Volume 17 Number 17 1989 Nucleic Acids Research

The *recR* locus of *Escherichia coli* K-12: molecular cloning, DNA sequencing and identification of the gene product

Akeel A.Mahdi and Robert G.Lloyd

Department of Genetics, University of Nottingham, Queens Medical Centre, Nottingham NG7 2UH, UK

Received August 2, 1989; Revised and Accepted August 8, 1989 EMBL accession no. X15761

ABSTRACT

The *recR* gene of *Escherichia coli*, which is associated with *recBC*-independent mechanisms of recombination and DNA repair, has been located between *dnaZX* and *htpG* on a 6.4 kb *Eco*RI fragment of DNA that has been cloned and analysed in λ and plasmid vectors. Nucleotide sequencing of this interval revealed two open reading frames that constitute an operon lying immediately downstream of *dnaZX*. The second of these two reading frames was identified as *recR*. It encodes a polypeptide with a predicted molecular weight of 21,965 Daltons that migrates on SDS gels as a 26 kDa protein. The first gene of the operon encodes a polypeptide of 12,015 Daltons. Its function is not known.

INTRODUCTION

recR is the latest addition to the complement of genes associated specifically with homologous recombination and DNA repair in *Escherichia coli*. We initially identified the locus in a *recB sbcB sbcC* genetic background by means of transposon insertions that were found to increase sensitivity to mitomycin C and UV light and to reduce the yield of recombinants recovered from conjugational and transductional crosses (1). This phenotype was subsequently extended to the *recB recC sbcA* background where recombination proceeds by a related *recBC*-independent mechanism. As with other genes (*recF*, *recJ*, *recN*, *recO*, *recQ* and *ruv*) needed for recombination in these *recBC*-deficient backgrounds (2,3), mutation of *recR* does not prevent conjugational or transductional recombination in $recBC^+$ sbc^+ strains, though plasmid recombination is largely abolished and there is a modest sensitivity to UV light and mitomycin C. From these observations and studies of interactions with other *rec* loci, we concluded that *recR* is involved in a *recBC*-independent recombinational process of DNA repair that relies also on *recF*, *recO* and possibly *recJ* (1).

Genetic crosses located *recR* very close to *dnaZX* at minute 11 on the standard genetic map (1). A further understanding of its role in DNA repair and recombination will require a molecular analysis of the gene and of its product. In this paper we describe the molecular cloning and DNA sequencing of *recR*, its relationship to the flanking *dnaZX* and *htpG* loci, and the identification of its product.

METHODS

Strains, plasmids and λ phages

E. coli K-12 strains are listed in Table 1. *recR252* is an insertion of a 'mini-*kan*' derivative of Tn*10*. pBR322 (7), pJRD184 (8) and pUC18 (9) were used as general cloning vectors. pHSG415 (10) was used as a low-copy-number vector. Since it exhibits temperature-sensitive replication, strains carrying this plasmid or its derivatives were grown at 30°C.

Table 1. *E.coli* K-12 strains.

Strain	Relevant features[1]	Source and/or Reference
AB1157	*rec*$^+$ *sbc*$^+$	(4)
AB2463	*recA13*	(4)
AB2480	*uvrA6 recA13*	S. Sedgwick (5)
JC7623	*recB21 recC22 sbcB15 sbcC201*	A. J. Clark (4)
JC8679	*recB21 recC22 sbcA23*	A.J. Clark (1)
AX727	*dnaZ2016* (T.S. growth)	B.J. Bachmann (1)
N2435	*rec*$^+$	(6)
N2477	*recN262*	(6)
AM134	*recB21 sbcB15 sbcC207 recR252*	(1)
AM158	as AM134 but (λcI857)$^+$	This work
AM197	*recB21 recC22 sbcB15 sbcC201 recR252*	(1)
AM207	*recR252 rpsL31* (Str^r)	(1)
AM265	*recB21 recC22 sbcA23 recR252*	(1)
AM305	*recN262 recR252*	(1)
NH4104	F42 *lacZ*$^+$ / *lac-1 rpsL*$^+$ (Str^s)	K.B.Low (5)
JM109	F128 *proAB*$^+$ *lacI*^{q} ZΔ*M15 traD36* / *thi-1 recA1* Δ(*lac-pro*) *endA1 gyrA96 hsdR17 supE44*	(9)

[1]The full genotype is listed in the reference cited.

Other plasmids constructed during this study are described in Fig. 1. The λ library screened for *recR* transducing phages was constructed by ligating a partial *Sau*3A digest of *E. coli* DNA into *Bam*HI-digested DNA from λPE11, a *c*I857 derivative of the λL47 cloning vector (11). The library was a gift from D. Bramhill and P. T. Emmerson. The phage sequencing vectors M13mp18 and M13mp19 and their recombinant derivatives were maintained in strain JM109 (9).

Media and general methods

LB broth and agar have been cited elsewhere along with methods for measuring sensitivity to UV light and mitomycin C and for conducting matings in liquid medium (12). Strains harbouring antibiotic resistance plasmids were grown in media supplemented with 20 μg/ml tetracycline (Tc), 50 μg/ml ampicillin (Ap), 40 μg/ml kanamycin (Km), or 25 μg/ml chloramphenicol (Cm), as appropriate. Media, strains and methods for propagating λ, constructing λ lysogens, and detecting recombinants of pUC18 have been described by Silhavy et al. (13) and Maniatis et al. (14).

Enzymes and DNA analysis

Restriction endonucleases, T4 DNA ligase, and Klenow polymerase were purchased from commercial sources and used as directed by the suppliers. Plasmid stocks were prepared from *recA* strain AB2463 by the rapid alkaline-sodium dodecyl sulphate lysis procedure of Ish-Horowicz and Burke (15). λ DNA was prepared by the method of Silhavy et al. (13). Procedures for analysis of restricted DNA by agarose gel electrophoresis, purification of DNA fragments by electroelution, the filling in of recessed 3′ DNA ends, DNA ligation and transformation followed recipes and protocols described by Maniatis et al. (14).

