JOURNAL OF BACTERIOLOGY, July 1985, p. 376–384
0021-9193/85/070376-09$02.00/0

Vol. 163, No. 1

Nucleotide Sequence Binding Specificity of the LexA Repressor of *Escherichia coli* K-12

KENNETH F. WERTMAN AND DAVID W. MOUNT*

Department of Molecular and Cellular Biology, University of Arizona, Tucson, Arizona 85721

Received 13 December 1984/Accepted 1 April 1985

The specificity of LexA protein binding was investigated by quantifying the repressibility of several mutant *recA* and *lexA* operator-promoter regions fused to the *Escherichia coli* galactokinase (*galK*) gene. The results of this analysis indicate that two sets of four nucleotides, one set at each end of the operator (terminal-nucleotide contacts), are most critical for repressor binding. In addition, our results suggest that the repressor-operator interaction is symmetric in nature, in that mutations at symmetrically equivalent positions in the *recA* operator have comparable effects on repressibility. The symmetry of this interaction justified reevaluation of the consensus sequence by half-site comparison, which yielded the half-site consensus (5')CTGTATAT. Although the first four positions of this sequence were most important, the last four were well conserved among binding sites and appeared to modulate repressor affinity. The role of the terminal-nucleotide contacts and the mechanism by which the internal sequences affected repressor binding are discussed.

A large number of proteins have been identified which bind to DNA, performing structural, enzymatic, or regulatory roles (3). The recognition events involved can be relatively nonspecific (independent of nucleotide sequence), as in the case of histones or procaryotic histonelike proteins (11), or the binding site may be specified by some novel structural feature of the nucleic acid (site specific). Site-specific interactions contain at least two components: a nonspecific interaction that is primarily due to electrostatic interactions with the backbone phosphates and a site-specific interaction which results from a matrix of DNA hydrogen bond donors and acceptors located in the DNA helical grooves, which interact with a complementary set of hydrogen bond acceptors and donors on the protein surface (30, 32). Several DNA-binding proteins have been shown to bind with specificity to multiple sites. The binding affinity for these sites varies, presumably as a function of the number of hydrogen bond pairs and van der Waals contacts formed in the complex (9).

We have studied the specificity of binding of the *Escherichia coli* LexA protein. The *lexA* gene of *E. coli* K-12 encodes a small protein (molecular weight, 22,300) which is responsible for the repression of several unlinked operons (15). The various affinities with which the repressor binds to these sites assist in establishing different levels of transcription for these operons in the uninduced state (12). The nucleotide sequence of several LexA repressor binding sites are known, and the compiled sequence data provide strong evidence for a 16-base-pair recognition site with approximate twofold rotational symmetry. Of these, only three nucleotides at each end of the palindrome (which we shall term the terminal trinucleotides) are highly conserved among binding sites.

We (37) and others (4, 8) previously identified nucleotide substitutions which result in the derepression of the *E. coli* *recA* and *lexA* promoters. Mutations which changed base pairs in the LexA protein-binding sites were inferred to act at the level of transcription initiation by decreasing the affinity of the repressor for these sites. We believed for several reasons that a detailed analysis of the *recA* operator (*recAo*) would provide the most insight into repressor binding. First, the repressor demonstrates high affinity for this site (2). Second, of seven *recAo* mutations which affect operator affinity, only three are within the terminal trinucleotides (4, 8, 37). Finally, the *recA* operator has near perfect rotational symmetry (2, 16). This report summarizes studies which addressed the mechanism by which substitutions at the *recA* operator act and quantitates the effect of these substitutions on operator and promoter activity.

MATERIALS AND METHODS

Bacterial strains, plasmids, and bacteriophages. The *E. coli* strains, plasmids, and bacteriophages used in these experiments are listed in Table 1. All in vivo manipulations were done by standard genetic protocols (24). The phage λ *lacZop*::*lexA* was constructed in vitro (see below).

Media. Growth conditions of M13 derivatives for testing the β-galactosidase alpha-complementing activity and for the preparation of single- and double-stranded DNA have been previously described (37). Luria agar and broth and MacConkey agar were prepared by standard recipes (24). Cells were grown in M9 minimal salts medium (24) supplemented with Casamino acids (Difco Laboratories, Detroit, Mich.) (0.5%), glycerol (0.5%), and ampicillin (100 μg/ml) for determination of galactokinase activity. Microbiological medium was purchased from Difco. The chromogenic reagent 5-bromo-4-chloro-3-indoyl-β-D-galactoside (X-gal) was obtained from Bachem Inc., Torrance, Calif. D-[1-^{14}C]galactose (CFA 435) was purchased from Amersham Corp., Arlington Heights, Ill. [α-^{32}P]dATP (800 Ci/mmol) was from New England Nuclear Corp., Boston, Mass. All other reagent-grade chemicals were purchased from Sigma Chemical Co., St. Louis, Mo.

In vitro constructions. All enzymes were purchased from New England BioLabs, Inc., Beverly, Mass., and used as prescribed by the manufacturer, except as otherwise specified. All nucleic acids used in in vitro constructions were

* Corresponding author.

TABLE 1. Bacterial strains, plasmids, and phage

Strain, plasmid or phage	Relevant features	Reference or source
E. coli strain no.		
JM103	Δ(*lac-pro*) *supE* F'*proAB lacI*^{Q} *lacZ* Δ*M15*	J. Messing (23)
UT481	*E. coli* B/K-12 hybrid; strong alpha acceptor activity	C. Lark
DM1415	*recA1 lexA51 galK2 sulA211*	Laboratory strain
DM3101	DM1415 λ *imm*^{434} cI *ind lexA*$^{+}$	This work
DM3102	DM1415 F'*lacI*^{Q1} *lacL8 lacZ4505*::Tn*5* λ *lacZop*::*lexA*$^{+}$	This work
N100	*recA13 galK2*	K. McKenney (21)
GC4572	Δ(*lac-pro*) *sfiA*::*lacZ lexA51*	R. D'Ari (12)
594	*sup*$^{+}$ *lacZ*	
Plasmid		
pJWL42	Ap^{r} *lexA*$^{+}$	J. Little (37) (formerly pJL42)
pKO1	Ap^{r} *galK*	K. McKenney
pDE-FL54	F'*lacI*^{Q1} *lacL8 lacZ4505*::Tn*5* *proA*$^{+}$*B*$^{+}$	D. Ennis
Phage		
M13mp8/P	Promoterless *lacZ*α	37
M13recAop::lacZα	*recAop*::*lacZ*α fusion (also mutant derivatives)	37
M131exAop::lacZα	*lexAop*::*lacZ*α fusion	37
λNM426	*c*Iam *Eco*RI fragment, deletion (44.5 to 54.3%) mutant in other *Eco*RI sites	N. Murray via K. Mizzuchi
λlacZop::lexA	*lexA* under *lacZop* control	This work

purified by standard techniques as previously described (13, 37).

