JOURNAL OF BACTERIOLOGY, Nov. 1994, p. 7024–7031
0021-9193/94/$04.00+0

Vol. 176, No. 22

Transcription of the *Escherichia coli recE* Gene from a Promoter in Tn*5* and IS*50*

ALVIN J. CLARK,[1]* LESLIE SATIN,[1] AND CHARLES C. CHU[2]

Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, California 94720-3202,[1] and Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20892[2]

Received 8 October 1993/Accepted 1 September 1994

Six *sbc*::Tn*5* insertions and one *sbc*::IS*50* insertion, which cause *recE* expression in *Escherichia coli*, have been cloned, and their DNA sequences have been determined. The sites of insertion are found at three positions in a 10-bp region: 58, 63, and 68 bp upstream of *recE*. Primer extension experiments with the cloned Tn*5* insertions demonstrate that *recE* transcripts start adjacent to the insertion elements of five of these mutations and both adjacent and one nucleotide within the insertion element for the sixth mutation. This supports the hypothesis that these mutations have inserted a promoter, and PCR analysis reveals an outward promoter within the distal 69 nucleotides of Tn*5*. Primer extension analysis of RNA from the uncloned Tn*5* and IS*50* mutants reveals three additional insertion sites close to the others. Because all the insertions lie in the spacer region between *racC* and *recE*, transcribed in *sbcA6* and *sbc-23* strains, we propose that these insertions be renamed *recEs*::Tn*5* and *recEs*::IS*50*.

recE is one of many genes involved in conjugational recombination in *Escherichia coli* (10, 33). *recE* resides on the defective lambdoid prophage Rac in the *E. coli* chromosome (13, 22, 23). In Rac$^+$ strains no trace of RecE protein is seen by use of immunological methods (24), implying that *recE* is either transcribed and not translated or not transcribed. Point mutations (24, 25) lead to the expression of *recE* as judged by the presence of RecE protein (14, 24) and its DNA exonuclease activity exonuclease VIII (5, 16, 17, 21). The few point mutations tested are *cis*-acting and dominant to their wild-type allele, implying that they are promoter, operator, or transcription terminator mutations (13, 25). These mutations are called *sbcA* mutations (e.g., *sbcA6*) because they suppress the Rec$^-$ and UV^s phenotypes of *recB* and *recC* mutations and because they may affect only one gene.

Tn*5* and IS*50* insertion mutations also lead to *recE* expression (13, 36). Like *sbcA6*, the insertion mutations tested were *cis* acting and dominant to the wild-type alleles, leading to the hypothesis that a new promoter was created or inserted (13, 36). Southern blot analysis placed the sites for insertion near the N-terminal end of *recE* (9, 36). The goal of this work is to locate the insertions by DNA sequencing and to examine *recE* transcripts in order to learn the location of any functional promoter.

MATERIALS AND METHODS

Nomenclature of insertion mutations. Fouts et al. (13) described eight Tn*5* and six IS*50* insertions which resulted in expression of *recE*. These were named *sbc* mutations and given *sbc* allele numbers because the mitomycin-resistant (Mit^r) phenotype by which they were detected resulted from suppressing *recB21 recC22*. The *sbc* cistron designation was omitted, however, because there was no evidence that the sites of insertion defined the location of *sbcA*. Thus, the mutations were known as *sbc-82*::Tn*5*, *sbc-83*::IS*50*, etc. (13, 36). The present work has shown that the insertions all lie in the transcribed spacer between the N terminus of *recE* and the C terminus of *racC*. To give this spacer the cistron designation *sbcE* implied to us that inactivation of this *sbc* allele was responsible for expression of *recE*. Since inactivation is not the mechanism for expression, we have chosen to rename the mutations *recEs* in accord with the recommendation of Bachmann and Low (4), although this necessitates changing allele numbers as well as the gene name. The correspondence of *sbc* and *recEs* allele numbers is given in Table 1.

Bacterial strains. All bacteria used in this work are derivatives of *E. coli* K-12. Those with *recEs* mutations are listed in Table 1. Other strains are described as they are mentioned. For genetic nomenclature, the conventions of Demerec et al. (12) and Bachmann (3) are followed.

Media and chemicals. Luria broth (35) supplemented with 50 μg of ampicillin per ml and/or 25 to 30 μg of kanamycin per ml as appropriate was used to grow strains for plasmid and RNA isolations. Restriction endonucleases were purchased from New England Biolabs (Beverly, Mass.). The reaction conditions used were those suggested by the manufacturer for high-, medium-, and low-salt buffers (11) with the exception that bovine serum albumin was omitted.

Cloning of insertion mutations. *recEs*::Tn*5* and *recEs*::IS*50* insertions were cloned by taking advantage of the fact that there are no *Eco*RI or *Cla*I restriction endonuclease sites within Tn*5* or IS*50*. The Tn*5* and IS*50* elements in these mutant insertions reside in a 2.2-kb *Cla*I-*Eco*RI fragment (see Fig. 1B), and were estimated to be 155 ± 100 bp from the *Cla*I site from the data of Willis et al. (36) and Chu et al. (9).

Tn*5* insertions were cloned by taking advantage of the additional fact that Tn*5* encodes *aphA*, a gene that confers kanamycin resistance to *E. coli* (30). Chromosomal DNA was prepared from bacterial strains containing *recEs*::Tn*5* mutations as described by Willis et al. (37). Eight micrograms of chromosomal DNA and 2 μg of pBR322 DNA were simultaneously digested with excess *Eco*RI and *Cla*I, treated with ligase, and then used to transform competent AB1157 to ampicillin and kanamycin resistance by standard methods (26).

* Corresponding author. Mailing address: Department of Molecular and Cell Biology, MCB: 410 Barker Hall, University of California at Berkeley, Berkeley, CA 94720-3202. Phone: (510) 642-0985. Fax: 510-643-5035. Electronic mail address: ecorec2@violet.berkeley.edu.