DNA sequencing

Overlapping DNA fragments from the region to be sequenced were inserted into the multiple cloning site of pUC18 from where they were then directed into M13mp18 and M13mp19. The DNA inserts were then sequenced by the dideoxy chain termination method of Sanger et al. (16) using kits from Pharmacia-LKB containing phage T7 DNA polymerase,

deoxynucleoside and dideoxynucleoside triphosphates, and M13 universal sequencing primer (17mer). [^{35}S]-dATP was obtained from Amersham International. Sequencing reactions containing 7-deaza dGTP or 7-deaza dITP instead of dGTP were used to help resolve any compressions in GC-rich regions. Except for the first 269 bp shown in Figure 5, the sequence determined was read from both strands and was compiled and analysed using microcomputer software packages from DNASTAR Ltd.

Transposon mutagenesis

pAM107 was transformed into the F42*lac* strain NH4104 and then mobilised to the *recR252* strain AM207 by mating for 2 h at 37°C. Selection was imposed for Tc^r(*rpsL*) transconjugant colonies and individual isolates were tested for sensitivity to mitomycin C and UV light. Sensitive isolates were assumed to carry Tn*1000* insertions in pAM107 that prevented expression of *recR*. Plasmid DNA samples extracted from two *recR*$^-$ and two *recR*$^+$ isolates from each of six independent matings were analysed by *Eco*RI, *Bam*HI and *Kpn*I single digests to confirm the presence of the 5.7 kb Tn*1000* insertion and to determine its location.

Identification of plasmid-encoded proteins

Plasmid encoded proteins were labelled with [^{35}S]methionine (Amersham) in strain AB2480 using the maxicell method (17). Proteins were separated by SDS-PAGE (18) and visualized by fluorography. The [^{14}C]methylated protein mixture used to provide molecular weight standards contained myosin (200,000), phosphorylase-b (92,500), bovine serum albumin (69,000), ovalbumin (46,000), carbonic anhydrase (30,000) and lysozyme (14,300), and was obtained from Amersham International.

Isolation of recR$^+$ transducing phages

The lysogenic *recB sbcBC recR252* strain AM158 was infected with the λPE11 library of *E. coli* DNA at a multiplicity of infection of about 0.1. Transductants resistant to 0.3 μg mitomycin C per ml in LB agar were selected at 30°C, purified, and heat-shocked to induce the prophages. The lysates produced were streaked against AM158 on LB agar supplemented with 0.2 μg mitomycin C per ml and exposed to 30 J/m^2 UV light before incubating at 30°C overnight. Lysates that gave growth of resistant cells were assumed to contain *recR*$^+$ transducing phages. *recR*$^+$ phages purified from 15 independent transductants were streaked against a λ*c*$^+$ lysogen of the temperature sensitive *dnaZ* mutant, AX727. All 15 allowed growth of the *dnaZ* strain at 42°C, which confirmed that they carried DNA inserts from the *dnaZ-recR* region of the chromosome. One of these phages, λAM234, was used as a source of DNA for sub-cloning *recR* into plasmid vectors.

RESULTS

Molecular cloning of recR

The cloning of *recR* into plasmid vectors is summarised in Fig. 1. A 6.4 kb *Eco*RI DNA fragment of λAM234 was first cloned into pHSG415. This low-copy-number construct, pAM100, conferred resistance to UV light in a *recR* single mutant as well as in *recBC sbcBC recR* and *recBC sbcA recR* strains (Fig. 2a−c). pAM102 carries the same fragment inserted into pBR322. It allows growth of the *dnaZ* mutant AX727 at 42°C, from which we concluded that the insert also carries *dnaZX*. However, it failed to restore UV resistance to a *recR* single mutant and increased the sensitivity of a *rec*$^+$ strain (Table 2a and data not shown). We had shown before (1) that a *recN recR* double mutant is much more sensitive to UV light than a *recR* single mutant, and is highly sensitive when compared with the relatively resistant *recN* mutant (see Table 2). The data in Table 2a show that introduction