Mutagenesis of the *recA* operator with synthetic oligonucleotides was performed essentially as described by Messing (23), but under the polymerization conditions of Zoller and Smith (38). A synthetic 20-mer (Applied Biosystems model 380A DNA synthesizer) was hybridized to single-stranded phage DNA (1 μg). This hybrid, which contained a central mismatch, was hybridized to a second oligonucleotide with homology to codons 11 to 16 in the *lacZ* gene (molar ratio of each primer/template, 20:1). This tribrid was treated for 12 h at 16°C with DNA polymerase I (Klenow fragment; 2 U) and T4 DNA ligase (200 U) in a 20-μl reaction mixture which contained 10 mM Tris hydrochloride (pH 7.5), 5 mM dithiothreitol, 0.5 mM ATP, and 1 mM each dATP, dGTP, dCTP, and dTTP. A 1-μl portion of this mixture was used to transform $CaCl_2$-treated cells (13).

Plasmids bearing transcriptional fusions of the *recA* regulatory region to the *E. coli galK* gene were constructed and analyzed by the procedures of McKenney et al. (21). DNA of plasmid pKO1 (100 ng) was admixed with 200 ng of M13 replicative-form I DNA (M13 *recAop*::*lacZ*α or a mutant derivative) and completely digested with *Eco*RI and *Hin*dIII. After phenol extraction and ethanol precipitation, this mixture was dissolved in buffer, treated with T4 DNA ligase, and used to transform competent *E. coli* N100 cells. Transformed cells were plated on MacConkey galactose medium containing ampicillin (100 μg/ml). After incubation at 37°C for 36 h, Gal^{+} isolates were used for plasmid DNA purification and restriction analysis.

A bacteriophage containing a *lacZop*::*lexA* fusion was constructed by inserting an *Eco*RI fragment from plasmid pJWL46 (J. Little, unpublished data) into the vector λNM426. This fragment contains a fusion of the *lacZ* operator and promoter to the *E. coli lexA* structural gene. *Eco*RI-digested plasmid and phage DNAs were mixed (molar ratio, 5:1), treated with T4 DNA ligase, and encapsidated in vitro (5). These phage were plaqued on strain N100, and isolates containing the *lexA* gene were identified by plaque hybridization (5). One such isolate was crossed with λ *imm*^{4} cI *ind lac*5, and a recombinant which made a turbid, Lac^{-} plaque on *E. coli* 594 was chosen. This isolate was shown to repress the *sfiA*::*lacZ* fusion in GC4572 and was designated λlacZop::lexA.

Bisulfite mutagenesis of the *lexA* operator. The deamination of single-stranded DNA was performed as previously described (33). Phage M13 *lexAop*::*lacZ*α (60 μg) in 1 ml of citrate buffer was mixed with 1 ml of 4 M sodium bisulfite and 80 μl of hydroquinone. After incubation (37°C, 8 h), 0.6 ml of the mixture was dialyzed as previously described (33). This deaminated DNA (10 μg in 50 μl) was hybridized to a synthetic oligonucleotide with homology to the *lacZ* gene, as described above. This hybrid was treated (37°C, 30 min) with DNA polymerase I (Klenow fragment, 12 U) in a mixture containing 10 mM Tris hydrochloride (pH 7.5), 5 mM dithiothreitol, 1 mM each dATP, dGTP, dCTP, and dTTP, and 30 μCi of [α-^{32}P]dATP. DNA polymerase was heat inactivated (90°C, 15 min), and 5 μl of concentrated restriction enzyme buffer (800 mM Tris hydrochloride [pH 7.4], 80 mM MgCl, 20 mM 2-mercaptoethanol, 1 mg of bovine serum albumin per ml) was added. The DNA was digested with *Eco*RI and *Hin*dIII endonucleases, and the reaction products were separated by polyacrylamide gel electrophoresis (18). A mixture of mutagenized 140-base-pair fragments was purified by using DEAE-cellulose paper (Schleicher & Schuell, Inc., Keene, N.H.), ligated into M13mp8/P (37), and used to transform strain JM103. Because bisulfite-induced C-to-T transitions in the template will result in G-to-A transitions to the complementary strand, we screened the clones only for the appearance of new dideoxyadenosine termination sites. The complete nucleotide sequence of the inserted fragment was determined for all isolates with mutations.

Galactokinase assays. The galactokinase assay procedures outlined by McKenney et al. (21) were used here, with the exception of the growth medium used. Overnight cultures in M9 medium (see above) were diluted 1/100 into fresh me-

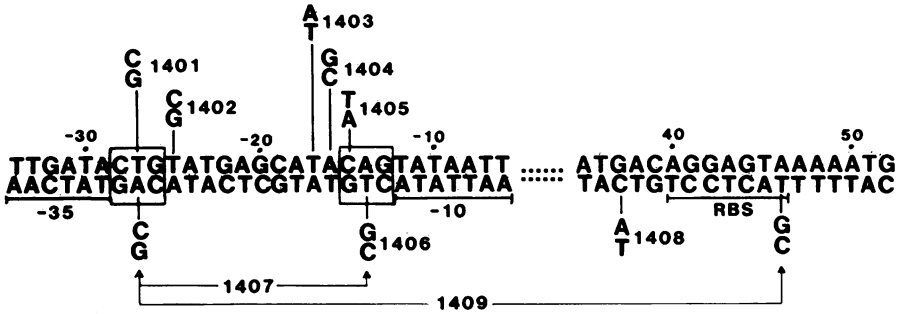

FIG. 1. Mutations in the regulatory region of the *recA* gene. The DNA sequence of the wild-type *recA* regulatory region is shown with the following notations: promoter elements (−10 and −35) and the ribosome-binding site are underlined, and the terminal trinucleotides of the operator are boxed. Nucleotide substitutions indicated above the sequence were identified after propagation in a mutator host (*mutD5*), and those shown under the sequence were introduced by oligonucleotide mutagenesis (see the text). The alleles *recAo1407* and *recAo1409* are double mutants (see the text). The coding sequence of *recA* begins at position 50.

dium, grown to an optical density of 0.6 at 650 nm and assayed.

