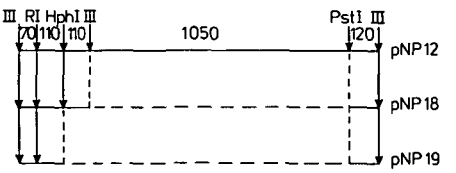

Figure 2. Schematical representation of part of the structure of plasmids pNP12, pNP18 and pNP19. Distances between relevant restriction sites are indicated. The cleavage sites for the enzymes EcoRI and HaeIII are given by the abbreviations RI and III. The horizontal dotted line indicates the absence of either the PstI-HaeIII fragment on pNP18 or the PstI-HphI fragment on pNP19, relative to plasmid pNP12.

selection for Km^r Uv^r clones,yielded plasmid pNP18 (illustrated in Fig. 2). A comparison of the HaeIII restriction patterns of pNP12 and pNP18 confirmed that a PstI-HaeIII fragment has been removed, resulting in the formation of plasmid pNP18.

Another derivative of pNP12 was constructed which lacks a more extended DNA segment from the PstI-EcoRI fragment than plasmid pNP18. For that purpose we employed a similar procedure as outlined for the construction of pNP18. It was found that the HaeIII-EcoRI fragment (220 bp.), located within the PstI-EcoRI fragment of pNP12, contains a single site for the enzyme HphI. Digestion of pNP12 DNA with HphI was subsequently followed by incubation with S1 nuclease to remove protruding ends and finally with EcoRI. This mixture of DNA fragments was ligated with linear 'vector' pNP12 DNA, containing one 'blunt' end (PstI terminus treated with S1 nuclease) and one cohesive EcoRI end. The resulting plasmid, denoted pNP19, is similar to pNP18, but lacks the PstI-HphI fragment (about 1,100 bp.) of pNP12. The composition of plasmid pNP19 was verified by restriction enzyme analysis. Transformants of strain HP3435 (Δ*uvrB*), carrying pNP19, display a Km^r Uv^r phenotype.

To assure that plasmids pNP18 and pNP19 have retained the intact regulatory elements of the *uvrB* gene we applied three criteria. i) These plasmids should provide for the same level of UV resistance in an UvrB deletion strain as pNP12. ii) Plasmids pNP18 and pNP19 should specifically encode a protein with a molecular weight of about 80,000 dalton which has been identified previously to represent the *uvrB* gene product, encoded by plasmids pNP10 and pNP12 (3). iii) The amount of UvrB protein encoded by plasmids pNP18 and pNP19 should be similar to the amount programmed by pNP12.

The UV survival of bacteria irradiated with various UV doses was determined for strain HP3435 (Δ*uvrB*) harbouring either pNP12, pNP18 or pNP19 (Fig. 3). It is clear that the UV resistance of *uvrB* mutant strains carrying either one of the plasmids does not significantly differ from that of a wild type $uvrB^+$ strain. Hence, according to this criterion we conclude that the intact regulatory region of the *uvrB* gene is present on these plasmids. It should be noted, however, that this observation is only a rough indication for the presence of the *uvrB* promoter, since substantial variations in the amount of UvrB protein may not affect the level of UV resistance.

To verify whether plasmids pNP18 and pNP19 encode the UvrB protein (MW of 80,000 dalton) we performed a 'maxi-cell' experiment as outlined before (3, 31)(Fig. 4). The data show that both plasmids indeed code for the UvrB protein Consequently, the translational regulatory elements are present on both pNP18

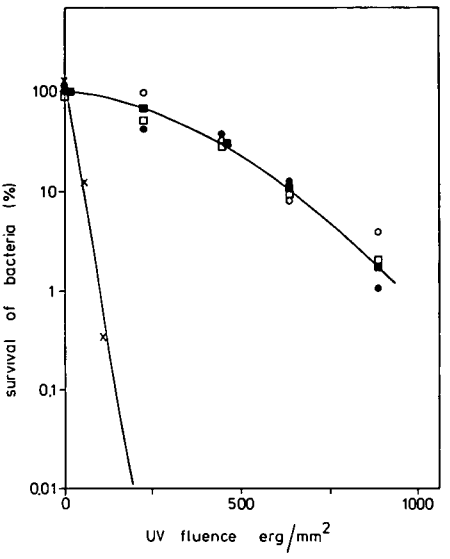

Figure 3. UV survival of strain HP3435 (Δ*uvrB*) transformed with either pNP12, pNP18 or pNP19 DNA. The strains were grown in L-broth supplemented with 50 µg/ml kanamycin (the antibiotic was omitted for strains XA106 and HP3435) until a density of 5 x 10^8 cells per ml was reached. After collection of the bacteria by centrifugation, they were washed twice with minimal medium (30). The cells were diluted in minimal medium, irradiated with various UV doses and dilutions were plated on L-broth with and without kanamycin. Symbols: O, HP3435/pNP12; ●, HP3435/pNP18; □, HP3435/pNP19; ■, XA106 ($uvrB^+$); X, HP3435.