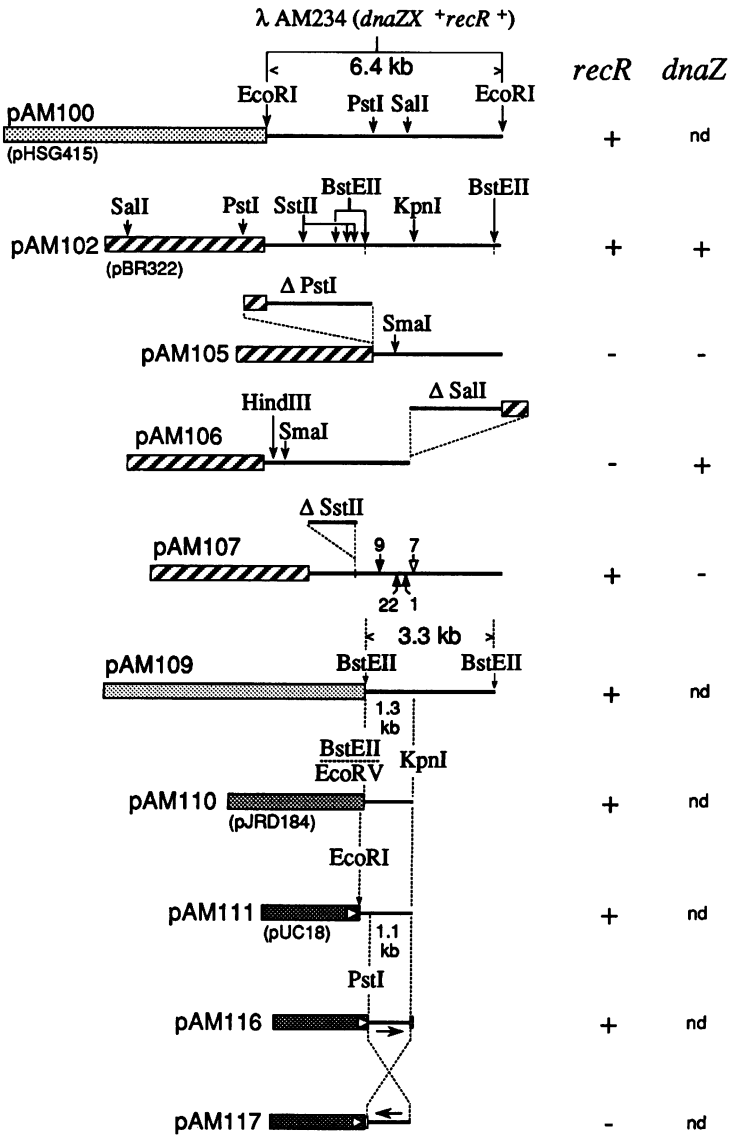

Figure 1. Summary of the construction, restriction mapping, and functional analysis of recombinant plasmids carrying DNA fragments from the *dnaZX*$^+$ *recR*$^+$ phage, λAM234. Vectors are shown as shaded boxes and are identified on first use. Chromosomal inserts are represented by solid lines and are aligned with respect to the *Pst*I site. Restriction sites are shown by vertical arrows. Except for the *Sal*I and *Pst*I sites in pAM102, restriction sites in the vector DNA are limited to those used for insertion of the cloned DNA. pAM105, pAM106 and pAM107 are deletions from pAM102. The locations of Tn*1000* insertions in pAM107 are shown by numbered open (*recR*$^+$) or closed (*recR*) triangles. Open triangles within the pUC18 constructs indicate the location and orientation of the *lac* promoter, while the short open boxes in pAM116 and pAM117 represent the inverted duplications of the *Kpn*I-*Pst*I region of the multiple cloning site. The ability or inability of the plasmid constructs to express *recR* and/or *dnaZ* is indicated on the right by + and − respectively (n.d. = not determined).

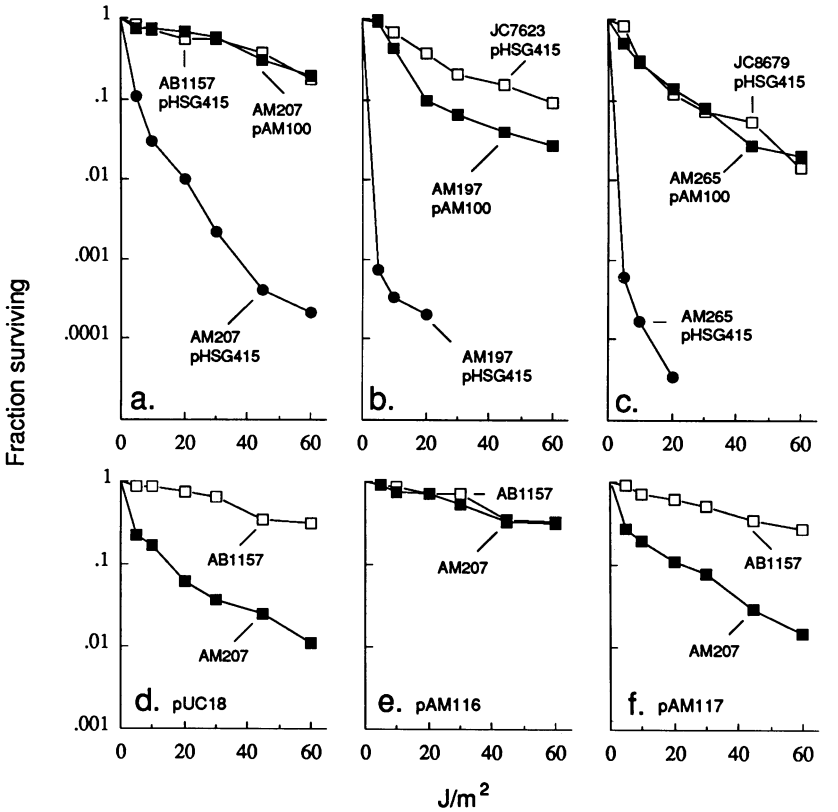

Figure 2. Effect of *recR* plasmids on the survival of UV-irradiated strains. The plasmids constructs, vector controls and strains used are identified within each panel.

of pAM102 into the *rec*$^+$, *recN* and *recN recR* backgrounds produces strains that show a very similar modest sensitivity to UV light but which are considerably more resistant than the *recN recR* strain carrying the vector plasmid. We concluded that *recR* is expressed from pAM102 but that the increased copy-number of some other factor masks the ability to restore UV resistance to a *recR* single mutant.

Table 2. Effect of *recR* plasmids on the survival of *rec*$^+$, *recN* and *recN recR* strains.

Plasmid	Fraction surviving (UV dose = 30 J/m^2) N2435 (*rec*$^+$)	N2477 (*recN*)	AM305 (*recN recR*)
a. pBR322	0.31	0.097	0.00012
pAM102	0.0027	0.0026	0.0057
b. pJRD184	0.50	0.22	0.00042
pAM110	0.30	0.25	0.20
c. pUC18	0.34	0.15	0.000072
pAM116	0.22	0.092	0.11
pAM117	0.20	0.11	0.000045

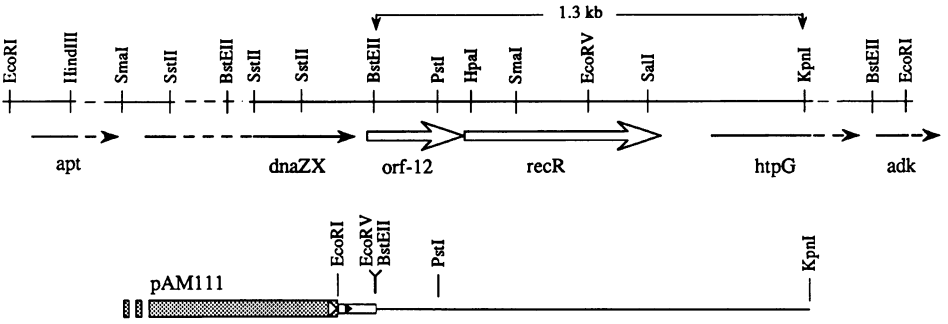

Figure 3. Restriction map showing the molecular organisation of *orf-12* and *recR* relative to flanking genes in the chromosome, and to the plasmid promoters in pAM111. The open arrows show the direction and extent of the reading frames determined from the DNA sequence presented in this paper. Vector plasmid sequences are indicated by shaded (pUC18) and open (pBR322) boxes. Promoters are indicated by open (*lac*) and closed (*tet*) triangles.