Plate test for M13 α-complementing activity. A 2-μl drop of a phage lysate (10^{10} phage per ml) was spotted on plates (yeast extract-tryptone agar) that were seeded with *E. coli* UT481 cells in an overlay agar which contained X-gal and isopropyl-β-D-thiogalactopyranoside (IPTG) at the described concentrations (23). These plates were incubated at

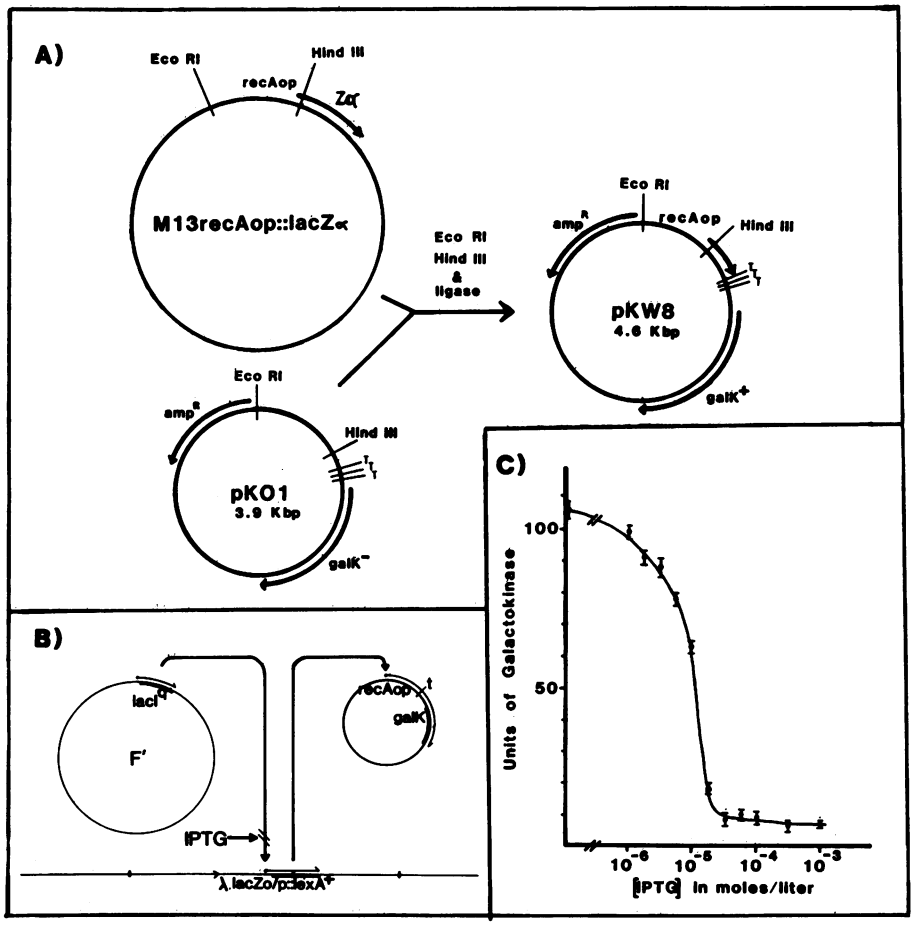

FIG. 2. Construction and use of *recAop*::*galK* fusion plasmids. (A) General strategy for the construction of the *galK* fusion plasmids. The derivation and properties of the phage M13 *recAop*::*lacZ* have been described previously (37). Details of the cloning are in the text. (B) Diagrammatic representation of the artificial regulatory circuit devised to test the effect of mutations in the operator upon the repressibility of the *recA* promoter. (C) Effect of IPTG addition on the expression of the wild-type *recAop*::*galK* fusion in cells containing an artificial regulatory circuit (DM3102 in panel B).

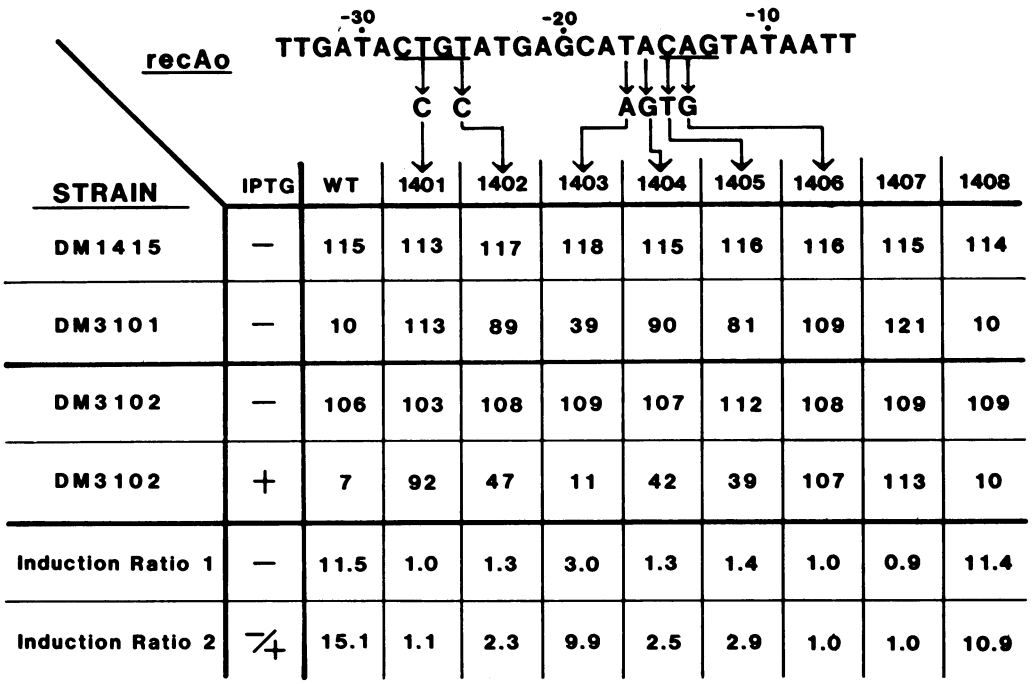

FIG. 3. Effect of mutations in the *recA* regulatory region on their in vivo repressibility by LexA protein. Steady-state levels of galactokinase were measured for each plasmidborne *recA* mutation (vertical columns) in strains expressing different levels of LexA repressor (horizontal rows): DM1415 (*lexA* [defective]); DM3101 (*lexA*$^+$); DM3102 (LexA synthesis under *lacI* control; see Fig. 2B). Galactokinase activities are expressed in picomoles of galactose phosphorylated per minute per optical density unit at 650 nm. Induction ratio 1 is the kinase activity expressed by a plasmid in strain DM1415 divided by its activity in strain DM3101. Induction ratio 2 is the ratio of the activities of a plasmid in strain DM3102 with and without IPTG (10^{-3} M). All determinations are the average of at least three experiments (each in triplicate), and a single standard error on any value was not greater than 7%.

37°C for 24 h. The blue color intensity of mutant phage spots was then compared with the reactions of control phage to test for M13 α-complementing activity.