and pNP19. The amount of UvrB protein, relative to the protein providing for kanamycin resistance (Km^r) (MW of about 28,000 dalton), synthesized on pNP12 is similar to the amount made on pNP18. However, we have observed consistently in 'maxi-cell' experiments that lower amounts of the UvrB protein are synthesized on plasmid pNP19 (lane 3 of Fig. 4). This result can be explained by assuming that the DNA segment, between the HaeIII- and the HphI site on the PstI-EcoRI fragment of pNP12, present on pNP18, but not on pNP19, is essential for the efficiency of *uvrB* transcription. The experiments presented in the proceeding sections will indicate the function of the HaeIII-HphI fragment in the promotion of transcription of the *uvrB* gene.

DNA sequence of the *uvrB* regulatory elements

We have shown that the *uvrB* regulatory region is located on a HaeIII-EcoRI fragment (220 bp.) of plasmid pNP12. To establish the structure of this

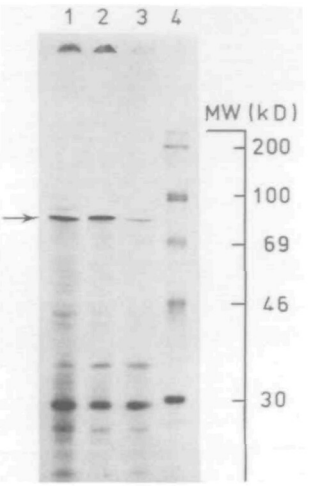

Figure 4. Plasmid encoded protein synthesis in maxi-cells. The procedure for the preparation of maxi-cells of strain SR362, transformed with various plasmids, has been outlined in the section Materials and Methods. Labelling of 1.6 x 10^9 cells in 4 ml M9 minimal medium, supplemented with 1% amino-acids, 1% glucose and thymine, was done for 150 min at 32°C with 30 µCi $|^{35}S|$-methionine. The cells were collected by centrifugation and lysed by heat-treatment for 2 min at 100°C in 1% SDS, 5% mercaptoethanol and 10% (v/v) glycerol. Samples, containing 9 x 10^4 cpm of acid-precipitable material, were loaded on a 10% polyacrylamide slabgel, containing 0.1% SDS. After electrophoresis the labelled protein bands were visualized by fluorography. Lane 1: proteins encoded by pNP12 DNA. Lane 2: proteins encoded by pNP18 DNA. Lane 3: proteins encoded by pNP19 DNA. Lane 4: $|^{14}C|$-labelled marker proteins with a molecular weight of, respectively, 200,000, 100,000, 92,500, 69,000, 46,000 and 30,000 daltons. The arrow indicates the position of the UvrB protein (MW about 80,000 dalton).

region the DNA sequence of the HaeIII-HaeIII fragment (290 bp.), containing the EcoRI site, was determined employing the chemical degradation method (11). The sequence strategy, showing the radioactive labelling of the 5' ends of both HaeIII sites, the EcoRI site and one of the TaqI sites, is outlined in Fig. 5A. The actual DNA sequence of most of the HaeIII-HaeIII fragment, namely the part for which both strands have been sequenced, is presented in Fig. 5B.

The righthand part of the nucleotide sequence contains an ATG codon (at position 133-135) followed by an open reading frame to the end of the fragment. The other two reading frames contain one or more non-sense codons and lack the proper initiation signal. In front of, and in phase with the ATG codon, is a TAG stopcodon. At a distance of eight nucleotides upstream of the ATG codon the codogenic strand harbours the sequence 5' AGGT. The correspond-