TABLE 1. Correspondence between old *sbc* and new *recEs* mutant allele numbers

Allele[a]		Strain[b]	
sbc	*recEs*	Mal$^-$ (λ^r)	Mal$^+$ (λ^s)
82::Tn*5* (IB)	*3082*::Tn*5* (IB)		JC12020
83::IS*50* (IIA)	*3083*::IS*50* (IIA)	JC12023	JC12050
84::Tn*5* (IB)	*3084*::Tn*5* (IB)	JC12024	JC12051
85::IS*50* (IIB)	*3085*::IS*50* (IIB)	JC12025	JC12052
86::Tn*5* (IB)	*3086*::Tn*5* (IB)	JC12026	JC12053
87::IS*50* (IIB)	*3087*::IS*50* (IIB)	JC12027	JC12054
111::Tn*5* (IB)	*3111*::Tn*5* (IB)		SDB1041
112::Tn*5* (IB)	*3112*::Tn*5* (IB)		SDB1045
113::Tn*5* (IB)	*3113*::Tn*5* (IB)		SDB1046
114::Tn*5* (IA)	*3114*::Tn*5* (IA)		SDB1049
115::IS*50* (IIA)	*3115*::IS*50* (IIA)		SDB1048
117::IS*50* (IIA)	*3117*::IS*50* (IIA)		SDB1051
118::IS*50* (IIB)	*3118*::IS*50* (IIB)		SDB1050
119::Tn*5* (IA)	*3119*::Tn*5* (IA)		SDB1047

[a] Roman numerals in parentheses refer to the insertion element, and letters refer to the orientation of the insertion element. I refers to Tn*5*, and II refers to IS*50*. A for Tn*5* refers to the orientation in which IS*50*R is proximal to *recE* and IS*50*L is distal. B for Tn*5* refers to the opposite orientation. In both orientations (IA and IB) the outer end of IS*50* is proximal to *recE*. B for IS*50* refers to the orientation in which the outer end of either IS*50*L or IS*50*R is proximal to *recE*. A for IS*50* refers to the orientation in which its inner end is proximal to *recE*.

[b] Referred to previously by Fouts et al. (13) and Willis et al. (36).

Plasmids containing six *recEs*::Tn*5* mutations were obtained. Three of these, pJC866, pJC931, and pJC944, have the expected 7.9-kb insert (2.2-kb *Cla*I-*Eco*RI fragment plus 5.7-kb Tn*5* insertion). Three others have additional DNA fragments. pJC899 contains an extra 4.3 kb of pBR322 vector fragment. pJC902 contains an extra 2.4 kb of *Eco*RI fragment. pJC929 contains two additional *Eco*RI-*Cla*I fragments that together amount to no more than 100 bp.

recEs3083::IS*50* was cloned from JC12023 chromosomal DNA as described above with the exception that the appropriate clone was selected by colony hybridization (35) instead of antibiotic resistance. An oligonucleotide called prJC8 was 5′ end labeled and used as a probe (36). The sequence of prJC8 is 5′ ATGCCAACCCTGATCCGG 3′, which is identical to a portion of *recE* (coordinates 1556 to 1573 in reference 9). Plasmid pJC951 containing the expected 3.7-kb insert (2.2-kb *Cla*I-*Eco*RI fragment plus 1.5 kb-IS*50* insert) in pBR322 was obtained.

Sequencing of insertion mutations. Plasmid DNA was prepared and sequenced by the chain termination method (7). Synthetic primers were purchased to sequence both junctions at the site of Tn*5* insertion: prJC10 has the sequence 5′ CTGGCAACTCTAATTGC 3′, which has coordinates 350 to 366 in reference 9, and prJC11 has the sequence 5′ TTCGC TTTCCGTAACAGG 3′, which is complementary to codons 8 to 12 and parts of codons 7 and 13 of *recE* and has coordinates 530 to 513 in reference 9 and Fig. 1A. Primers were purchased from the laboratory of Jack F. Kirsch (University of California, Berkeley) and were purified by thin-layer chromatography based on a method described by Alvarado-Urbina et al. (1).

RNA preparation. RNA was isolated from *E. coli* by the rapid method described by Summers (34) with two exceptions: (i) after lysis and centrifugation, the pellet of cell debris was removed with a toothpick, and (ii) the final step was a phenol extraction to remove RNase activity. Ethanol-precipitated RNA was stored in diethyl pyrocarbonate-treated water at −20°C.

Primer extension. The procedure utilized for primer extension was that of McKnight et al. (27). The reaction was carried out at 42°C for 1 h. The mixture was extracted with phenol-chloroform and precipitated with ethanol. The dried pellet was dissolved in sequencing gel loading buffer, boiled for 3 min, and electrophoresed on a 5% sequencing gel. M13 DNA was sequenced by using the −40 universal primer from New England Biolabs and loaded on the same gel as molecular weight standards. Actinomycin D for the McKnight procedure was purchased from Calbiochem (San Diego, Calif.). Avian myeloblastosis virus reverse transcriptase for both procedures was purchased from Life Sciences, Inc. (St. Petersburg, Fla.) or Promega (Madison, Wis.). Oligonucleotide prJC11 was the primer for *recE* transcripts (see Fig. 1A). It was 5′ end labeled by a kinase forward reaction by the method of Woods (38). T4 polynucleotide kinase for the Woods method was purchased from Boehringer Mannheim (Indianapolis, Ind.).

Preparation, identification, and cloning of PCR fragments. DNA extracted from *recEs3112* strain JC15203 by the method of Ish-Horowicz and Burke (15) was incubated with pairs of DNA primers according to the method of Mullis and Faloona (29). prJC75 was paired with prJC71 in one reaction and with prJC73 in a second reaction to make DNA fragments of 68 and 214 bp, respectively. Thirty cycles (consisting of 2 min at 94°C, 2 min at 55°C, and 2 min at 72°C preceded by 5 min at 94°C and succeeded by 5 min at 72°C) were used. prJC70 was paired with prJC72 in a third reaction to make a fragment of 144 bp. Forty cycles of the preceding protocol were used. Fragments were isolated from polyacrylamide gels and treated as recommended in the SureClone kit (Pharmacia) prior to ligation to dephosphorylated pUC18 DNA cut with *Sna*I. Strain DH5α was transformed to Amp^r by using each DNA preparation. Individual colonies were screened for plasmids containing inserts of the appropriate size. Plasmid DNA was obtained, and the sequence of the DNA inserted was determined by the chain termination method of Chen and Seeberg (7).

Plasmid DNA of each pUC18 derivative was purified and treated as follows to reclone each insert and a few flanking nucleotides from the multicloning site of pUC18, using pK04 as vector. Plasmid DNA was treated with four restriction nucleases: *Bam*HI and *Eco*RI to remove the insert for subcloning and *Ssp*I and *Sca*I to cleave the *bla* gene of the pUC18 vector. DNA was purified by phenol extraction and ethanol precipitation. It was added to a gel-isolated 3.6-kb fragment of *Eco*RI-*Bam*HI-treated pK04 DNA. Ligase was added to each mixture. After suitable incubation, the DNA mixtures were used to transform *galK* mutant AB1157 to Amp^r by using galactose indicator MacConkey agar. Purple and white colonies were picked separately, pooled, and tested for the size of fragment inserted. The inserts in DNAs from individual clones were sequenced by the chain termination method of Chen and Seeberg (7).