RESULTS

Mutagenesis of the *recA* operator. In previous work we identified five nucleotide substitutions in the *recA* operator which result in an operator-constitutive phenotype *recAo1401-1405*, (Fig. 1). To examine the role of symmetry in repressor-operator interaction, we obtained additional mutations in the *recA* operator by site-directed mutagenesis. A synthetic oligonucleotide was hybridized to the single-stranded DNA of an M13 phage which carries a fusion of the *recA* regulatory region and the N-terminal coding sequence of a fragment from the *E. coli lacZ* gene (alpha donor peptide gene). This hybrid, which would serve as primer-template for DNA synthesis, extended from positions −5 to −24 relative to the mRNA start site and contained a central base pair mismatch at position −14. After treatment with DNA polymerase I and transformation of JM103(pJWL42), 6 isolates were chosen (of 5,000 total) that demonstrated constitutive expression of alpha donor activity. DNA sequence analysis confirmed the expected nucleotide substitution (*recAo1406*) in four of the six. This A · T-to-G · C transition was related by rotational symmetry to the T · A-to-C · G transition at position −27 (*recAo1401*). The other two isolates contained unpredicted substitutions, one at position −16 (previously isolated *recAo281* [4]) and the other at position 37 (*recA1408*).

A symmetric doubly mutant operator containing both the *recAo1401* and *recAo1406* mutations was constructed by using the mismatched oligonucleotide to prime DNA synthesis on a single-stranded phage chromosome that already contained the *recAo1401* mutation. Using the same screening procedure, we chose 6 isolates (of 5,000 total) with stronger alpha donor activity. DNA sequence analysis of these isolates revealed that all of them retained the template mutation but only one had obtained the *recAo1406* substitution. This double mutant was designated *recAo1407*. Of the other five, one had an unpredicted substitution at position 46 (*recA1409*). We were unable to identify additional mutations in the other four isolates (sequence data spanned positions 65 to −80 relative to the mRNA start site).

In vitro construction of *recAop*::*galK* operon fusions. McKenney et al. (21) described the construction and use of a multicopy cloning vector, pKO1 (Fig. 2). This vector provides unique restriction endonuclease cleavage sites for the insertion of promoters in front of a promoterless *galK* gene. Translation of this gene is dependent on the *galK* ribosome-binding site, and expression of the gene is unaffected by translation initiation originating within the inserted promoter fragment. Therefore, at steady state, the amount of galactokinase present in cells is proportional to the activity of the inserted promoter. The construction of plasmids containing transcriptional fusions of the *recA* regulatory region to the *galK* gene is diagrammed in Fig. 2A. Plasmids used and their regulatory mutation sites were as follows: pKW8, wild type; pKW872, *recAo1401*; pKW865, *recAo1402*; pKW863, *recAo1403*; pKW826, *recAo1404*; pKW860, *recAo1405*; pKW880, *recAo1406*; pKW890, *recAo1407*; and pKW843, *recAo1408*.

Phenotypes of the *recAop*::*galK* plasmids. Galactokinase expression from *recAop*::*galK* plasmids is sensitive to the levels of LexA protein present in the cell and thus provides

A)

		1/2 Site 1 2 3	4 5 6 7 8	ΔC
recA	Left	C T G	T A T G A	4
	Right		T A T G C	
lexA1	L		T A T A T	3
	R		T G A G T	
lexA2	L		T A T A T	3
	R		G G T G T	
sulA	L		T A C A T	3
	R		T A T G G	
uvrA	L		T A T A T	3
	R		A A T G A	
uvrB	L		T T T T T	4
	R		G A T A A	
uvrD	L		T A T A T	2
	R		G G T A T	
umuDC1	L		C T G G C	8
	R		T T C T T	
umuDC2	L		T A T A T	2
	R		T T T T T	
mucAB	L		T A T A A	2
	R		T T T A T	
clo13	L		T G T A T	1
	R		T A T A T	
cI e1 1	L		T A T A T	3
	R		G T T T T	
cI e1 2	L	G T G	G T T A T	4
	R		T A C A T	
col1B	L		T A T A T	3
	R	A T G	G A T A C	

	1	2	3	4	5	6	7	8	8	7	6	5	4	3	2	1
	+	+	+	+	+	+	+	+	−	−	−	−	−	−	−	−
Consensus	C	T	G	T	A	T	A	T	A	T	A	T	A	C	A	G
	G	A	C	A	T	A	T	A	T	A	T	A	T	G	T	C
	−	−	−	−	−	−	−	−	+	+	+	+	+	+	+	+

B)

	1	2	3	4	5	6	7	8
G	1	0	28	6	4	1	7	1
A	1	0	0	1	17	1	17	4
T	0	28	0	20	7	23	4	20
C	26	0	0	1	0	3	0	3

FIG. 4. Derivation of an operator half-site consensus. (A) DNA sequences of several operator regions (19; reviewed in reference 36) are aligned with respect to their terminal-trinucleotide sequences. Left and right refer to the upstream and downstream halves of their dyad symmetries relative to the direction of transcription. The terminal trinucleotide CTG is present in both ends of all sites (except *cI*e*I* 2 and *coI*I b) but for the sake of clarity is not reiterated here. The parameter ΔC is the number of deviations from the consensus sequence in the entire binding site. (B) Frequency with which each nucleotide is found at each operator half-site position. Positions are numbered as in panel A.

an in vivo assay for the repressibility of the various mutant operators. When strain DM1415 was transformed with these plasmids, they all yielded strong Gal$^+$ reactions on MacConkey galactose medium because this strain does not produce active LexA protein (26). However, in the same strain lysogenized with λ *lexA*$^+$ (DM3101), the Gal phenotypes of these plasmids differed; the plasmid bearing *recAo*$^+$ was Gal$^-$ at 24 h, whereas all plasmids containing mutant operators were uniformly Gal$^+$ (*recAo1401-1407*).

This apparent inability to discriminate among *recA* operator mutants suggested that to detect residual repressibility of the mutant operators it might be necessary to construct a host that overproduced LexA protein. We therefore constructed the phage λ *lacZop*::*lexA* (see above) and used it to lysogenize a strain DM1415 derivative (DM3102) that harbored an F′ that overproduces the Lac repressor. Transformation of the strain with a *recA*::*galK* plasmid established an artificial regulatory circuit (Fig. 2B). The *lacI* gene product repressed the *lexA* gene, and the *lexA* gene product repressed *galK*. The effect of IPTG addition on the expression of galactokinase in strain DM3102 harboring a wild-type *recAop*::*galK* fusion plasmid is shown in Fig. 2C.

Repressibility of mutant *recA* operators. Using the above strains, we assayed the levels of galactokinase expression (Fig. 3).

In cells containing little or no LexA protein (strains DM1415 or DM3102 without IPTG), all of the plasmids directed the synthesis of comparable levels of galactokinase, indicating that none of these mutations affect promoter activity.