Location of the DNA encoding recR

Analysis of deletion derivative of pAM102 (Fig. 1) suggested that the DNA encoding *recR* spans the *Pst*I and *Sal*I sites to the centre of the cloned fragment, with *dnaZX* extending across the *Sst*II sites to the left. This location for *recR* was confirmed when 12 Tn*1000* insertions that inactivated *recR* in pAM107 (see Methods) were all located to a 1.0 kb region leftwards of the *Kpn*I site. One Tn*1000* insertion (number 7, Fig. 1) that did not inactivate *recR* was located close to the same *Kpn*I site. Since pAM109 is *recR*$^+$, we

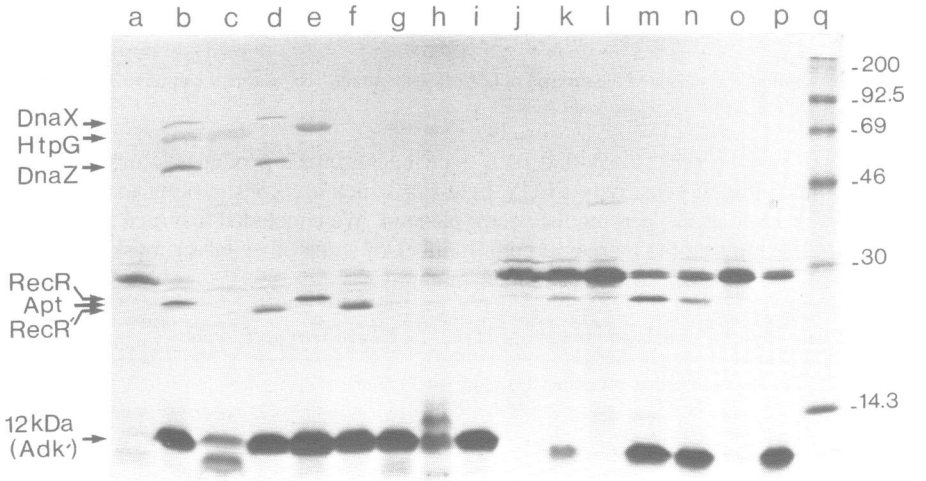

Figure 4. Fluorograph of labelled proteins from maxicell extracts separated on 15% poly acrylamide-sodium dodecyl sulphate gels. The plasmids used were a, pBR322; b, pAM102; c, pAM105; d, pAM106; e, pAM107; f, pAM108; g, pAM107::Tn*1000*-1; h, pAM107::Tn*1000*-9; i, pAM107::Tn*1000*-22; j, pJRD184; k, pAM110; l, pUC18; m, pAM111; n, pAM112; o, pAM114; p, pAM113. Molecular weight markers (lane q) are identified by size (kDa) on the right. Products encoded by the cloned DNA inserts are arrowed on the left.

suspected that *recR* was encoded within the central 1.3 kb *Bst*EII-*Kpn*I fragment of pAM102 (Fig. 1). To confirm this, the 3.3 kb *Bst*EII fragment was blunt-ended, digested with *Kpn*I, and the smaller of the two fragments generated was inserted into *Eco*RV-*Kpn*I digested pJRD184 (Fig. 1). The construct, pAM110, suppresses the UV sensitivity of a *recN recR* strain and does not sensitize a *rec*$^+$ strain (Table 2b). The latter observation confirms that the sensitization observed with pAM102 is not due to multiple copies of *recR*.

Transcriptional orientation of recR

The *Bst*EII-*Kpn*I fragment was excised from pAM110 by digestion with *Eco*RI and *Kpn*I and inserted into similarly digested pUC18 to generate pAM111. The 1.1 *Pst*I-*Kpn*I section of this insert was excised with *Pst*I and re-inserted into pUC18 in both orientations, giving pAM116 and pAM117 (Fig. 1). Examination of the effect of these constructs on the UV-sensitivity of *recR* strains revealed that while pAM116 (and pAM111) expresses *recR*$^+$, pAM117 does not (Table 2c, Fig. 2d−f). From these results we conclude that while the DNA encoding *recR* lies within the 1.1 kb *Pst*I-*Kpn*I fragment, the normal promoter does not. Presumably, *recR* is expressed from the *lac* promoter of pUC18 in pAM116, and possibly from the promoter of the Tc-resistance gene cloned with *recR* in pAM111. This would mean that *recR* is normally transcribed from left to right as drawn in Fig. 1, and from a promoter that is located to the left of the *Pst*I site.

Identification of the recR product

The restriction map of the 6.4 kb *Eco*RI fragment accumulated during the previous studies revealed (see Fig. 3) an overlap not only with the *apt-dnaZX* region to the left of the *Pst*I site but also with the *htpG-adk* region to the right of the *Sal*I site, both of which have been sequenced (19−22). We took advantage of these overlaps to identify the product of *recR* among the proteins encoded by pAM102.

Maxicell analysis of pAM102 and its derivatives (Fig. 4, lanes b−i) revealed that the 6.4 kb *Eco*RI fragment specifies at least seven polypeptides, five of approximately 74, 68, 56, 26, and 25 kDa, respectively, and two of about 12 kDa. From their mobilities and the effects of the various deletions on their production, we identified the 74 kDa and 56 kDa proteins as the products of *dnaZX*, the 68 kDa protein as the product of *htpG*, and the 25 kDa protein as the product of *apt*. The latter is more clearly visible in the absence of the 26 kDa protein (lanes g−i) and is also quite prominent in Fig 6, lane g.

The 12 kDa band appears to contain two proteins. One is removed by the *Pst*I deletion in pAM105 (Fig. 4, lane c) and by Tn*1000* insertion 9 (lane h), which is very close to the *Pst*I site, and the leftmost of the insertions in pAM107 that inactivate *recR*. It is unlikely to be the product of *recR* since it appears to be present in the *recR* mutant plasmids generated by the *Sal*I deletion (lanes d and f) or Tn*1000* insertions 1 and 22 (lanes g and i). The 12 kDa band remaining in these lanes is most probably the 109 amino-acid, N-terminal fragment of adenylate kinase produced as a result of the interruption of *adk* at the *Eco*RI site (Fig. 3; 22).