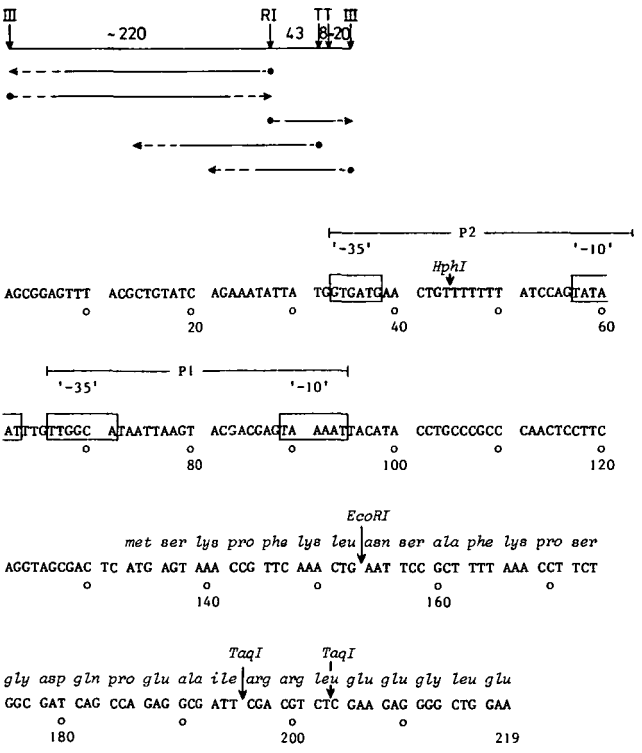

Figure 5A. Restriction map and sequence strategy for the HaeIII fragment of plasmid pNP12, containing the *uvrB* regulatory elements. Restriction sites are given for the enzymes HaeIII, EcoRI and TaqI which are represented by the symbols: III, RI and T. The arrows show the sequence runs, while dotted regions are parts of the DNA sequence which are not firmly established.

Figure 5B. DNA sequence of a part of the HaeIII fragment of plasmid pNP12, containing the *uvrB* regulatory elements. DNA sequencing of both strands was performed using the chemical degradation method (11). For ease of survey only the nucleotide sequence of one strand, which is codogenic for both *bioD* and *uvrB*, is given. The regions which can serve as promoters for the *uvrB* gene, are indicated below the sequence. The putative '-10 regions' extend from, respectively, position 57 till 62 and position 89 till 94. The '-35 regions' extend from, respectively, position 33 till 38 and position 66 till 71. The DNA sequence for these regions have been compared with the model sequence, proposed by Rosenberg and Court (1), notably 5'TTGACA for the '-35 region' and 5'TATAAT for the '-10 region'. Cleavage sites for restriction endonucleases within this DNA sequence are given.

ing RNA sequence 5' AGGU is complementary to the 3' end of 16S ribosomal RNA. Interactions between mRNA and the ribosome, based on nucleic acid complementarity, are believed to be implicated in initiation of translation (32). We therefore believe that the ATG codon represents the initiation triplet of the *uvrB* gene which is transcribed towards the EcoRI site as shown previously.

The promoters of the *uvrB* gene

In general, *E.coli* promoters exhibit a striking similarity in their DNA sequence. A statistical analysis of the nucleotide sequences of 46 promoters has been carried out, resulting in a 'model' sequence for *E.coli* promoters (1) which is very similar to the one derived by computer analysis (33). The model sequence consists of a RNA polymerase binding site ('-10 region'; ref. 34) and a RNA polymerase recognition site ('-35 region; ref. 35). The model sequence, which is given in the legend to Fig. 5B, matches two regions within the stretch of DNA between the beginning of the *uvrB* gene and the left end of the presented DNA sequence. The two possible promoters are: i) 5' GTGATG ('-35 region' from 33 till 38) and TATAAT ('-10 region' from 57 till 62). This promoter is designated P2. ii) 5' TTGGCA ('-35 region' from 66 till 71) and TAAAAT ('-10 region' from 89 till 94). This promoter is designated P1.

A comparison of the properties of plasmids pNP18 and pNP19 has shown that the removal of the HaeIII-HphI fragment (from the left side till position 45) which is present on pNP18, but not on pNP19, causes a pronounced effect on the extent of UvrB protein synthesis. Based on theoretical considerations this removal will eliminate the '-35 region' of promoter P2 and would abolish its functioning. Hence, we propose that both promoters are essential for transcription of the *uvrB* gene.