In cells containing normal levels of repressor (strain DM3101), plasmids containing wild-type operators were repressed (wild type and *recA1408*), yielding an induction ratio comparable to chromosomal *recA* mRNA after irradiation with UV light (19a, 22). All of the plasmids with mutations in the *recA* operator demonstrated an insensitivity to the repressor. The *recAo1401* and *recAo1406* mutations eliminated operator activity, as did the combination of these mutations (*recAo1407*). All other substitutions had lesser effects.

At high levels of LexA protein (strain DM3102 with IPTG), no residual repressor-binding affinity could be detected for operators containing the *recAo1401* and *recAo1406* mutations. The results for other mutations in the operator were affected quantitatively but not qualitatively.

Mutations which are related to one another by twofold rotational symmetry (*recAo1401* and *recAo1406* or *recAo1402* and *recAo1404*) had comparable effects on *galK* repressibility. These data indicate that *recAo1402* and *recAo1404* retained comparable repressor-binding activity. This conclusion cannot be drawn unequivocally for *recAo1401* or *recAo1406*, because residual repressor affinity for these sites was below our limit of detection. Nonetheless, the data strongly suggest that the repressor-operator complex was rotationally symmetric.

Half-site analysis of LexA binding sites. The symmetry inferred from the above results justified the reevaluation of the consensus binding sequence by half-site comparison. The assumption on which this analysis was based was that a repressor dimer binds symmetrically to the operator so that each monomer makes identical sets of contacts, one set on either side of the symmetry axis. Thus, the basic unit of recognition is the operator half site. In the case of a 16-base-pair LexA binding site, positions 1 through 8 were compared with the sequence that was complementary to positions 16 through 9. A comparison of 28 LexA binding half sites and the derivation of a half-site consensus are shown in Fig. 4. This analysis demonstrated that, although the terminal trinucleotide CTG (positions 1, 2, and 3) was greatly conserved, significant homology existed among positions that were more central to the binding sites. These conservations suggest that an ideal operator site would have the sequence shown at the bottom of Fig. 4A.

Chemical mutagenesis of the *lexA* operator. The LexA repressor bound less tightly to the two repressor binding sites (*lexA1o* and *lexA2o*) in the *lexA* operator, presumably because these sites had less homology to the consensus sequence. Specifically, there were four C · G base pairs, two in each of the *lexA* sites, that did not match the consensus. Transitions at any of these positions would have increased

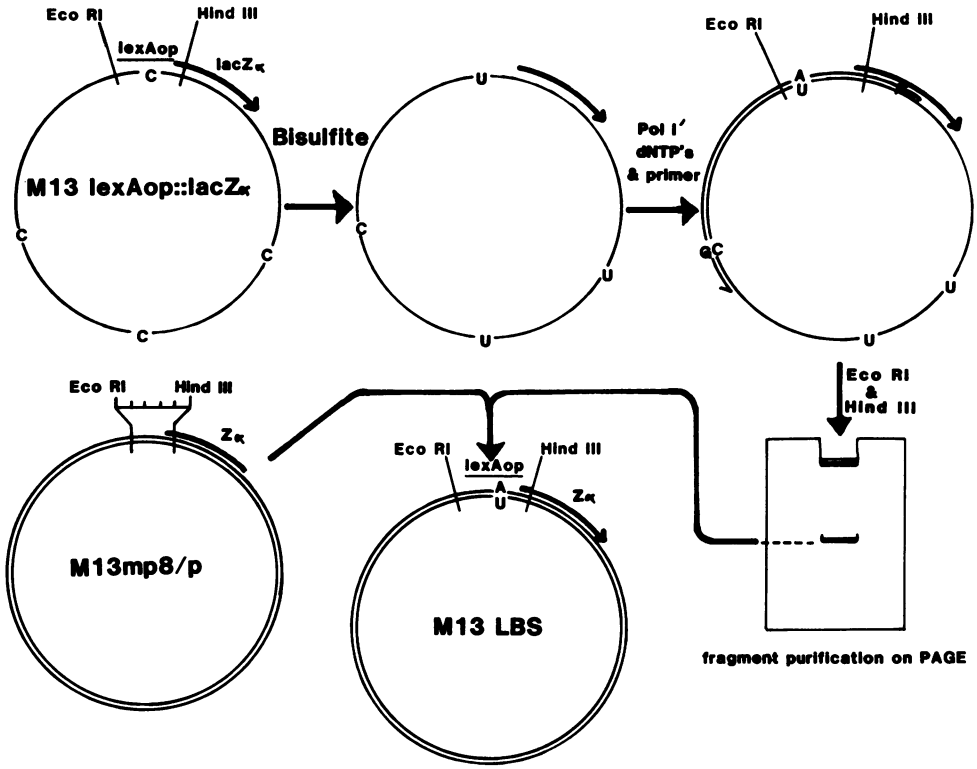

FIG. 5. Strategy for chemical mutagenesis of the *lexA* operator. Single-stranded M13 DNA bearing the *lexA* operator was treated with sodium bisulfite and copied with DNA polymerase I (Klenow fragment), and the synthetic duplex region was excised with restriction endonucleases. This mutagenized operator-bearing fragment was purified by polyacrylamide gel electrophoresis (PAGE) and then subcloned into an M13 promoter cloning vector (37). The product of this construction was a fusion of the N terminus of LexA protein to the alpha-complementing gene of M13mp8/P. See the text for details.

the homology of the binding site to the consensus and, if our modeling was correct, would have increased the binding affinity of the repressor. We therefore selectively deaminated cytosines in the sense strand of the *lexA* regulatory region by treating a recombinant M13 phage with sodium bisulfite (Fig. 5). Because this strategy resulted in the construction of an M13 phage containing fusions of the mutagenized *lexA* regulatory region to the alpha donor peptide gene, we were able to readily determine both the genotype and phenotype of the mutants (Fig. 6). Only C · G-to-T · A transitions were found. Two isolates (M13 LBS C11 and M13 LBS C15) were identified which contained transitions at two of the aforementioned positions (cytosines −6 [C,-6] and 14 [C,14]). Because these substitutions increased the homology of *lexA1o* or *lexA2o* to consensus, they were chosen for further study.

Construction and analysis of *lexAop*::*galK* fusions. The wild-type *lexA* operator-promoter fragment and the fragments bearing the C,−6 and C,14 mutations were subcloned into pKO1 (Fig. 2). Strain DM3102 cells bearing any one of these plasmids grew as small red colonies on MacConkey galactose medium and as white colonies on the same medium supplemented with 1 mM IPTG. At 0.1 mM IPTG, the plasmids bearing the C,−6, and C,14 mutations appeared to be more effective in repressing galactokinase, based on the following results: pKW12 (*lexAo*$^+$), red; pKW13 (*lexAo* C,−6) or pKW14 (*lexAo* C,14), pink. However, because these results could also be interpreted as a reduction in *lexA* promoter activity (C,−6 mutation changes Pribnow −10), we quantified the production of galactokinase from these plasmids in the absence (strain DM1415) and the presence (strain DM3101) of the LexA repressor (Table 2). The data indicate that the C,−6 mutation reduced the activity of the *lexA* promoter and increased its repressibility. The C,14 mutation had no effect on promoter activity, but like the C,−6 mutation, reduced kinase expression below that of the wild type in *lexA*$^+$ cells. These data provide direct support for the consensus sequence.