The 26 kDa protein is the only one affected by the *Pst*I deletion in pAM105 (Fig. 4, lane c), the *Sal*I deletion in pAM106 (lane d), and the three pAM107::Tn*1000* insertions examined (lanes g−i), all of which prevent expression of *recR*. We concluded that this protein must be the product of *recR*. Given that *recR* spans the *Sal*I site and is transcribed in the *Pst*I to *Sal*I orientation, the *Sal*I deletion in pAM106 and pAM108 should remove the DNA coding for the carboxy-terminus. Examination of Fig. 4, lanes d and f, reveals that the 26 kDa polypeptide is replaced by one of 24 kDa. The identity of RecR was confirmed when we found that of the proteins identified in pAM102, only the 26 kDa

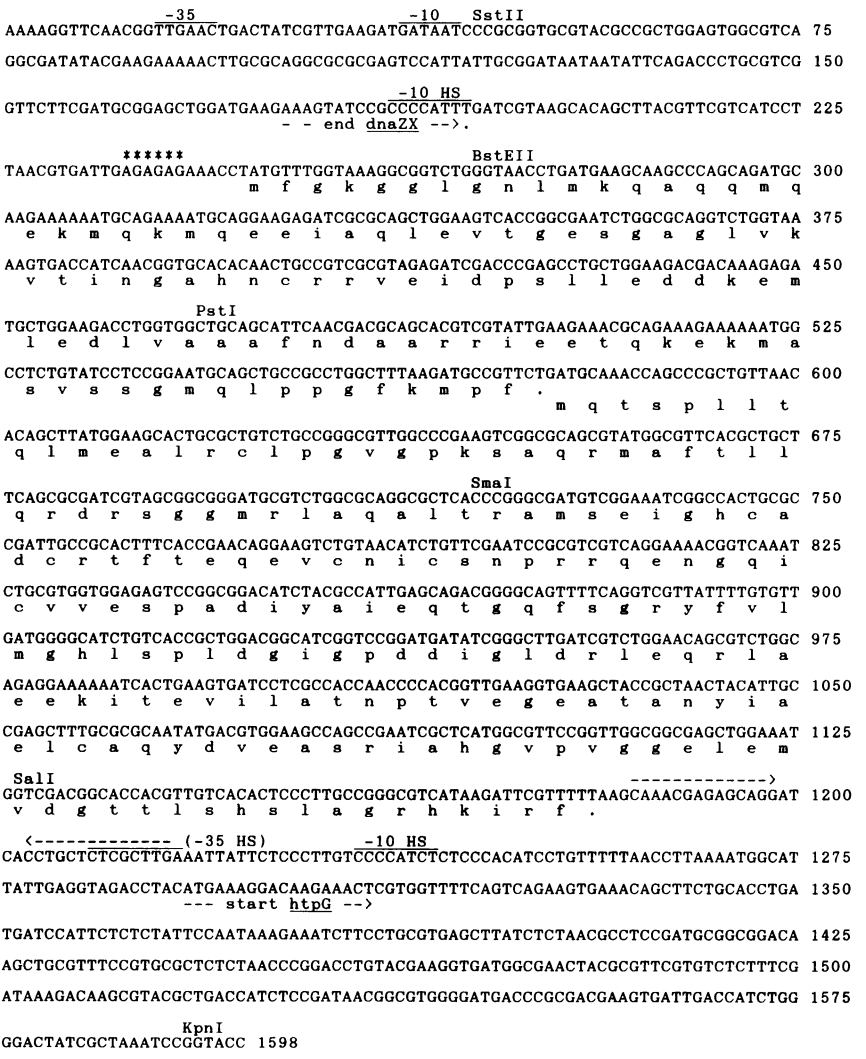

Figure 5. DNA sequence of *recR* and flanking regions. Restriction enzyme sites are listed above the first nucleotide of the recognition sequence. Putative promoter −10 and −35 sequences are overlined and labelled as such (HS = heat shock). Possible ribosome binding sites are identified by asterisks. Arrows define regions of dyad symmetry. Stop codons are labelled with a period. Translation of the *dnaZX* and *htpG* sequences flanking *recR* is omitted for clarity.

product is encoded by the 1.3 kb *Bst*EII-*Kpn*I region in pAM110 and pAM111 (Fig. 4, lanes j−m).

Nucleotide sequence of recR

The DNA spanning the *Pst*I site is clearly needed for expression of *recR* and of a previously unknown gene encoding a protein of 12 kDa. The observation that Tn*1000* insertion 9 in pAM107, which is the closest to the *Pst*I site (Fig. 1), eliminates the 26 kDa RecR

protein as well as the 12 kDa protein (Fig. 4, lane h) suggests that these two genes form an operon. To investigate this possibility directly, we sequenced the *Bst*EII-*Kpn*I region.

Figure 5 shows the nucleotide sequence of the *recR* region. The *Bst*EII-*Kpn*I section sequenced on both strands extends 1,329 nucleotides from bp 270 to the 3′ end. Inspection of this sequence revealed overlaps at both the 5′ and 3′ ends with previously reported sequences from the *dnaZX* and *htpG* regions, respectively (19−21). The overlap at the 5′ end extends from the *Bst*EII site to the *Pst*I site at bp 469. This region lies immediately downstream of the 3′ end of *dnaZX*. For reasons that will become obvious, Figure 5 shows an additional 269 bp of sequence at the 5′ end which extends from the end of *dnaZX* through to the *Bst*EII site (19,20). This part of the sequence matches perfectly with both of the published sequences (19,20), but was read from one strand only. The overlap at the 3′ end extends from the *Sal*I site at bp 1127 to the end of the sequence. This region is reported to include the 5′ end of *htpG*, which begins with the ATG at bp 1293 (21).

Analysis of the sequence between the end of *dnaZX* and the beginning of *htpG* revealed two major open reading frames. The first begins with an ATG at bp 249 and extends 327 nucleotides to a TGA stop codon at bp 576. It corresponds exactly with the open reading frame beginning at the end of the sequence reported by Flower and McHenry (20). It is preceded by the sequence AGAGAG which as suggested by Flower and McHenry could provide a ribosome binding site. Translation of this open reading frame would give a polypeptide of 109 amino acids with a predicted molecular weight of 12,015 Daltons. We concluded that it must be the structural gene for the 12 kDa protein located in this interval. Since the 5′ end spans the *Bst*EII site, the gene would be foreshortened in pAM110 and pAM111. This would account for the observation (Fig. 4, lanes k and m) that these two plasmids encode a new polypeptide of 11 kDa. From now on, we shall refer to this gene, whose function remains unknown, as *orf-12*.