Support for our proposal on the existence and the localization of the *uvrB* promoters is provided by the data on length measurements of 'run-off' transcripts, using restriction fragments which contain the regulatory elements. *In vitro* transcription with purified RNA polymerase and either the HaeIII-HaeIII fragment (290 bp.) or the HaeIII-EcoRI fragment (220 bp.) as template was monitored by polyacrylamide-urea gelelectrophoresis (Fig. 6). Major transcripts are synthesized on the HaeIII-HaeIII fragment with a length of approximately 50, 130, 160, 220 nucleotides and 'end-to-end' transcripts which are often found in *in vitro* transcription systems, using restriction fragments as template. To establish the orientation of these chains, we also performed 'run-off' transcription on the HaeIII-EcoRI fragment of pNP12, which was shortened at one side by about 70 bp. relative to the HaeIII-HaeIII fragment. On this template major RNA chains are synthesized with a length of about

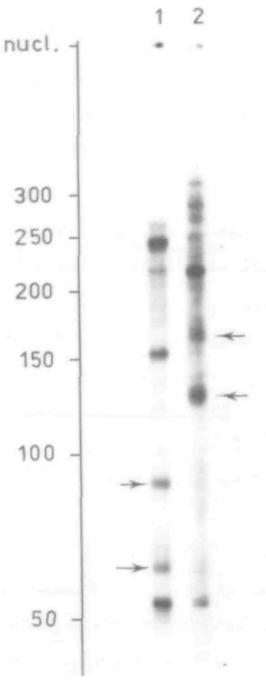

Figure 6. Analysis of *in vitro* transcripts synthesized on restriction fragments, containing the *uvrB* regulatory elements. 'Run-off' transcription was carried out on either the HaeIII fragment (290 bp.) or on the HaeIII-EcoRI fragment (220 bp.) of pNP12 DNA as a template. Conditions for *in vitro* RNA synthesis are as outlined in the section Materials and Methods. RNA, labelled with α-$|^{32}P|$-UTP, was analyzed on a 8% polyacrylamide, 7 M urea slabgel and the bands are visualized by autoradiography. $|^{32}P|$-labelled length markers were a 104 nucleotides transcript, synthesized *in vitro* on ColE1 DNA as a template (36) and a 49 nucleotides RNA fragment, produced by digestion of $|^{32}P|$-labelled 16S ribosomal *E.coli* RNA with cloacin DF13 (37). Furthermore, we used denatured HaeIII restriction fragments of $|^{32}P|$-labelled phage M13 with a length of 313, 311, 309, 158, 142, 106 and 69 nucleotides (longer fragments are not indicated) (38). Lane 1: the HaeIII-EcoRI fragment (220 bp.) of pNP12 DNA used as template. The arrows indicate the position of the 60 nucleotides and the 90 nucleotides transcripts. Lane 2: the HaeIII-HaeIII fragment (290 bp.) of pNP12 DNA used as a template. The arrows indicate the position of the 130 nucleotides and the 160 nucleotides transcripts.

50, 60, 90, 150 nucleotides and 'end-to-end' transcripts. The RNA chains of 130 and 160 nucleotides, made on the HaeIII-HaeIII fragment, match the DNA segments between the right-ended HaeIII-terminus and the two '-10 regions' at, respectively, 57-62 and at 89-94. Likewise, the transcripts of 60 and 90

nucleotides, synthesized on the HaeIII-EcoRI fragment, match the DNA segments between the EcoRI-terminus and the two '-10 regions'. From these data we estimate that the transcripts of 130 and 60 nucleotides are transcribed from left to right and are initiated at a few basepairs downstream of the proposed '-10 region' 5' TAAAAT (P1; 89-94), whereas the RNA chains of 160 and 90 nucleotides are initiated near the other '-10 region' 5' TATAAT (P2; 57-62) and synthesized in the same direction. We consider the transcripts of 220 nucleotides (on HaeIII-HaeIII DNA) and the corresponding chains of 150 nucleotides (on HaeIII-EcoRI DNA) as artifacts of the *in vitro* system, since the region preceeding the starting site of these chains does not reveal obvious similarities with the consensus sequences for the '-35' and the '-10 regions' (results not shown). Furthermore, in a following section evidence is presented which shows that *in vivo* no RNA molecules are initiated at this site (approximately at position 1 in Fig. 5B). Finally, we cannot assign an origin and an orientation to the short 50 nucleotides transcripts, since these molecules are synthesized on both templates. Hence, we conclude that *in vitro* RNA polymerase can utilize both promoters for the expression of the *uvrB* gene. The *in vivo* maxi-cell experiments, in which we have compared the amount of UvrB protein synthesized either on plasmid pNP18 or on plasmid pNP19, are consistent with the conclusion that both promoters contribute to the expression of the *uvrB* gene.