DISCUSSION

The specificity of LexA protein binding was first inferred from homologies in the *recA* and *lexA* regulatory regions and from the effect of the purified repressor on the enzymatic and chemical reactivities of these sequences. These studies indicated that the bound repressor covers approximately 26 base pairs of the DNA and makes close contact at positions in the major groove which are related by rotational symmetry (2, 16). We undertook these studies to obtain evidence identifying the position and relative importance of the nucleotide sequence determinants of LexA binding. Recently, the nucleotide sequence of several LexA binding sites have been determined. Here we summarize our results in light of these data, in the expectation that an overview of the repressor-half site interaction will provide a basis for assessing the affinity that the repressor might demonstrate for sites which have not yet been assayed or putative sites inferred from DNA sequence alone.

The following discussion considers each position in the

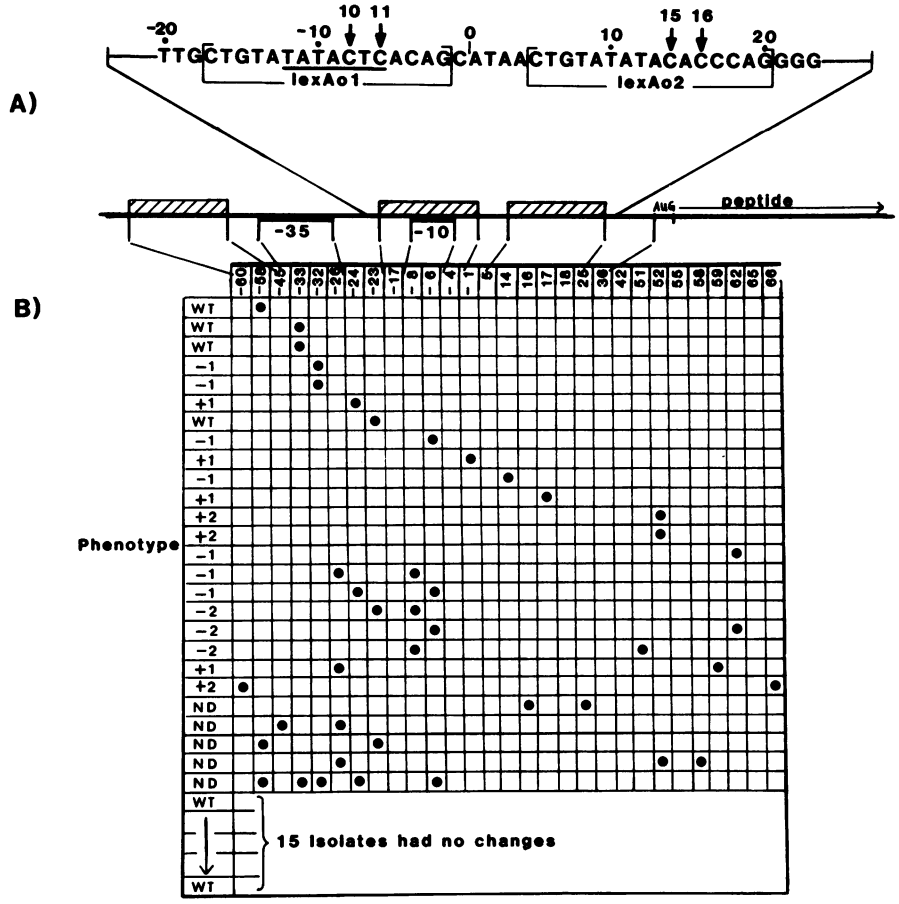

FIG. 6. Genotypes and phenotypes of phage-bearing mutations in the *lexA* regulatory region. (A) Diagrammatic representation of the *Eco*RI-*Hin*dIII fragment bearing the *lexA* regulatory region and DNA sequence of the sense strand of this fragment in the area of the *lexA* operators. The arrows above this sequence indicate the numeric position of cytosines which if deaminated would result in a closer match to consensus. Sequences with homology to *E. coli* RNA polymerase-binding sites (−10 and −35) and the N-terminal sequences of the hybrid alpha donor gene (peptide) are shown. In addition, sequences with homology to the LexA protein-binding-site consensus are indicated as hatched boxes. (B) Positions of C-to-T transitions (●) in 41 independent isolates after bisulfite mutagenesis. Each vertical column represents a cytosine residue (deamination target) in the fragment numbered relative to the mRNA start site. Each horizontal row is an independent isolate. The phenotype scored is the ability of the phage to provide alpha donor peptide to infected cells as indicated by the hydrolysis of X-gal in a plaque assay (see the text). The numbers used to describe the phenotypes reflect the intensity of blue plaque color: −2 (Lac$^-$) < −1 < WT (wild type) < +1 < +2. ND, Not determined.

half-site consensus sequentially. The reader is referred to the numbering scheme and nucleotide sequences in Fig. 4A.

Position +1 of LexA half sites is occupied by cytosine with 93% conservation (Fig. 4) and therefore appears to be important. Although no mutations at this position have been identified in the *recA* operator, substitutions at this position in either end of *lexA1o* result in a constitutive phenotype (37). In addition, binding places the repressor in close proximity to this site, as indicated by the protection of guanine −1 from alkylation by dimethyl sulfate (2).

A thymine at position +2 appears to be absolutely required for LexA binding. Introduction of cytosine at this position in the *recA* operator (*recAo1401* or *recAo1406*) eliminates repressibility in vivo (Fig. 3). As well, a T-to-A transversion at this site reduces repressor affinity by greater than 85-fold relative to the wild type (4). Position 2 is also important for binding at *lexA2o*, because the introduction of cytosine at position +2 results in a constitutive phenotype (37). Finally, all known half sites have a thymine at this position (Fig. 4).

Several lines of evidence confirm the significance of the complete conservation of guanine at position +3. A transition here virtually eliminates repressor binding at normal repressor concentrations (Fig. 3, *recAo1405* in strain DM3101). Our earlier isolation of *lexA* operator-constitutive mutants identified two independent substitutions in *lexA2o* at position 3 (37). Therefore, it is not surprising that the repressor bound to either *recAo* or *lexAo* protects this guanine from chemical alkylation (2). However, position 3 is

TABLE 2. Galactokinase activities of *lex*A operator-promoter fusion plasmids

Host strain	Plasmid		
	pKW12 (wild type)	pKW13 (*lexAo* C,−6)	pKW14 (*lexAo* C,14)
DM1415	26.5 ± 0.4	18.5 ± 1.5	26.8 ± 0.9
DM3101	3.0 ± 0.1	1.0 ± 0.3	0.9 ± 0.1
Induction ratio	8.8	18.5	29.8

not as important as position 2, because position 3 mutants demonstrate residual binding (cf. *recAo1405* to *recAo1401* or *recAo1406* in Fig. 3).