Following are the sequences of the five primers used for PCR analysis as pairs:

(i) *Eco*RI cleavable

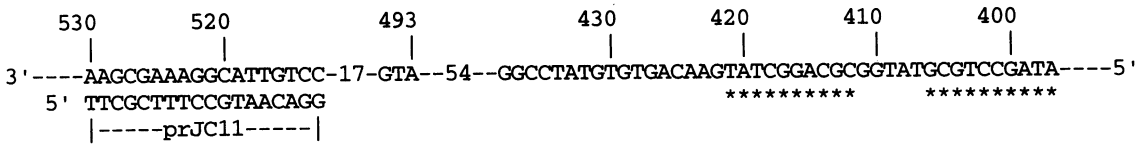
PrJC71 5′ G · GAATTC · AGTATGTATCCGGCAAG
prJC72 5′ G · GAATTC · CATGGAAGTCAGATCCT
prJC73 5′ G · GAATTC · CACAGATTTAGCCCAGT

(ii) Intended to be *Bam*HI cleavable

prJC70 5′ GC · CTTAGG · AACTCAGATTAAAATTC
prJC75 5′ GC · CCTAGG · AACTCAGATTAAAATTC

The dots set off the sequences which were to constitute restriction sites for cloning the PCR fragments. Mistakes in

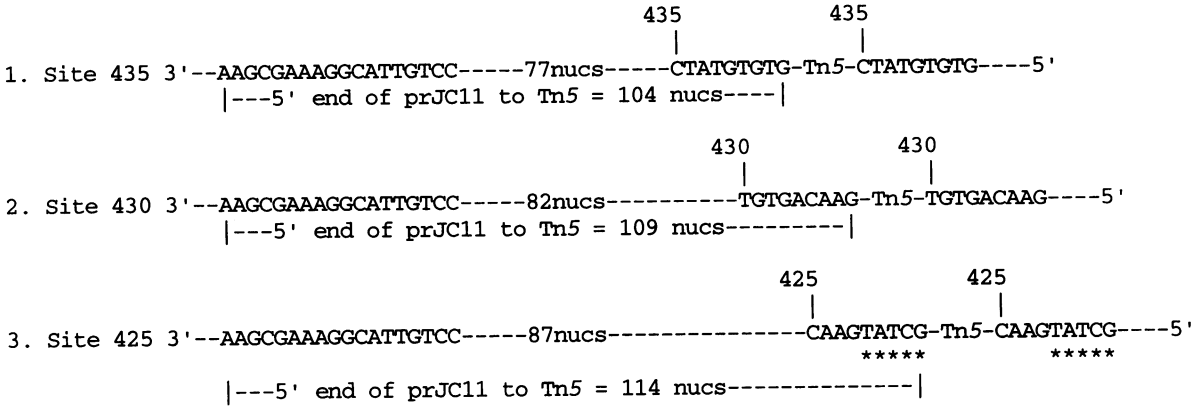
A. Map of relevant portions of *recE* and *recEs*



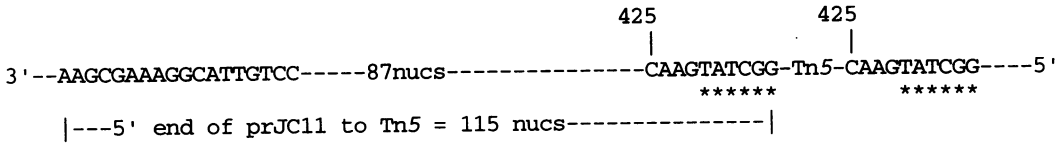
B. Tn*5*/IS*50* Insertion sites

1. Determined from DNA sequence: 435 430 425
2. Inferred from primer extension: 435 428 427 425 418

C. Maps of Tn*5* insertions which make 9 nucleotide duplications at sites 435,430 and 425



D. Map of Tn*5* insertion which makes a 10 nucleotide duplication at site 425



E. Maps of IS*50* insertions with putative insertion sites at 418, 427 and 428

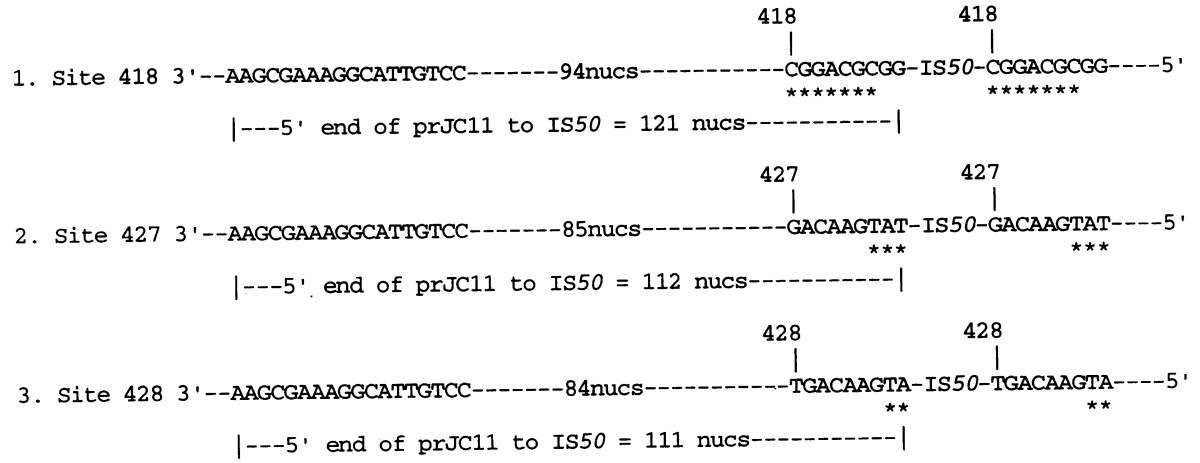


FIG. 1. Insertion sites of *recEs*::Tn*5* and *recEs*::IS*50* mutations and description of primer extension analysis. (A) Relevant portions of *recE* and *recEs* are mapped as a single-strand sequence written unconventionally 3′ to 5′. Numbers above the sequence refer to coordinates from Fig. 2 of reference 9. The triplet at position 493 is the translational initiation codon of *recE*; *recE* is to the left, and *recEs* is to the right. Numbers in the sequence refer to the number of omitted bases. Asterisks mark an inverted repeat sequence. The second line of sequence is primer prJC11 and is written under its complement in *recE*. prJC11 was used for primer extension analysis. (B) Tn*5* and IS*50* insertion sites are marked by vertical lines and numbers referring to the sequence in panel A. (C) Maps of Tn*5* insertions which have made 9-nucleotide duplications at three different sites. *recE* and *recEs* sequences are written in two parts on the first line of sequence. On the left is the portion of *recE* complementary to prJC11. On the right is Tn*5* flanked by direct repeat sequences. The number of nucleotides separating the two parts is stated between them. Asterisks mark nucleotides from the inverted repeat sequence mentioned in the description of panel A. The second line of sequence contains prJC11 and the number of nucleotides between its 3′ terminus and the first nucleotide of Tn*5*. Sequence numbers refer to panel A. (D) Map of a Tn*5* insertion which has made a 10-nucleotide duplication at site 425. Conventions are those employed in panel C. (E) Representation of IS*50* inserted at three different sites as inferred from primer extension data. Conventions are those employed in panel C.