Several other cases have been reported in which tandem promoters contribute to the expression of an adjacent gene or an operon. Notably, transcription of DNA, specifying *E.coli* ribosomal RNA, is directed by two distinct promoters (39,40). Also, transcription of the *E.coli* galactose (*gal*) operon is mediated by two promoters (41). Interestingly, the activity of the *gal* promoters is regulated quite differently. One promoter is dependent on the presence of cyclic AMP and cyclic AMP receptor protein (CRP), whereas the other is not and maintains a constant basal level of *gal*-mRNA. Hence, tandem promoters may provide for separate sites which serve different regulatory mechanisms. In the following sections we will discuss these possibilities for regulation of the expression of the *uvrB* gene.

Inverted repeat DNA sequences within the regulatory region of the *uvrB* gene.

In general, regulatory proteins which act at the level of transcription interact with a characteristic DNA sequence within the regulatory elements or with the RNA transcribed from it. We assume that this property also applies to the UvrC protein and the LexA protein which control the expression of the *uvrB* gene (2,3). The LexA protein is regarded as a repressor of the *recA* gene and

other genes as well (34,42,43,44). It is conceivable that this protein also functions as a repressor for the *uvrB* gene. The mechanism of action by which the UvrC protein reduces the synthesis of *uvrB*-mRNA is at present not known. Usually, however, negative control of transcription takes place either by repression or by termination of transcription. The first option often requires inverted repeat sequences within the regulatory region, whereas the second option might be caused by a transcriptional termination signal contained by the mRNA molecule.

To explore these possibilities we have subjected the DNA sequence of the *uvrB* regulatory region to a computer analysis, devised to search for inverted repeat sequences (45). The most striking inverted repeats, which are separated by a limited number of basepairs, occur within the elements involved in initiation of transcription and translation (see Fig. 5B). These observations can be summarized as follows: i) Several inverted repeats are found in and around the '-10 region' of promoter P1. A six basepair repeat is present, *i.e.* 5'AAATTA (91-96), while its counterpart 5'TAATTT (58-63) is located adjacent to the '-35 region' of P1. Furthermore, a four basepair repeat is observed, *i.e.* 5'GTAA (88-91) which is separated by only two basepairs from its counterpart 5'TTAC (94-97). ii) The region complementary to the 3' end of 16S ribosomal RNA (32) is part of a six basepair inverted repeat 5'CAGGTA (120-125), while the sequence 5'TACCTG (98-104) is present in the region which probably codes for the 5' end of *uvrB*-mRNA. If, the UvrC protein carries out a 'repressor-like' function then the occurrence of the inverted repeat sequences, might be of importance.

Furthermore, it should be noted that the DNA sequence encoding the 5' end of *uvrB*-mRNA, initiated at the P1 promoter, is extremely G-C rich. In this region (100-111) ten out of twelve basepairs are G-C's, a composition which will affect the melting properties of this segment during RNA synthesis (46) and may obstruct the progress of RNA polymerase molecules during RNA synthesis. If, the UvrC protein acts as a transcription termination factor, thereby reducing the expression of the *uvrB* gene, then this G-C rich region might be the site of interaction. At present, we are attempting to introduce localized alterations in the P1 promoter region. This approach may give us insight into the significance of the peculiar DNA sequences and their possible function in regulation of the expression of the *uvrB* gene by the UvrC protein.

Commonalities within DNA sequences of the regulatory regions of the *recA*, *lexA* and *uvrB* genes.

It has been proposed that the LexA protein is a repressor for a set of

genes which are induced upon treatment of the cell with DNA damaging agents (47). Such treatments would release an effector which in turn would activate a proteolytic property of the RecA protein. Cleavage of the *lexA* gene product by the RecA protein is considered to account for induction of many genes, *e.g.* *uvrA* *lexA* and *recA* (7,8,42,43,44). Recently it was indicated that the *uvrB* gene is also subject to regulation by the *lexA-recA* control circuit (2). It is predictable that the regulatory regions of genes, which are repressed by the *lexA* gene product, contain commonalities in their DNA sequences implicated in an interaction with this protein. For that purpose we have searched for common DNA sequences, using the reported DNA sequences of the regulatory regions of the *recA* gene (48,49), the *lexA* gene (50) and the DNA sequence of *uvrB*. The results are presented in Fig. 7.