The high degree of conservation of sequence at positions 1 through 3 distinguishes them from positions 4 through 8 and suggests that these nucleotide pairs provide the primary contacts for LexA recognition. However, a thymine at +4 is as important as the +3 guanine, because transitions at +4 and +3 behave comparably (*recAo1405* versus *recAo1402* or *recAo1404*). This +4 substitution (*recAo1404* = *recAo281*) reduces repressor affinity approximately 35-fold (4). Consensus data indicate that thymine and guanine are preferred (71 and 21%, respectively) and adenine and cytosine are greatly underrepresented. A possible explanation for this disparity is that the former two nucleotides place a hydrogen bond acceptor deep in the major groove, while the latter two place a donor in this location. All operator-constitutive mutations at position +4 insert a cytosine (4, 8, 37; Fig. 1). The significance of this observation is unclear because this position in the downstream end of *recAo* may be a mutational hot spot (10 independent C-to-T transitions from three different mutagenic treatments).

Substitutions at position 5 had modest effects on repressor binding. The *recAo1403* mutation retained adequate binding to effect a 3-fold repression of the *recA* promoter at normal repressor levels and a 10-fold repression at high concentrations (Fig. 3, strains DM3101 and DM3102 plus IPTG). Similarly, a transition at this site (*recAo339*) reduces repressor affinity only 10-fold (4). A summary of LexA sites indicates that adenosine is preferred (61%). Thymine and guanine were found among binding sites at 25 and 14%, respectively, but the substitution of either of these for adenine +5 in the *recA* operator resulted in a constitutive phenotype (*recAo339*, *recAo340*, and *recAo1403*). Indeed, a guanine-to-adenine change in *lexA2o* appears to increase repressor affinity (Table 2). Cytosine was not found at the +5 position.

No mutations have been identified at half-site positions 6, 7, or 8. However, these positions were as well conserved as positions 4 and 5. Position +6 was usually occupied by thymine (82%), but all four nucleotides were represented. Position 7 was similar in conservation profile to position 5, including the absence of cytosine. Position 8 was comparable to position 6.

Three lines of evidence argue that positions 6, 7, and 8 contribute to repressor binding. First, *recAo* and *sulAo* differ in their affinity for repressor (25) but are identical in nucleotide sequence except at these positions. Second, the *clo-13* operator, which effects a 60-fold induction ratio (35), matches the consensus sequence except for a guanine at +5. Because the introduction of guanine at +5 in the *recAo* resulted in constitutivity, we infer that the perfect match of *clo-13* to consensus at positions 6, 7, and 8 allows for the position 5 deviation. Third, our analysis of the *lexA2o* C,14 mutation provides direct evidence that position 7 contributes to repressor binding (Table 2).

The high-resolution structures of three DNA binding proteins were recently determined by crystallography (1, 20, 27, 28), and from these data models have been proposed which infer a common secondary structural feature of these proteins (31, 34). Specifically, lambda *cro* and *c*I proteins and the *E. coli* catabolite gene activator protein all contain a helix-turn-helix motif (bihelical unit) which is responsible for site-specific binding. The spatial organization of this bihelical unit allows for the placement of the second helix (α3 in *c*I or *cro* and helix F in the catabolite gene activator protein) into the major groove of B-DNA, making specific contacts with base-pair edge groups. The first helix (α2 in *c*I or *cro* and helix E in the catabolite gene activator protein) is apparently involved in an electrostatic interaction with the DNA backbone phosphates and in maintaining the orientation of the other helix. Pabo and Sauer (29) have compared the protein sequences of several DNA-binding proteins and suggest an extensive conservation of the bihelical unit. Their alignment of the LexA repressor sequence predicts a helix-turn-helix arrangement which is in good agreement with computer-assisted secondary structural predictions (unpublished observations) and implies that site-specific binding of LexA protein to operator DNA involves an alpha helix-major groove interaction. This model is supported by the number and spatial arrangement of key repressor contacts in the *recA* operator (positions 1 through 4) in that an alpha-helical domain could interact favorably with 4 to 6 consecutive base pairs in B-DNA. If this model is correct, then the overall view of the LexA repressor-operator complex is very similar to that proposed for *cro* and *c*I binding in that (i) a repressor dimer binds to one side of a B-DNA helix, (ii) this complex demonstrates twofold rotational symmetry, and (iii) alpha helix-major groove interactions are responsible for site recognition.

The role of half-site positions 6, 7, and 8 is less certain. The known structural and functional similarities of LexA and *c*I repressors (14, 15) make it attractive to speculate that LexA also uses an extreme N-terminal arm to extend, via the major groove, around the helix toward the dyad axis. The N-terminal extension of LexA protein (i.e., before the putative bihelical sites) is highly charged and closer to the length of λ repressor than phage 434 repressor, which does not contain an N-terminal arm (29, 31). In this regard, the results of Brent and Ptashne (2) indicate a change in the reactivity of the major groove near the center of the *recA* binding site and the second *lexA* binding site. Alternately, position 6, 7, or 8 may confer subtle alterations upon the classical B-DNA structure (6, 7, 10, 17), which facilitate repressor interaction at positions 1 through 4. The final resolution of the interaction of LexA repressor with operator DNA awaits further genetic and structural characterization of LexA repressor function.

ACKNOWLEDGMENTS

We thank D. Ennis, J. Ito, J. Little, N. Ossanna, and K. Peterson for critically reading this manuscript. In addition, we thank C. Lark for *E. coli* UT481.

This work was supported in part by grant PCM8208440 from the National Science Foundation and by Public Health Service grant GM24496 from the National Institutes of Health.

LITERATURE CITED

1. **Anderson, W. F., D. H. Ohlendorf, Y. Takeda, and B. W. Matthews.** 1981. Structure of the cro repressor from bacteriophage λ and its interaction with DNA. Nature (London) **290:** 754–758.
2. **Brent, R., and M. Ptashne.** 1981. Mechanism of action of the *lexA* gene product. Proc. Natl. Acad. Sci. U.S.A. **78:**4204–4208.
3. **Champoux, J. J.** 1978. Proteins that affect DNA conformation. Annu. Rev. Biochem. **47:**449–479.
4. **Clark, A. J.** 1982. *recA* operator mutations and their usefulness. Biochimie **64:**669–675.
5. **Davis, R. W., D. Botstein, and J. R. Roth.** 1980. Advanced bacterial genetics, p. 130. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
6. **Dickerson, R. E., and H. R. Drew.** 1981. Structure of a B-DNA dodecamer. J. Mol. Biol. **149:**761–786.