The second open reading frame begins with an ATG at bp 578 and extends 603 nucleotides to a TAA termination codon at bp 1181. The proposed start codon overlaps the termination codon for *orf-12* by 1 bp. Translation of this second reading frame would give a polypeptide of 201 amino acids with a predicted molecular weight of 21,965 Daltons. Although RecR protein migrates on SDS gels with an apparent molecular weight of about 26 kDa, we suspected that this reading frame must be the *recR* gene.

To try and confirm the identity of *recR*, we took advantage of the multiple cloning site of pAM111 to delete progressively through the *Kpn*I site (leftwards as shown in Fig. 3) to the *Sal*I (pAM112), *Sma*I (pAM113) or *Pst*I (pAM114) site in the insert. All three deletions inactivated *recR* (data not shown). The *Pst*I deletion (Fig. 4, lane o) removed both the 26 kDa and the 11 kDa proteins. The *Sma*I deletion (lane p) removed only the 26 kDa protein, which confirms that it is encoded downstream of the 11 kDa protein made by pAM111. However, the *Sal*I deletion (lane n) did not appear to affect either protein. To examine the migration of the respective proteins in more detail, we ran gels in which mixtures of pAM111 and pAM112 or pAM108 and pAM112 maxicell extracts were loaded in the same wells alongside the relevant single extracts as controls. The results (Fig. 6) confirm that the 26 kDa band in pAM112 is unaffected by the *Sal*I deletion, whereas it is reduced to about 24 kDa in pAM108. The data presented also confirm that the 11 kDa protein encoded by pAM111 (and pAM110) does migrate faster than the 12 kDa band encoded by pAM107 and pAM108 (and pAM102). Given the sequence of *lacZ*α in pUC18 downstream of the *Sal*I, it became clear that the *Sal*I deletion in pAM112 would lead to a fusion protein with the carboxy terminus of LacZα that happens to have the same molecular

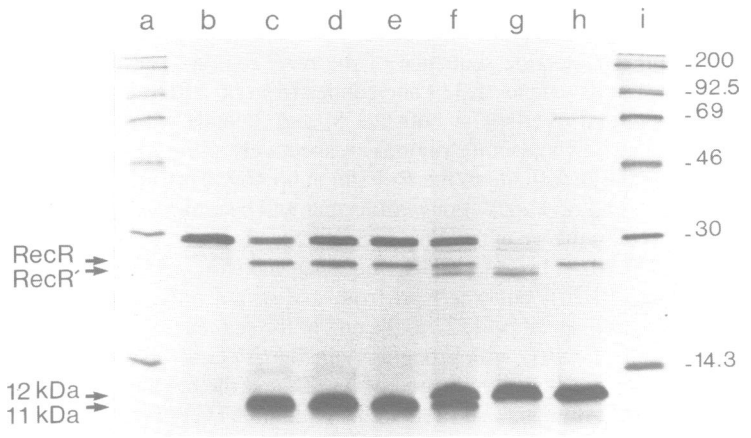

Figure 6. Fluorograph showing the relative mobilities of radio-labelled proteins from mixtures (approximately 1:1) of pAM111 and pAM112 (lane d) or pAM108 and pAM112 (lane f) maxicell extracts. pUC18 (lane b), pAM111 (lane c), pAM112 (lane e), pAM108 (lane g) and pAM107 (lane h) maxicell extracts were used as controls. Molecular weight markers (lanes a and i) are identified by size in kDa on the right. The pairs of proteins under examination are identified on the left.

weight as RecR. From these studies we are confident that the second reading frame is indeed *recR*. The region immediately downstream of the proposed stop codon for *recR* contains a GC-rich inverted repeat that could provide a signal for transcriptional termination (23).

Transcriptional and translational coupling of orf-12 and recR

Examination of the DNA sequence immediately upstream of *recR* revealed no obvious signals that could promote the initiation of transcription. The nearest sequences with reasonable matches to known promoters (24) are those within the 3′ end of *dnaZX* (see Fig. 5) pointed to by Flower and McHenry (20). Another sequence at the very end of *dnaZX* is similar to the −10 regions of heat-shock promoters (25,26). If the promoter for *recR* does indeed lie within this region, it would mean that *orf-12* and *recR* form a single operon as suggested above.

The overlap between the TGA termination codon for *orf-12* and the ATG initiation codon for *recR* would allow for these two genes to be also coupled at translation. The absence of any obvious ribosome binding site (27) at the appropriate distance from the start of *recR* suggests that synthesis of RecR may depend on the ribosomes translating *orf-12* messages being relocated to *recR* after reaching the stop codon. A comparison of the expression on *recR* from the plasmid constructs pAM111 and pAM116 provides support for the view that expression of *recR* does depend on translation of upstream sequences. While both plasmids are equally capable of restoring UV-resistance to *recR* mutants, it is clear that RecR is made in much smaller amounts from pAM116 (Fig. 7). The difference between the expression of *recR* in the two plasmids is particularly clear if the intensity of the RecR band is considered relative to the bands of the ampicillin resistance proteins. Since both are derivatives of pUC18, this difference is unlikely to be due to a simple effect of plasmid copy number.