composing the sequences in group ii prevented cleavage by *Bam*HI, however, and so we used a blunt-end cloning procedure. Sequence analysis shows that the 214- and 144-bp PCR fragments were cloned as if they were blunt ended and that the 68-bp fragment had a terminal nucleotide removed. For subsequent cloning into pK04 from the pUC18 intermediate, the *Bam*HI site of the multicloning site was used. A substitution was inadvertently placed at position 11 in prJC71 during synthesis. The sequence shows an A, but the sequence of Tn*5* (see Fig. 4) shows a G at this position. Inspection of the sequences shows that the single-base differences between fragments generated by enzyme action or human error do not affect the interpretation of the results.

RESULTS

Insertion sites of Tn*5* and IS*50* by DNA sequence analysis. Location of one IS*50* and six Tn*5* insertions that cause expression of *recE* was done by DNA sequence determination. Figure 1A shows a sequence map of the N-terminal portion of *recE* (on the left) and a portion of the spacer (*recEs*) between *racC* and *recE* (on the right). The sequence is drawn in 3′ to 5′ orientation to highlight the complementarity with the primer prJC11, whose use is discussed in the next section. Figure 1B shows three insertions sites (435, 430, and 425) determined by DNA sequencing. The sites are between *recE* and a hyphenated inverted repeat sequence marked by asterisks in Fig. 1A. The location of each Tn*5* element, flanked by a 9-nucleotide duplication at each position, is shown in Fig. 1C. As demonstrated previously (8), one Tn*5* element at site 425 is flanked by a 10-nucleotide duplication (Fig. 1D).

Promoter insertion by Tn*5* and IS*50* insertions. To test whether Tn*5* inserted or created a promoter in the six cloned Tn*5* mutants, primer extension experiments were performed to determine the locations of the 5′ ends of any RNA transcripts. The primer used, prJC11, is complementary to codons 8 to 12 (and parts of codons 7 and 13) in *recE* and hence detects *recE* transcripts. RNA was prepared from the six sequenced Tn*5*

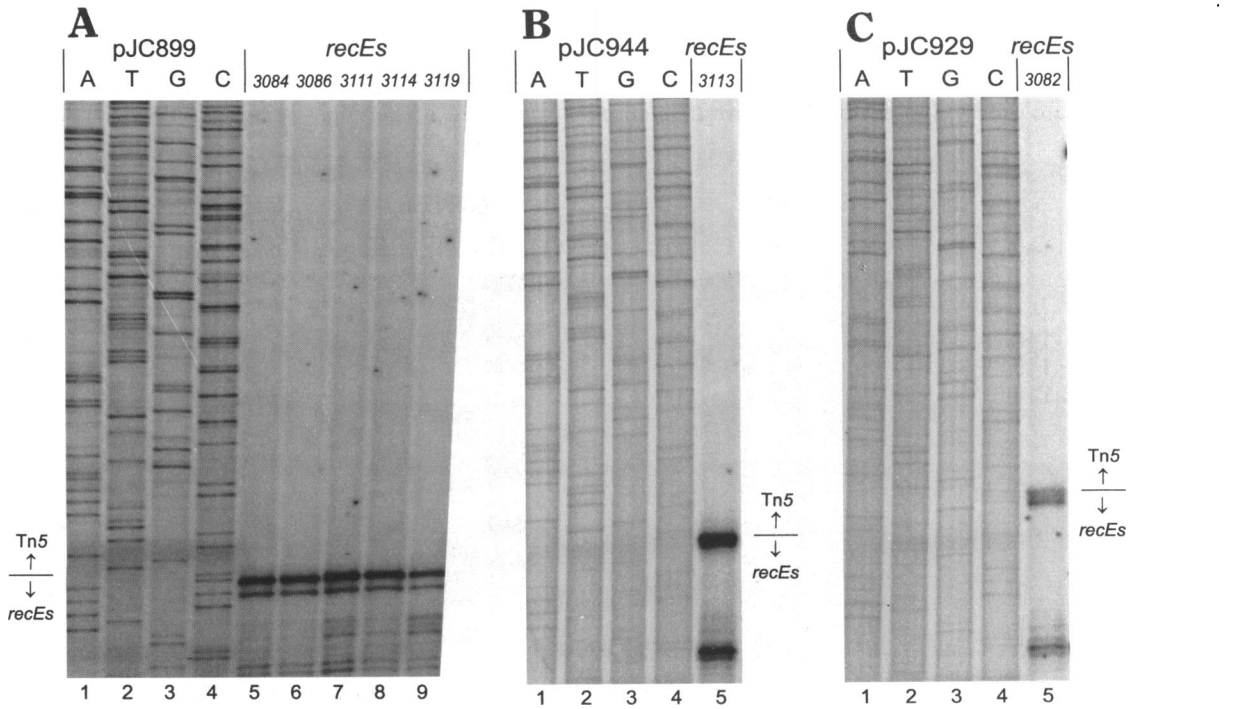


FIG. 2. Primer extension analysis of Tn*5* insertion mutations at site 435 (9-base duplication) (A), site 430 (9-base duplication) (B), and site 425 (10-base duplication) (C). Sequencing ladders (lanes 1 to 4) are derived from plasmid DNA carrying a cloned insertion at the cognate site and primer prJC11. The boundary between *recEs* and Tn*5* sequence is marked in the margin next to each ladder. The remaining lanes contain DNA derived by extending primer prJC11 hybridized to RNA derived from each insertion mutant strain.