7. **Drew, H. R., R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura, and R. E. Dickerson.** 1981. Structure of a B-DNA dodecamer: conformation and dynamics. Proc. Natl. Acad. Sci. U.S.A. **78**:2179–2183.
8. **Ginsburg, H., S. H. Edmiston, J. E. Harper, and D. W. Mount.** 1982. Isolation and characterization of an operator-constitutive mutation in the *recA* gene of *E. coli* K-12. Mol. Gen. Genet. **187**:4–11.
9. **Gussin, G. N., A. D. Johnson, C. O. Pabo, and R. T. Sauer.** 1983. Repressor and cro protein: structure, function, and role in lysogenization, p. 93–121. *In* R. W. Hendrix, J. W. Roberts, F. W. Stahl, and R. A. Weisberg (ed.), Lambda II. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
10. **Hogan, M., J. LeGrange, and B. Austin.** 1983. Dependence of DNA helix flexibility on base composition. Nature (London) **304**:752–754.
11. **Hubscher, U., H. Lutz, and A. Kornberg.** 1980. Novel histone H2A-like protein of *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. **77**:5097–5101.
12. **Huisman, O., R. D'Ari, and S. Casaregola.** 1982. How *Escherichia coli* sets different basal levels in SOS operons. Biochimie **64**:709–712.
13. **Kahn, M., R. Kolter, C. Thomas, D. Figurski, R. Meyer, E. Remaut, and D. Helinski.** 1979. Plasmid cloning vehicles derived from plasmids ColEl, F, R6K, RK2. Methods Enzymol. **68**: 268–280.
14. **Little, J. W.** 1984. Autodigestion of *lexA* and phage λ repressors. Proc. Natl. Acad. Sci. U.S.A. **81**:1375–1379.
15. **Little, J. W., and D. W. Mount.** 1982. The SOS-regulatory system of *Escherichia coli*. Cell **21**:11–22.
16. **Little, J. W., D. W. Mount, and C. R. Yanisch-Perron.** 1981. Purified *lexA* protein is a repressor of the *recA* and *lexA* genes. Proc. Natl. Acad. Sci. U.S.A. **78**:4199–4203.
17. **Lomonossoff, G. P., P. J. G. Butler, and A. Klug.** 1981. Sequence-dependent variation in the conformation of DNA. J. Mol. Biol. **149**:745–760.
18. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning, p. 173–181. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. **Mankovich, J. A., P. Lai, N. Gokul, and J. Konisky.** 1984. Organization of the colicin Ib gene. J. Biol. Chem. **259**:8764–8768.

19a. **Markham, B. E., J. E. Harper, and D. W. Mount.** 1985. Physiology of the SOS response: Kinetics of *lexA* and *recA* transcriptional activity following induction. Mol. Gen. Genet. **198**:207–212.

20. **McKay, D. B., and T. A. Steitz.** 1981. Structure of catabolite activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA. Nature (London) **290**:744–749.
21. **McKenney, K., H. Shimatake, D. Court, U. Schmeissner, C. Brady, and M. Rosenberg.** 1981. Gene amplification and analysis, p. 383. *In* J. G. Chirikjian and T. S. Papas (ed.), Structural analysis of nucleic acids. Elsevier/North-Holland Publishing Co., New York.
22. **McPartland, A., E. Green, and H. Echols.** 1980. Control of *recA* gene RNA in *E. coli*: regulatory and signal genes. Cell **20**:731–737.
23. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
24. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Mizusawa, S., D. Court, and S. Gottesman.** 1983. Transcription of the *sulA* gene and repression by LexA. Proc. Natl. Acad. Sci. U.S.A. **80**:358–362.
26. **Mount, D. W.** 1977. A mutant of *Escherichia coli* showing constitutive expression of the lysogenic induction and error prone repair pathways. Proc. Natl. Acad. Sci. U.S.A. **74**:300–304.
27. **Pabo, C. O., W. Krovatin, A. Jeffrey, and R. T. Sauer.** 1982. The N-terminal arms of λ repressor wrap around the operator DNA. Nature (London) **298**:441–443.
28. **Pabo, C. O., and M. Lewis.** 1982. The operator-binding domain of λ repressor: structure and DNA recognition. Nature (London) **298**:443–447.
29. **Pabo, C. O., and R. T. Sauer.** 1984. Protein-DNA recognition. Annu. Rev. Biochem. **53**:293–321.
30. **Record, M. T., T. M. Lohman, and P. L. deHaseth.** 1976. Ion effects on ligand-nucleic acid interactions. J. Mol. Biol. **107**:145–158.
31. **Sauer, R. T., R. R. Yocum, R. F. Doolittle, M. Lewis, and C. O. Pabo.** 1982. Homology among DNA-binding proteins suggests use of a conserved super-secondary structure. Nature (London) **298**:447–451.
32. **Seeman, N., J. Rosenberg, and A. Rich.** 1976. Sequence-specific recognition of double helical nucleic acids by proteins. Proc. Natl. Acad. Sci. U.S.A. **73**:804–808.
33. **Shortle, D., and D. Nathans.** 1978. Local mutagenesis: a method for generating viral mutants with base substitutions in preselected regions of the viral genome. Proc. Natl. Acad. Sci. U.S.A. **75**:2170–2174.
34. **Steitz, T. A., D. H. Ohlendorf, D. B. McKay, W. F. Anderson, and B. W. Matthews.** 1982. Structural similarity in the DNA-binding domains of catabolite gene activator and *cro* repressor proteins. Proc. Natl. Acad. Sci. U.S.A. **79**:3097–3100.
35. **van den Elzen, P. J. M., J. Matt, H. H. B. Walters, E. Veltkamp, and H. J. J. Nijkamp.** 1982. The nucleotide sequence of the bacteriocin promoters of plasmids CloDF13 and ColE1: role of LexA repressor and cAMP in the regulation of promoter activity. Nucleic Acids Res. **10**:1913–1928.
36. **Walker, G. C.** 1984. Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*. Microbiol. Rev. **48**:60–93.
37. **Wertman, K. F., J. W. Little, and D. W. Mount.** 1984. Rapid mutational analysis of regulatory loci in *Escherichia coli* K-12 using bacteriophage M13. Proc. Natl. Acad. Sci. U.S.A. **81**: 3801–3805.
38. **Zoller, M. J., and M. Smith.** 1982. Oligonucleotide directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutants in any fragment of DNA. Nucleic Acids Res. **10**:6487–6500.