Since the normal promoter is missing, we assume that expression of *recR* in pAM116

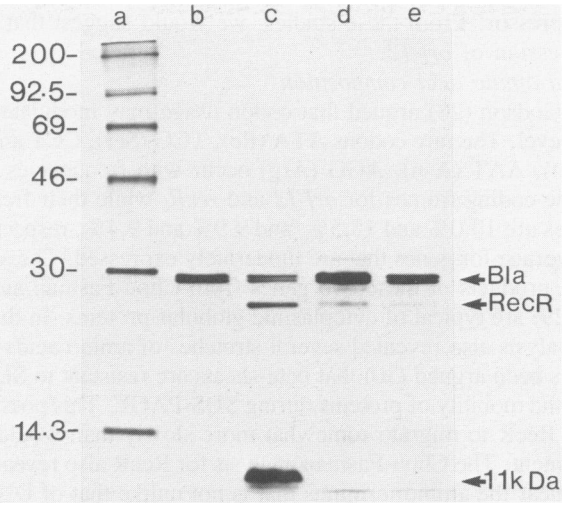

Figure 7. Fluorogram showing [^{35}S]methionine-labelled proteins encoded by pUC18 (lane b), pAM111 (lane c) and pAM116 (lanes d and e) in AB2480 maxicells grown with (lane e) or without (lanes b−d) 1.0 mM IPTG (isopropyl β-D-thiogalactoside). Molecular weight markers (lane a) are labelled in kDa. Proteins encoded by the chromosomal inserts in pAM111 and pAM116 are arrowed on the right, as are the two bands of β-lactamase (Bla) encoded by the ampicillin resistance gene of the vector.

is directed from the *lac* promoter (9). Examination of the DNA sequence in the *lacZ* to *orf-12* fusion created at the *Pst*I site in pAM116 reveals that translation from *lacZ* would terminate within the *orf-12* sequence at the TGA at bp 500 (Fig. 5), which is well before the start codon for *recR*. Presumably, this has a polar effect that reduces expression of *recR* quite severely, though not completely. pAM111 carries the slightly larger *Bst*EII-*Kpn*I fragment filled in at the *Bst*EII end and abutted to the 187 bp *Eco*RI-*Eco*RV fragment from (originally) pBR322 (Figs. 1 and 3). We sequenced across the *Eco*RV to *Bst*EII abutment in this construct and found 5′-GATGTAACC-3′ as expected. Examination of the known sequence for pBR322 revealed that expression of *recR* in this construct must depend on the tetracycline resistance gene (*tet*) promoter located within the *Eco*RI-*Eco*RV fragment. In this case, translation across the (*Eco*RV to *Bst*EII) fusion to the *orf-12* sequence would terminate at the two consecutive TGA stop codons at bp 277 and 280 shortly after the *Bst*EII site (Fig. 5). However, the ribosomes could presumably relocate to the ATG within this sequence, which happens to be in the right reading frame for *orf-12*. Translation initiated from this ATG would certainly account for the 11 kDa product encoded by this construct. What is more important is that it would enable the ribosomes to reach the start of *recR* and therefore allow for its increased expression in this case relative to pAM116. We cannot rule out the possibility that the different levels of RecR produced by these constructs reflect the relative strengths of the *lac* and *tet* promoters, especially since the genetic background used for the maxicell analysis (strain AB2480) is wild-type for the *lacI* repressor gene. However, we found that the level of RecR produced from pAM116 in AB2480 cells was not affected by the presence of the *lac* operon inducer, IPTG, in the growth medium (Fig. 7). This is not really surprising since the very high copy number of this pUC18 derivative would provide a sufficient excess of *lac* promoter sequences to

titrate out the repressor. From these studies, we would suggest that expression of *recR* depends on expression of *orf-12*.

Codon usage and amino acid composition

Konigsberg and Godson (28) argued that codon usage may modulate gene expression at the translational level. The rare codons ATA (Ile), TCG (Ser), CCT and CCC (Pro), ACG (Thr), CAA (Gln), AAT (Asn), AGG (Arg) occur with frequencies of 5.4% and 6.0% respectively in the coding frames for *orf-12* and *recR*, while their frequencies in the two non-coding frames are 10.0% and 15.5%, and 9.9% and 9.4%, respectively. These values are around the average for genes that are moderately expressed. The secondary structures predicted for the products of these two genes from Chou-Fasman analysis of the amino acid sequences (29) are typical of cytoplasmic globular proteins. In the case of RecR, the Chou-Fasman analysis also revealed several stretches of amino acids capable of forming beta-sheets. It has been argued (30) that beta-sheets are resistant to SDS binding and may therefore reduce the mobility of proteins during SDS-PAGE. This possibility could explain the tendency for RecR to migrate somewhat more slowly than would be predicted from its amino acid content. The Chou-Fasman analysis for RecR also revealed a possible helix-turn-helix motif near the amino-terminus that is not unlike that of DNA binding proteins (31,32). We have no direct evidence that RecR does bind DNA, but in view of its involvement in DNA repair and recombination it would not be surprising if it did.

DISCUSSION

We have shown that *recR* resides between *dnaZX* and *htpG* within a 6.4 kb *Eco*RI fragment of the *E. coli* chromosome and directs the synthesis of a polypeptide of about 26 kDa. The molecular dissection of the *recR* region also revealed a previously unknown locus between *dnaZX* and *recR* that encodes a polypeptide of about 12 kDa. Our studies indicate that these two genes form an operon transcribed in the direction of *htpG* as indicated in Fig. 3. A search of the GenBank (release 59) DNA sequence and NBRF-PIR (release 20) protein sequence databases revealed no major similarities with *orf-12* and provided no obvious clues to the function of the 12 kDa product. The fact that the gene lies within the same operon as *recR* is probably significant but does not necessarily mean that it too is involved in DNA repair.

Examination of the DNA sequence immediately downstream of *dnaZX* reveals no obvious signals for the termination of transcription, which raises the possibility that transcription from the *dnaZX* promoter region (19,20) may extend into the *recR* operon. The good match to a −10 region for a heat shock promoter at the very end of *dnaZX*, raises the additional possibility that expression of the *orf-12-recR* operon is induced along with *htpG* as part of the heat-shock response (21,25). The current list of heat-shock proteins of unknown origin (25) includes candidates of about the right sizes. The increased UV-sensitivity of a *recR* single mutant (AM207) at low temperature (compare the vector controls in panels a and d of Fig 2) could therefore be of some significance. We have since examined the survival of identical *recR252* strains at different temperatures and confirmed that it is indeed considerably reduced at 30°C relative to 37°C. Since the mutation involved is a transposon insertion, it seems unlikely that this phenotype is a reflection of a cold-sensitive RecR protein. While we cannot rule out a simple physiological artifact, it seems possible that some expression of the heat-shock response may help to alleviate the defect in DNA repair in *recR* mutants.