TABLE 2. Primer extension analysis of Tn*5* and IS*50* insertions in *recEs*

Site	Length of extension product(s)[a]	*recEs* Allele[b]	
		Sequenced	Unsequenced
435	ND[c]	*3083*::IS*50* (IIA)	None
	104	*3084*::Tn*5* (IB)	*3086*::Tn*5* (IB)[d]
		3114::Tn*5* (IA)	*3087*::IS*50* (IIB)
		3119::Tn*5* (IA)	*3111*::Tn*5* (IB)[e]
430	109	*3113*::Tn*5* (IB)	None
428	111	None	*3085*::IS*50* (IIB)
			3118::IS*50* (IIB)
427	194	None	*3117*::IS*50* (IIA)
425	114, 115, 116	*3082*::Tn*5* (IB)	None
	114	*3112*::Tn*5* (IB)	
418	203	None	*3115*::IS*50* (IIA)

[a] Lengths (nucleotides) for products from sequenced mutations at the following sites were measured against a sequence ladder generated with prJC11 and a plasmid carrying a Tn*5* mutation at that site: site 435, plasmid pJC899; site 430, plasmid pJC944; and site 425, plasmid pJC929 or pJC866, depending on whether a 10-base or a 9-base duplication had occurred. The length of the product for the mutation at site 427 was measured against a ladder generated from pJC866. Lengths for products from unsequenced mutations at sites 435, 428, and 418 were measured against a sequence ladder generated with M13 phage and primer no. 1212 from New England Biolabs. They were corrected by one nucleotide as indicated by the error in lengths for the other sites measured against the M13 ladder. Lengths of products from two of the unsequenced mutations at site 435 were measured against a ladder generated from plasmid pJC899 carrying a sequenced mutation at site 435.

[b] (IA) and (IB) represent the two orientations of Tn*5* (35). (IIA) and (IIB) represent the two orientations of IS*50* (35); *recEs3114*::Tn*5* and *recEs3119*::Tn*5* are possible siblings (14).

[c] ND, not determined.

[d] RNA was purified from JC12026 and JC12053 for primer extension. Both produced extension products of 104 nucleotides.

[e] RNA was purified from SDB1041 and KF1053 for primer extension. Both produced extension products of 104 nucleotides. In Fig. 1B and C of reference 8, a sequence labeled *sbc-111* at site 425 is shown. This is an error. The figures should have been labeled *sbc-112* or, as renamed here, *recEs3112*.

insertion mutants and JC5491, the strain from which the insertion mutants were derived (13). If a promoter is created by or is contained within Tn*5*, the extension products from mutant samples are expected to exhibit distinctive differences in length. The results for five of the six are shown in Fig. 2A, lanes 5, 8 and 9; B, lane 5; and C, lane 5. The nucleotide lengths of the most slowly moving bands observed with mutant strain RNA and not wild-type strain RNA are listed in Table 2. The nucleotide sequence ladder used to measure the length of each extension product was derived from a plasmid carrying the same or a cognate insertion (Fig. 2). The extension products consist of the 18-nucleotide primer plus the complement of any additional RNA between the 3′ terminus of the primer and the 5′ RNA terminus. This length is specified in Fig. 1C. In four cases the extension product was long enough that the 5′ terminus of its template lay immediately adjacent to the first nucleotide of Tn*5*. One exception was the mutation at site 425 (Fig. 2C, lane 5). In this case multiple bands were seen. One band represented a product extending one nucleotide into Tn*5*. This mutation carries an abnormal 10-nucleotide duplication flanking Tn*5* (8). Another mutation (*recEs3112*::Tn*5*) also at site 425 carries the normal 9-nucleotide duplication. We compared the primer extension products from these two mutations alongside sequencing ladders from their cognate plasmid clones (Fig. 3). The results show one extension product for the normal mutation (Fig. 3, lanes 6 and 12) with its 5′ base adjacent to the first base in Tn*5*. Three products are again seen for the abnormal mutation.

Insertion sites of Tn*5* determined by primer extension analysis. Two Tn*5* insertions (*recEs3086*::Tn*5* and *recEs 3111*::Tn*5*) were not cloned. We determined the site of attachment by primer extension. Extension products were found to migrate exactly like those of sequenced site 435 mutations (Fig. 2A, lanes 6 and 7). Therefore, we concluded that the Tn*5* elements in these mutations lie at site 435.

Location of outward promoter of Tn*5*. The 5′ terminus of *recE* mRNA located immediately adjacent to Tn*5* implies that the responsible promoter lies within Tn*5*. Berg et al. (6) located an outward promoter in the distal 186 nucleotides of Tn*5*. One possibility is that this promoter is close enough to the outside ends of Tn*5* to initiate transcription at the nucleotide adjacent to the first nucleotide of Tn*5*. Another possibility is that the promoter is more internally located and RNase action produced a 5′ terminus adjacent to the first nucleotide of Tn*5*.

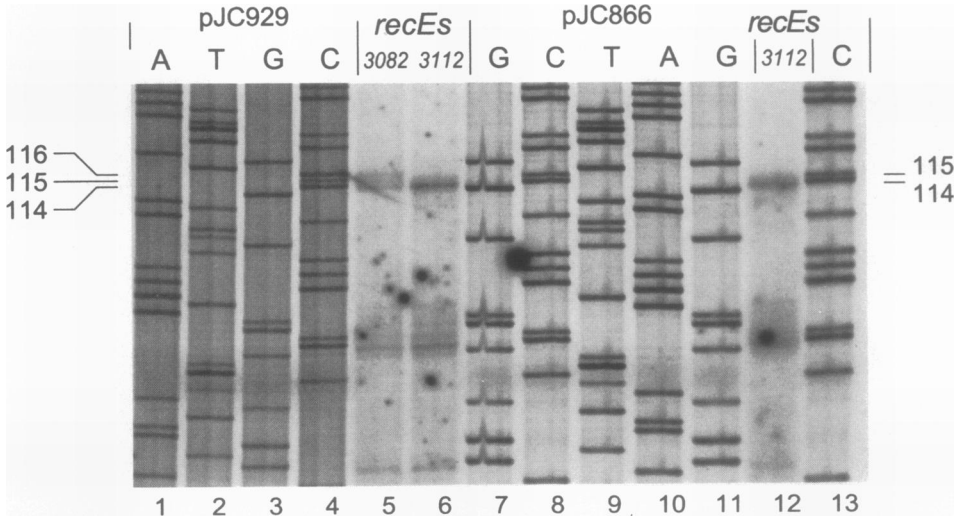


FIG. 3. Comparison of primer extension products from two Tn*5* mutations at site 425, i.e., *recEs3082* and *recEs3112*, containing 10- and 9-bp duplications, respectively. Sequencing ladders in lanes 1 to 4, 7 to 11, and 13 are derived from prJC11 and plasmid DNAs containing the cloned insertions; lanes 8 and 13 contain identical samples. Lanes 5, 6, and 12 contain DNA extended from prJC11 by using RNA templates isolated from mutant strains. Numbers on the left and right indicate lengths in nucleotides.