The promoter region of the *recA* gene, which contains a 10 basepair inverted repeat, harbours DNA segments common to both the regulatory region of the *lexA* gene and the '-10 region' of the P2 promoter of the *uvrB* gene. These commonalities include the sequences: 5'CTGT, T, 5'CAG, 5'ATAA, T. Furthermore, it should be noted that the DNA sequences around and within the '-10 regions' of the *recA* gene and the *uvrB* gene are almost identical (5'AT$^{A}_{C}$CAGTATAATT). Finally, we have indicated that the inverted repeats, present in all three DNA sequences, are centered at a corresponding point. We consider this observation to be suggestive for a site of interaction of the LexA protein.

If our assumption is correct then it can be predicted that *in vivo*, using a strain with active LexA protein, transcription of the *uvrB* gene is prevented

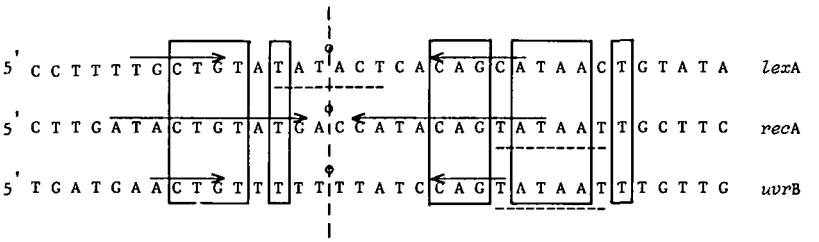

Figure 7. Comparison of the DNA sequences of regulatory regions preceding the *lexA*, *recA* and *uvrB* genes. The DNA sequence of the *recA* gene was taken from Horii *et al.* (48) and Sancar *et al.* (49), whereas the *lexA* DNA sequence has been reported by Miki *et al.* (50). The arrows indicate the position of inverted repeat sequences. The dotted horizontal lines represent the '-10 regions' of, respectively, the *lexA*, *recA* and *uvrB* genes. The boxed areas indicate the position of common basepairs in all three DNA sequences. The dotted vertical line represents the axis of symmetry of the inverted repeats within the DNA sequences.

from promoter P2 and that *uvrB*-mRNA synthesis is mediated only by promoter P1. To test this prediction we performed a S1 nuclease protection experiment, using *in vivo* made RNA from an UvrB deletion strain harbouring plasmid pNP12 or from the plasmidless UvrB strain and uniformily labelled $|^{32}P|$-HaeIII-

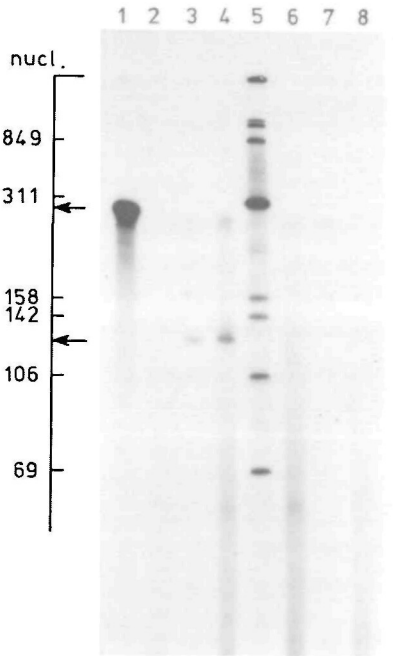

Figure 8. S1 nuclease protection of $|^{32}P|$-HaeIII DNA (290 bp.) containing the *uvrB* promoters by *in vivo* made RNA. Isolation of RNA from strains HP3435 and HP3435/pNP12, DNA-RNA hybridization with the $|^{32}P|$-labelled HaeIII fragment of pNP12 and S1 nuclease incubations are outlined in Materials and Methods. Samples in 70% (v/v) formamide were analyzed by electrophoresis on a 8% polyacrylamide, 7 M urea slab gel. Lane 1: no RNA added (intact DNA fragment of about 290 nucleotides). Lane 2: DNase treatment of the *in vivo* made nucleic acid preparation of strain HP3435/pNP12 omitted. Lane 3: protection with RNA from strain HP3435/pNP12 during incubation with 20 units of S1 nuclease per ml. Lane 4: as for lane 3, but with 10 units of S1 nuclease per ml. Lane 5: marker fragments of M13 RF DNA, labelled with $|^{32}P|$ by 'nick translation' (13) and digested with HaeIII, yielding fragments of 69, 106, 142, 158, a triplet of 309, 311 and 313, 849, 1623 and 2527 nucleotides (38). Lane 6: RNA from strain HP3435 devoid of pNP12 and S1 nuclease incubation with 10 units of enzyme per ml. Lane 7: as for lane 6, but with 20 units of S1 nuclease per ml. Lane 8: DNase treatment of the *in vivo* made nucleic acid preparation of strain HP3435 omitted. The arrows indicate the position of the intact HaeIII-HaeIII fragment (290 nucleotides) and of the protected fragment (130 nucleotides).