Since *htpG* appears to have its own promoter (21) and is preceded by a possible rho-

dependent (23) transcriptional terminator (Fig. 5), we assume that this locus can be transcribed independently of *recR*. There is an overlap between the inverted repeat of this terminator and the proposed −35 region of the (heat-shock) promoter for *htpG*, but the significance of this overlap is not clear. Bardwell and Craig (33) reported that while an *htpG* deletion mutant is viable, it grows poorly at high temperature. Neither of our two *recR* mutations (*recR252* and *recR256*), both of which are transposon insertions (1), has any discernible effect on growth at temperatures ranging from 18°C to 42°C (unpublished work). We conclude that the phenotype we have associated with these mutations is due to the loss of the 22 kDa RecR protein, and is unlikely to be related to any polar effect on *htpG*.

The location of *recR* downstream of *dnaZX*, which encodes two subunits of DNA polymerase III holoenzyme (26), is also reminiscent of the location of *recF* downstream of *dnaN* (34), which encodes another polymerase subunit. While these locations may be coincidental, they could point to a relationship between DNA replication and recombination of the type envisaged in phage T4 where break-join events are thought to prime much of the late DNA synthesis by creating new replication forks (35).

ACKNOWLEDGEMENTS.

We are grateful to D. Bramhill and P. T. Emmerson for providing the λ library used to search for *recR*. We would also like to thank Carol Buckman for technical support and Gary Sharples, Stuart Morton, Isam Naom and Fernando Gibson for advice and encouragement. This work was financed by grants to R.G.L. from the Science and Engineering Research Council. A.A.M was in receipt of a postgraduate training award from the Government of Iraq.

REFERENCES

1. Mahdi, A.A. and Lloyd, R.G. (1989) Mol. Gen. Genet. **216**, 503−510.
2. Clark, A.J. and Low, K.B. (1988), In: Low KB (Ed), The Recombination of Genetic Material, Academic Press, New York London, pp 155−215.
3. Smith, G.R. (1988) Microbiol. Rev. **52**, 1−28.
4. Bachmann, B.J. (1987), In: Neidhardt, F.C., Ingraham, J.L., Low, K.B., Magasanik, B., Schaechter, M. and Umbarger, H.E. (Eds), *Escherichia coli* and *Salmonella typhimurium* cellular and molecular biology. American Soc. for Microbiol., Washington DC, pp 1190−1219.
5. Attfield, P.V.A., Benson, F.E. and Lloyd, R.G. (1985) J. Bacteriol. **164**, 276−281.
6. Lloyd, R.G., Evans, N.P. and Buckman, C. (1987) Mol. Gen. Genet. **209**, 135−141.
7. Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heynecker, H.L., Boyer, H.W., Crosa,J.H. and Falkow, S. (1977) Gene **2**, 95−113.
8. Heusterspreute, M., Thi, V.H., Emery, S., Tournis-Gamble, S., Kennedy, N. and Davidson, J. (1985) Gene **39**, 299−304.
9. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Gene **33**, 103−119.
10. Hashimoto-Gotoh, T., Franklin, F.C.H., Nordheim, A. and Timmis, K.N. (1981) Gene **16**, 227−235.
11. Arthur, H.M., Bramhill, D., Eastlake, P. and Emmerson, P.T. (1982) Gene **19**, 285−295.
12. Lloyd, R.G. and Buckman, C. (1985) J. Bacteriol. **164**, 836−844
13. Silhavy, T.J., Berman, M.L. and Enquist, L.W. (1984) Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
14. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
15. Isch-Horowicz, D. and Burke J.F. (1981) Nucleic Acids Res. **9**, 2989−2998.
16. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463−5467.
17. Sancar, A., Wharton, R.P., Seltzer, S., Kacinski, B.M., Clarke, N.D. and Rupp, W.D. (1981) J. Mol. Biol. **148**, 45−62.
18. West, S.C. and Emmerson, P.T. (1977) Mol. Gen. Genet. **151**, 57−67.

19. Yin, K., Blinkowa, A. and Walker J.R. (1986) Nucleic Acids Res. **14**, 6541–6549.
20. Flower, A.M. and McHenry, C.S. (1986) Nucleic Acids Res. **14**, 8091–8101.
21. Bardwell, J.C.A. and Craig, E.A. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 5177–5181.
22. Brune, M., Schumann, R. and Wittinghofer, F. (1985) Nucleic Acids Res. **13**, 7139–7151.
23. Brendel, V., Hamm, G.H. and Trifonov, E.N. (1986) J. Biomolec. Structure and Dynamics **3**, 705–723.
24. Harley, C.B. and Reynolds, R.P. (1987) Nucl. Acids Res. **15**, 2343–2361.
25. Neidhardt, F.C. and VanBogelen, R.A. (1988) In: Neidhardt, F.C., Ingraham, J.L., Low, K.B., Magasanik, B., Schaechter, M. and Umbarger, H.E. (Eds), *Escherichia coli* and *Salmonella typhimurium* cellular and molecular biology. American Soc. for Microbiol., Washington DC, pp 1334–1345.
26. McHenry, C. S. (1988) Ann. Rev. Biochem. **57**, 519–550.
27. Shine, J. and Dalgarno, L. (1974) Proc. Natl. Acad. Sci. U.S.A. **71**, 1342–1346
28. Konigsberg, W. and Godson, G.N. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**,687–691.
29. Chou, P.Y. and Fasman, G.D. (1978) Adv. Enzymol. **47**, 45–148.
30. Chu, C.C., Templin, A. and Clark, A.J. (1989) J. Bacteriol. **171**, 2101–2109.
31. Schleif, R. (1988) Science **241**, 1182–1187.
32. Pabo, C.O. and Sauer, R.T. (1984) Ann. Rev. Biochem. **53**, 293–321.
33. Bardwell, J.C.A. and Craig, E.A. (1988) J. Bacteriol. **170**, 2977–2983.
34. Armengod, M.E., Garciá-Sogos, M. and Lambies, E. (1988) J. Biol. Chem. **263**, 12109–12114.
35. Mosig, G. (1987) Ann. Rev. Genet. **21**, 347–371.

This article, submitted on disc, has been automatically converted into this typeset format by the publisher.