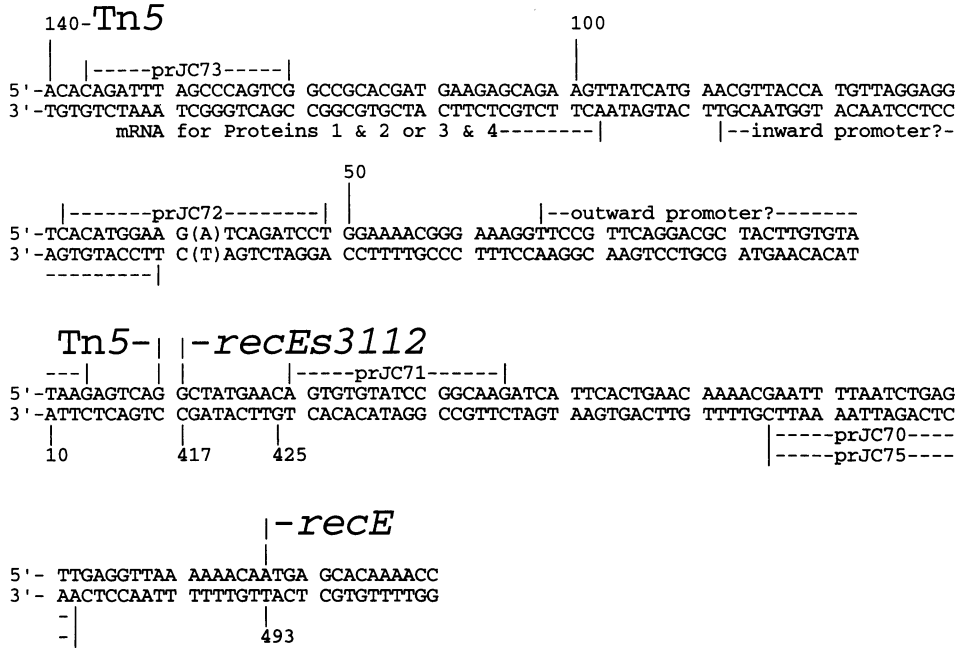


FIG. 4. Sequence of a portion of *recEs3112*::Tn*5*, showing the locations of five primers used for PCR cloning of regions of Tn*5* and *recEs*. The start of mRNA for proteins 1 and 2 or 3 and 4 at nucleotide 98 (Tn*5*) is taken from reference 19. Hypothetical inward and outward promoter locations were determined by eye. Tn*5* nucleotide coordinates are from reference 2; *recEs* and *recE* nucleotide coordinates are from reference 9. The nucleotide pair in parentheses between nucleotides 59 and 60 of Tn*5* was found in the fragments generated by PCR using both prJC72 and prJC73 with prJC75. It may stem from a difference between the Tn*5* used by Fouts et al. (13) and that sequenced by Auerswald et al. (2).

To locate the outward promoter more precisely, we used a PCR method to clone two distal regions of Tn*5* attached to the spacer region between one of our Tn*5* inserts (*recEs3112*::Tn*5*) and *recE*. As a control we cloned just the spacer region (Fig. 4). This did not lead to the expression of reporter gene *galK* in vector pK04 as judged by colony color on MacConkey galactose agar (not shown). However, both clones containing Tn*5* DNA in addition to the spacer DNA did lead to *galK* expression as judged by the purple colony color (data not shown). Thus, we conclude that there is an outward promoter located within the distal 69 nucleotides of Tn*5*.

Insertion sites of IS*50*. One IS*50* mutation (*recEs3083*::IS*50*) was cloned, sequenced, and found to lie at site 425. Primer extension analysis of RNA from the mutant strain did not show any major large band, however (Fig. 5A, lane 1). RNA from two other IS*50* mutants, carrying IS*50* in the same orientation as in *recEs3083*::IS*50* (i.e., orientation IIA [36]), did have major bands (Fig. 5A, lanes 2 and 3). These are long enough to include a large number of nucleotides transcribed from IS*50*. The orientation of IS*50* in all three of these mutations presents the *aphA* promoter in the correct orientation to transcribe toward *recE* (36). The *aphA* promoter is thought to comprise the sequence TTGCCA-17 nucleotides-TAAGGT (32). If transcription began at an adenine residue 6 nucleotides from this promoter, the RNA so formed would consist of 82 nucleotides copied from the IS*50* sequence (2). Using this assumption, we estimated that *recEs117*::IS*50* would be located at site 418 and that *recEs3115*::IS*50* would be located at site 427 (Fig. 1D and 2 and Table 2).

Three other IS*50* insertions (*recEs3087*::IS*50*, *recEs 3085*::IS*50*, and *recEs3118*::IS*50*) have the orientation opposite to that of the insertions just discussed. These present the same promoter sequence for *recE* transcription presented by Tn*5*. The lengths of their extension products are shown in Fig. 5A. The products from *recE3087*::IS*50* are the same length as those from sequenced mutation *recE3084*::Tn*5*, which lies at site 435 (data not shown). Hence, we infer that IS*50* in *recE3087* is at the same site (Table 2). The largest product from the other two IS*50* insertions is 7 nucleotides larger (Fig. 5A, lanes 4 and 6; compare with lane 5); hence, we infer that these insertions are located at site 428 (Fig. 1E and Table 2).

DISCUSSION

Nucleotide sequence and primer extension studies located the sites of 14 Tn*5* and IS*50* insertion mutations which lead to expression of *recE* (13, 36). In four cases in which both nucleotide sequence and primer extension analysis were done for the same strains, the data revealed RNA 5′ termini adjacent to the first nucleotide of Tn*5*. In one case RNA 5′ termini were both adjacent and one nucleotide within Tn*5*. This contradicts the suggestion of Willis et al. (36) that Tn*5* insertion creates a new promoter by fusing a −35 sequence in Tn*5* to a −10 sequence in the adjacent DNA. Instead it is consistent with the contention of Berg et al. (6) that there is a weak outward transcribing promoter within the terminal 186 nucleotides of Tn*5*. We narrowed that region to 69 nucleotides.

The 5′ RNA termini which we detected could have been formed by RNase action or by transcription initiation. Further work is required to distinguish between these possibilities. Nonetheless, we would like to speculate that the sequence 5′ TTCCGT-17 nucleotides-TATAAG 3′ is a possible candidate for a promoter. This sequence terminates 6 nucleotides from the terminal nucleotide of Tn*5* and IS*50* (2) and consequently could have initiated transcription at the first nucleotide adjacent to Tn*5*/IS*50* as we found. This sequence yields a score of 40 on the scale of Mulligan et al. (28), which would mark it as a weak promoter.