HaeIII DNA containing both promoter P1 and P2. The results are given in Fig. 8. Clearly, protection occurred when RNA was employed from the strain harbouring plasmid pNP12, whereas no protection was detected with RNA from the strain devoid of plasmid DNA. The length of the DNA fragment which is resistant against S1 nuclease digestion is about 130 nucleotides. This value agrees well with the distance between the '-10 region' of promoter P1 and the right-ended HaeIII-terminus and with the length of one of the *in vitro* transcripts (see Fig. 6). These results indicate that, under these conditions, *in vivo* transcription is mediated only by promoter P1. Apparently, no transcription is initiated from P2, presumably since this site is blocked by LexA protein. Furthermore, there are also no transcripts initiated upstream of P2, as are found during *in vitro* transcription.

In conclusion, we favor the following working hypothesis for an interaction of the LexA and the UvrC proteins with the regulatory region of the *uvrB* gene. Based on both our *in vivo* and *in vitro* data reported here, we propose that the *uvrB* gene is preceded by two distinct promoters. This property allows the possibility of separate interaction sites for regulatory proteins which will overlap with either one of the promoters. We have argued in the preceding sections that the *lexA* gene product probably interacts with the '-10 region' of the promoter (P2) distal to the N-terminus of *uvrB*, while the UvrC protein might affect transcription initiated near the '-10 region' (P1) proximal to the beginning of the *uvrB* gene. This proposal is consistent with our recent finding that the UvrC protein also reduces the biosynthesis of the UvrB protein, programmed by plasmid pNP19 which lacks the promoter distal to the N-terminus of the *uvrB* gene (J. Zwetsloot and H. Pannekoek, manuscript in preparation).

Acknowledgements

We appreciate the critical reading of the manuscript by Peter H. Pouwels and we thank Mrs. Nicoline van Hoek for typing of this paper. This work was supported by Euratom contract 102-72-1 BIAN.

†Correspondence to: H.Pannekoek, Biochemical Laboratory, Department of Molecular Genetics, Wassenaarseweg 64, 2333 AL Leiden, The Netherlands

REFERENCES

1. Rosenberg, M. and Court, D. (1979) Ann. Rev. Genet. 13, 319-353.
2. Fogliano, M. and Schendel, P.F. (1981) Nature 289, 196-198.