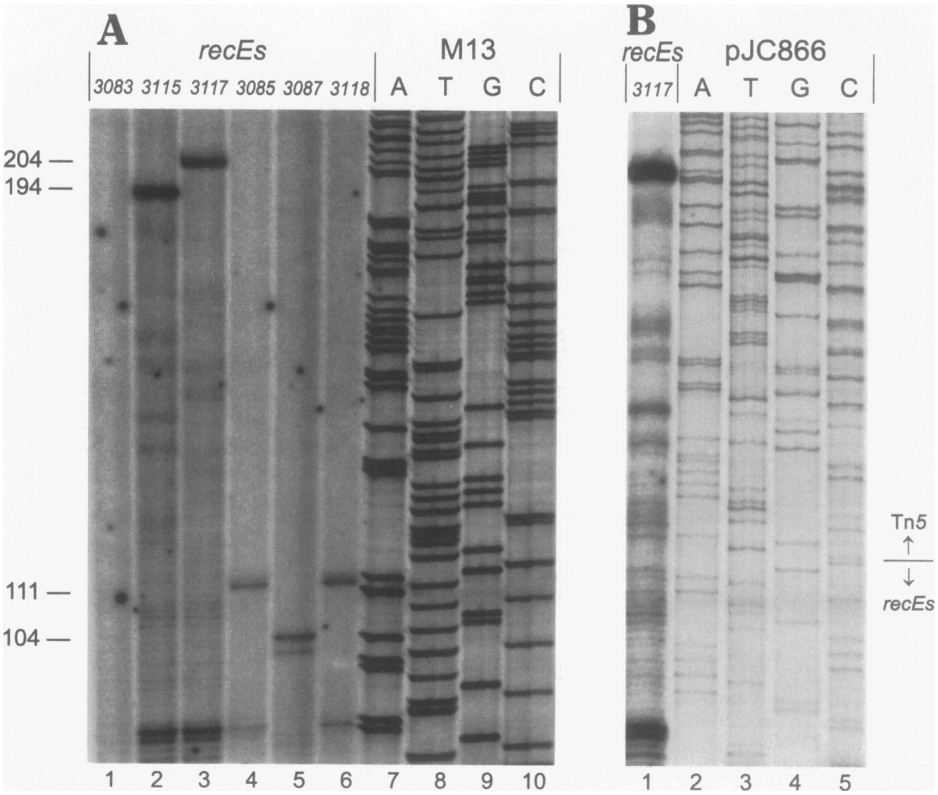


FIG. 5. Primer extension analysis of IS*50* insertion mutations of *recEs*. Measurements were made by using an M13 sequencing ladder (A) and a plasmid *recE3112*::Tn*5* (site 425, 9-bp duplication) sequencing ladder (B). Numbers along the left margin refer to the length (in nucleotides) of the largest DNA extension products visible in lanes 2 to 6. In the right margin the boundary between *recEs* and Tn*5* in the sequence of pJC866 is marked.

All 14 Tn*5* and IS*50* insertions described by Fouts et al. (13) and Willis et al. (36) were found to be located within a 19-nucleotide region. There are several explanations for the small size of the region. Prominent among these is avoidance of the dyad symmetry upstream from the insertions. This sequence consists of an 11-nucleotide-pair hyphenated inverted repeat located in the spacer (*recEs*) between *recE* and *racC* (9) and could be a terminator sequence (31). Equally prominent is the possibility that expression of the weak outward promoter of Tn*5* requires an appropriately placed downstream transcriptional enhancer-like sequence. Such a requirement would rationalize the fact that only one-third of Tn*5* insertions show expression of the outward promoter (6). Other possible reasons for the narrow region for productive insertions include preferential insertion sites and instability of *recE* transcripts initiated elsewhere.

The IS*50* insertion sequenced (*recEs3083*::IS*50*) belongs to orientation group IIA of Willis et al. (36). If a promoter internal to IS*50* is operative in this insertion, it cannot be an outward promoter of Tn*5* because in orientation IIA that promoter of IS*50* would direct transcription away from *recE*. Presumably the promoter operative in orientation IIA is the internal promoter which differentiates the two copies of IS*50* in Tn*5* (20). Although we were unable to clone two other IS*50* insertions in group IIA, we did observe 5′ RNA termini consistent with use of the internal promoter. Failure to observe a 5′ RNA terminus in RNA from the *recEs3083*::IS*50* (IIA) strain is consistent with the weak expression of *recE* in that mutant relative to the strong expression in the other two IS*50* (IIA) mutants (13, 36). Willis et al. (36) hypothesized that the weak expresser contains IS*50*R and the strong expressers contain IS*50*L. Because the putative insertion sites of IS*50* in the two strongly expressing strains differ from that in the weakly expressing strain, however, we cannot conclude that the differences in expression derive from genetically different copies of IS*50*.

In summary, this study confirms the suggestion of Willis et al. (36) that the Tn*5* and IS*50* insertions which activate expression of *recE* do not reveal the location or mechanism of action of *sbcA* point mutations. Instead the insertions confirm that the addition of an appropriate promoter, such as p_L of lambda (18, 19) or *bla-p2* of pBR322 (37), can activate expression of *recE*.

ACKNOWLEDGMENTS

This work was supported in part through the Department of Genetics by National Institutes of Health predoctoral institutional training grant GM07127, by Public Health Services grant AI05371, and by grant NP-237 from the American Cancer Society.

We thank Doug Berg and Bill Reznikoff for up-to-date information on the promoters of Tn*5*; Richard Harland, Kathy Flynn, and Bryan Rogers for advice on primer extension; Jessica Geller and Marjorie Wagman for help with cloning; members of our laboratory (especially Ann Templin) for technical support; Jean-Pierre Bouché for computer analysis of the putative outward Tn*5* promoter; Nelle Neighbor-Alonzo for preparation of the manuscript; and God for making all things possible.

REFERENCES

1. **Alvarado-Urbina, G., G. M. Sather, W.-C. Liu, M. F. Gillen, P. D. Duck, R. Bender, and K. K. Ogilvie.** 1981. Automated synthesis of gene fragments. Science **214**:270–274.

2. **Auerswald, E.-A., G. Ludwig, and H. Schaller.** 1980. Structural analysis of Tn*5*. Cold Spring Harbor Symp. Quant. Biol. **45:**107–113.
3. **Bachmann, B. J.** 1983. Linkage map of *Escherichia coli* K-12, edition 7. Microbiol. Rev. **47:**180–230.
4. **Bachmann, B. J., and K. B. Low.** 1980. Linkage map of *Escherichia coli* K-12, edition 6. Microbiol. Rev. **44:**1–56.
5. **Barbour, S. D., H. Nagaishi, A. Templin, and A. J. Clark.** 1970. Biochemical and genetic studies of recombination proficiency in *Escherichia coli*. II. Rec+ revertants caused by indirect suppression of Rec− mutations. Proc. Natl. Acad. Sci. USA **67:**128–135.
6. **Berg, D. E., A. Weiss, and L. Crossland.** 1980. Polarity of Tn*5* insertion mutations in *Escherichia coli*. J. Bacteriol. **142:**439–446.
7. **Chen, E. Y., and P. H. Seeberg.** 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4:**165–170.
8. **Chu, C. C., and A. J. Clark.** 1989. A 10- rather than 9-bp duplication associated with insertion of Tn*5* in *Escherichia coli* K-12. Plasmid **22:**260–264.
9. **Chu, C. C., A. Templin, and A. J. Clark.** 1989. Suppression of a frameshift mutation in the *recE* gene of *Escherichia coli* K-12 occurs by gene fusion. J. Bacteriol. **171:**2101–2109.
10. **Clark, A. J., and K. B. Low.** 1988. Pathways and systems of homologous recombination in *Escherichia coli*, p. 155–215. *In* K. Low (ed.), The recombination of genetic material. Academic Press, San Diego.
11. **Davis, R. W., D. Botstein, and J. R. Roth.** 1980. Advanced bacterial genetics, p. 227. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
12. **Demerec, M., E. A. Adelberg, A. J. Clark, and P. E. Hartman.** 1966. A proposal for a uniform nomenclature in bacterial genetics. Genetics **54:**61–76.
13. **Fouts, K. E., T. Wasie-Gilbert, D. K. Willis, A. J. Clark, and S. D. Barbour.** 1983. Genetic analysis of transposon-induced mutations of the Rac prophage in *Escherichia coli* K-12 which affect expression and function of *recE*. J. Bacteriol. **156:**718–726.
14. **Gillen, J. R., A. E. Karu, H. Nagaishi, and A. J. Clark.** 1977. Characterization of the deoxyribonuclease determined by lambda reverse as exonuclease VIII of *Escherichia coli*. J. Mol. Biol. **113:**27–41.
15. **Ish-Horowicz, D., and J. Burke.** 1981. Rapid and efficient cosmid cloning. Nucleic Acids Res. **9:**12989–12999.
16. **Joseph, J. W., and R. Kolodner.** 1983. Exonuclease VIII of *Escherichia coli*. I. Purification and physical properties. J. Biol. Chem. **258:**10411–10417.
17. **Joseph, J. W., and R. Kolodner.** 1983. Exonuclease VIII of *Escherichia coli*. II. Mechanism of action. J. Biol. Chem. **258:**10418–10424.
18. **Kaiser, K., and N. Murray.** 1979. Physical characterization of the "Rac prophage" in *E. coli* K12. Mol. Gen. Genet. **175:**159–174.
19. **Kaiser, K., and N. E. Murray.** 1980. On the nature of *sbcA* mutations in *E. coli* K12. Mol. Gen. Genet. **179:**555–563.
20. **Krebs, M. P., and W. S. Reznikoff.** 1986. Transcriptional and translational initiation sites of IS*50*: control of transposase and inhibitor expression. J. Mol. Biol. **192:**781–791.
21. **Kushner, S., H. Nagaishi, and A. J. Clark.** 1974. Isolation of exonuclease VIII: the enzyme associated with the *sbcA* indirect suppressor. Proc. Natl. Acad. Sci. USA **71:**3593–3597.
22. **Lloyd, R. G., and S. D. Barbour.** 1974. The genetic location of the sbcA gene of *Escherichia coli*. Mol. Gen. Genet. **134:**157–171.
23. **Low, B.** 1973. Restoration by the *rac* locus of recombinant forming ability in *recB*$^-$ and *recC*$^-$ merozygotes of *Escherichia coli* K-12. Mol. Gen. Genet. **122:**119–130.
24. **Luisi-DeLuca, C., A. J. Clark, and R. D. Kolodner.** 1988. Analysis of the *recE* locus of *Escherichia coli* K-12 by using polyclonal antibodies to exonuclease VIII. J. Bacteriol. **170:**5797–5805.
25. **Mahajan, S. K., C. C. Chu, D. K. Willis, A. Templin, and A. J. Clark.** 1990. Physical analysis of spontaneous and mutagen induced mutants of *Escherichia coli* K-12 expressing DNA exonuclease VIII activity. Genetics **125:**261–273.
26. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
27. **McKnight, S. L., E. R. Gavis, R. Kingsbury, and R. Axel.** 1981. Analysis of transcriptional regulatory signals of the HSV thymidine kinase gene: identification of an upstream control region. Cell **25:**385–398.
28. **Mulligan, M. E., D. K. Hawley, R. Entriken, and W. R. McClure.** 1984. *Escherichia coli* promoter sequences predict in vitro RNA polymerase selectivity. Nucleic Acids Res. **12:**789–800.
29. **Mullis, K., and F. Faloona.** 1987. Specific synthesis of DNA *in vitro* via a polymerase catalysed chain reaction. Methods Enzymol. **155:**335–350.
30. **Pansegrau, W., L. Miele, R. Lurz, and E. Lanka.** 1987. Nucleotide sequence of the kanamycin resistance determinant of plasmid RP4: homology to other aminoglycoside 3′-phosphostransferases. Plasmid **18:**193–204.
31. **Platt, T.** 1986. Transcription termination and the regulation of gene expression. Annu. Rev. Biochem. **55:**339–372.
32. **Rothstein, S. J., and W. S. Reznikoff.** 1981. The functional differences in the inverted repeats of Tn*5* are caused by a single base pair nonhomology. Cell **23:**191–199.
33. **Smith, G. R.** 1988. Homologous recombination in prokaryotes. Microbiol. Rev. **52:**1–28.
34. **Summers, W.** 1991. Preparation of bacterial RNA, p. 4.4.4–4.4.7. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology, vol. 1, suppl. 15. Greene Publishing Associates and Wiley-Interscience, New York.
35. **Willetts, N. S., A. J. Clark, and B. Low.** 1969. Genetic location of certain mutations conferring recombination deficiency in *Escherichia coli*. J. Bacteriol. **97:**244.
36. **Willis, D. K., K. E. Fouts, S. D. Barbour, and A. J. Clark.** 1983. Restriction nuclease and enzymatic analysis of transposon-induced mutations of the Rac prophage which affect expression and function of *recE* in *Escherichia coli* K-12. J. Bacteriol. **156:**727–736.
37. **Willis, D. K., L. H. Satin, and A. J. Clark.** 1985. Mutation-dependent suppression of *recB21* and *recC22* by a region cloned from the Rac prophage of *Escherichia coli* K-12. J. Bacteriol. **162:**1166–1172.
38. **Woods, D.** 1984. Oligonucleotide screening of cDNA libraries. Focus **6:**1–3.