3. Pannekoek, H., van Sluis, C.A., Brandsma, J.A., Dubbeld, D. and Noordermeer, I.A. (1981) Nature (submitted).
4. Howard-Flanders, P., Boyce, R.P. and Theriot, L. (1966) Genetics 53, 1119-1136.
5. Hanawalt, P.C., Cooper, P.K., Ganesan, A.K. and Smith, C.A. (1979) Ann. Rev. Biochem. 48, 783-836.
6. Seeberg, E. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 2569-2573.
7. Kenyon, C.J. and Walker, G.C. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 2819-2823.
8. Kenyon, C.J. and Walker, G.C. (1981) Nature 289, 808-810.
9. Pannekoek, H., Noordermeer, I. and van de Putte, P. (1979) J. Bacteriol. 139, 48-53.
10. Pannekoek, H., Noordermeer, I.A., van Sluis, C.A. and van de Putte, P. (1978) J. Bacteriol. 133, 884-896.
11. Maxam, A. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 560-564.
12. Elzen, P.J.M. van den, Konings, R.N.H., Veltkamp, E. and Nijkamp, H.J.J. (1980) J. Bacteriol. 144, 579-591.
13. Maniatis, T., Jeffry, A. and Kleid, D.G. (1975) Proc. Natl. Acad. Sci. U.S.A. 72, 1184-1188.
14. Berk, A.J. and Sharp, P.A. (1977) Cell 12, 721-732.
15. Burgess, R.R. and Jendrisak, J.J. (1975) Biochemistry 14, 4634-4638.
16. Burgess, R.R. (1969) J. Biol. Chem. 244, 6160-6167.
17. Sternbach, H., Engelhardt, R. and Lezius, A.G. (1975) Eur. J. Biochem. 60, 51-55.
18. Ormondt, R. van, Maat, J., de Waard, A. and van der Eb, A.J. (1978) Gene 4, 309-328.
19. Pannekoek, H., Noordermeer, I. and van de Putte, P. (1979) J. Bacteriol. 139, 54-63.
20. Pannekoek, H., Hille, J. and Noordermeer, I. (1980) Gene 12, 51-61.
21. Rodriguez, R.L., Bolivar, F., Goodman, H.M., Boyer, H.W. and Betlach, M.C. in D.P. Nierlich, W.J. Rutter and C.F. Fox (eds.), Molecular Mechanisms in the control of gene expression, Academic Press, New York, 1976, pp. 471-477.
22. Rupp, W.D., Sancar, A., Kennedy, W.J., Ayers, J. and Griswold, J. in P. C. Hanawalt, E.C. Friedberg and C.F. Fox (eds.), DNA repair mechanisms ICN-UCLA symp. on Molecular and Cellular Biology, Vol. IX, Academic Press, New York, 1978, pp. 229-235.
23. Clark, L. and Carbon, J. (1976) Cell 9, 91-99.
24. Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H. L. Boyer, H.W., Crosa, J.H. and Falkow, S. (1977) Gene 2, 95-113.
25. Chang, A.C.Y. and Cohen, S.N. (1978) J. Bacteriol. 134, 1141-1156.
26. Sutcliffe, J.G. (1978) Cold Spring Harbor Symp. Quant. Biol. 43, 77-90.
27. Berg, D.E. in A.I. Bukhari, J.A. Shapiro and S.L. Adhya (eds.), DNA insertion elements, plasmids and episomes, Cold Spring Harbor Laboratory New York, 1977, pp. 205-212.
28. Pannekoek, H., Maat, J., van den Berg, E. and Noordermeer, I. (1980) Nucleic Acid Res. 8, 1535-1550.
29. Vogt, V.M. (1973) Eur. J. Biochem. 33, 192-200.
30. Vogel, H. and Bonner, D. (1956) J. Biol. Chem. 218, 97-106.
31. Sancar, A., Hack, A.M. and Rupp, W.D. (1979) J. Bacteriol. 137, 692-693.
32. Shine, J. and Dalgarno, L. (1975) Nature 254, 34-38.
33. Scherer, G.E.F., Walkinshaw, M.D. and Arnott, S. (1978) Nucleic Acid Res. 5, 3759-3773.
34. Pribnow, D. (1975) Proc. Natl. Acad. Sci. U.S.A. 72, 784-788.
35. Gilbert, W. in R. Losick and M. Chamberlin (eds.), RNA polymerase, Cold Spring Harbor Laboratory, New York, 1976, pp. 193-205.

36. Chan, P.T., Lebowitz, J. and Bastia, D. (1979) Nucleic Acid Res. 7, 1247-1262.
37. Charldorp, R. van, Heus, H.A. and van Knippenberg, P.H. (1981) Nucleic Acid Res. 9, 267-275.
38. Wezenbeek, P.M.G.F. van, Hulsebos, T.J.M. and Schoenmakers, J.G.G. (1980) Gene 11, 129-148.
39. De Boer, H.A., Gilbert, S.F. and Nomura, M. (1979) Cell 17, 201-209.
40. Young, R.A. and Steitz, J.A. (1979) Cell 17, 225-234.
41. Musso, R.E., Di Lauro, R., Adhya, S. and de Crombrugghe, B. (1977) Cell 12, 847-854.
42. Mount, D.W. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 300-304.
43. Little, J.W. and Harper, J.E. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 6147-6151.
44. Brent, R. and Ptashne, M. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 1932-1936.
45. Staden, R. (1978) Nucleic Acid Res. 5, 1013-1015.
46. Saucier, J.M. and Wang, J.C. (1972) Nature New Biol. 239, 167-170.
47. Witkin, E.M. (1976) Bacteriol. Rev. 40, 867-907.
48. Horii, H., Ogawa, T. and Ogawa, H. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 313-317.
49. Sancar, A., Stachelek, C., Koningsberg, W. and Rupp, W.D. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 2611-2615.
50. Miki, T., Ebina, Y., Kishi, F. and Nakazawa, A. (1981) Nucleic Acid Res. 9, 529-543.