DNA Sequence Determinants of λ Repressor Binding *in Vivo*

Nicholas Benson, Paul Sugiono and Philip Youderian

Department of Biological Sciences, University of Southern California, Los Angeles, California 90089-1481

Manuscript received June 10, 1987
Revised copy accepted September 28, 1987

ABSTRACT

The critical operator determinants for λ repressor recognition have been defined by analyzing the binding of wild-type repressor to a set of mutant operators *in vivo*. Base pair substitutions at six positions within the λ operator half-site impair binding severely, and define these base pairs as critical for operator function. One mutant operator binds repressor better than the consensus operator, and is a superoperator. The model proposed by M. Lewis in 1983 for the binding of λ repressor to its operator accurately predicts the observed operator requirements for binding *in vivo*, with several minor exceptions. The order of affinities of the six natural λ operators has also been determined.

TEMPERATE coliphage λ encodes two repressor proteins, the products of the *cI* and *cro* genes, that regulate the decision between lytic and lysogenic development [see PTASHNE *et al.* (1980) and GUSSIN *et al.* (1983) for reviews]. The λ repressor and Cro proteins act to inhibit the transcription of two promoter complexes by binding specifically to three 17-base pair (bp) sites within each complex. Although the six bound sites within the two λ promoter/operator regions share extensive sequence homology, they all possess unique sequence features (Figure 1).

The amino-terminal half of λ repressor has been crystallized and is composed of a disordered amino-terminal arm attached to five α-helices joined by short surface turns (PABO and LEWIS 1982). Two of these α-helices (2 and 3) comprise the conserved "helix-turn-helix" secondary substructure characteristic of at least three other repressors, λ Cro protein (OHLENDORF *et al.* 1982, 1983), *Escherichia coli* Trp repressor (SCHEVITZ *et al.* 1985), and coliphage 434 repressor (ANDERSON, PTASHNE and HARRISON 1987), and are thought to be involved directly in sequence-specific recognition of the λ operator. The majority of mutations in the *cI* gene that specifically affect the binding of repressor change amino acid residues within these two α-helices (HECHT, NELSON and SAUER 1983; NELSON, HECHT and SAUER 1983).

On the basis of the crystal structure of the DNA-binding domain of λ repressor (PABO and LEWIS 1982), LEWIS *et al.* (1983) have proposed a detailed model for how the helix-turn-helix portion of a repressor monomer might form specific, weak chemical bonds with an operator half-site (Figure 2). Specific interactions are predicted to involve the formation of hydrogen bonds and van der Waals contacts between functional group donors and acceptors on amino acid side chains and corresponding acceptors and donors on the operator DNA [see PABO and SAUER (1984) for review]. The proposed network of interactions suggests that, at particular positions within the operator, some base pair substitutions should preserve the ability of repressor to bind. These are substitutions that replace wild-type base pairs with base pairs that display similar functional groups in the major groove involved in specific interactions. Other operator substitutions are predicted to interfere with binding, since they should present functional groups to the binding surface of repressor predicted to disrupt the network of specific weak bonds.

To determine what sequence features of the operator are critical for repressor binding, we have made all possible symmetric variants of the consensus λ operator. These mutant operators differ by pairs of single base pair substitutions from consensus. In this paper, we examine the interaction of wild-type repressor with the mutant operators to define the operator sequence requirements for repressor binding *in vivo*, and show that these requirements are nearly perfectly consistent with the model proposed for repressor binding (LEWIS *et al.* 1983).

MATERIALS AND METHODS

Bacteria and phage: Bacterial strains are derivatives of *Salmonella typhimurium* LT2 or *E. coli* K12. Salmonella strain MS1363 (*leuA-am*414 *sup E*) (SUSSKIND 1980) was used for the permissive growth of challenge phages. MS1868 (*leuA-am*414 *sup*^{o})/F′*lacI*^{Q1} was used as a host for challenge phage infections (BENSON *et al.* 1986). Recombinant challenge phages were constructed as described (BENSON *et al.* 1986). To confirm the operator genotypes of challenge phages, challenge phage DNAs were sequenced by the method of SANGER, NICKLEN and COULSON (1977) as modified by ZAGURSKY *et al.* (1985), using a synthetic oligonucleotide complementary to the 5′ end of the *arc* gene (5′ CGGCATTTTGCTCATTCC 3′) as primer. Plasmids were constructed and used to transform *E. coli* strains MM294 (MESELSON and YUAN 1968) or DH1 (HANAHAN 1983).

Media, enzymes, and chemicals: Media and general

phage techniques have been described (BENSON *et al.* 1986). Antibiotics were from Sigma Chemical Company. Restriction endonucleases, *E. coli* DNA polymerase I large fragment, T4 polynucleotide kinase, and T4 DNA ligase came from New England Biolabs. Deoxyribonucleoside triphosphates and ATP were purchased from PL Biochemicals. α-[^{32}P]dATP (700 Ci/mmol), used in DNA sequence analyses, was from ICN. Synthetic binding sites were synthesized on a model 381A Applied Biosystems automated DNA synthesizer, usually as mixed pairs of 17-mers with both G and C at the 9th (central) position. Oligonucleotides representing the natural λ operators were the kind gifts of BILL METZGER and PONZY LU.

Plasmids: The plasmid source of λ repressor, pLRB, carries an operon fusion of the *lac*UV5 promoter to both the *cI* gene and the pBR322 *tetA* gene (NELSON and SAUER 1985). Operator sequences were cloned into the *Xho*I site of plasmid pPY140 (BENSON *et al.* 1986). T4 DNA ligase (600 units) was added to a mixture of 6–60 μg of unphosphorylated oligonucleotide and 0.1–1.0 μg of *Xho*I-digested pPY140 DNA in 20 μl of 2 mM ATP, 20 mM dithiothreitol, 66 mM Tris (pH 7.5), 10 mM $MgCl_2$, and 200 μg/ml bovine serum albumin, and incubated at 25° for more than 6 hr. DNAs were precipitated by the addition of 5 μl 3 M sodium acetate (pH 4.5) and 80 μl of ethanol, pelleted by microcentrifugation for 10 min at 25°, and vacuum-dried. DNAs were resuspended in 20 μl 10 mM Tris (pH 7.5) and incubated at 65° for 5 min. An aliquot of 66 μl of 10 mM Tris (pH 7.5), 100 mM NaCl, 5 mM sodium ethylenediaminetetraacetate was added, mixtures were incubated at 65° for 10 min, placed at room temperature for 10 min, and used to transform *E. coli* MM294 or DH1. More than 70% of the transformants were found to carry a single insert of the oligonucleotide.

Construction and use of challenge phages: Challenge phages were constructed from clones of synthetic operators in plasmid pPY140 and analyzed as described by BENSON *et al.* (1986). To determine the efficiency of lysogeny of MS1868/F'*lacI*^{Q1} (pLRB) upon infection with a challenge phage, cells were grown in LB medium at 37° to a density of 1×10^8/ml, β-isopropylthiogalactoside (IPTG) was added to a specified concentration, and cells were grown out to a density of 2×10^8/ml. Challenge phage were added to give an input multiplicity of infection (MOI) of 20 phage/cell and allowed to adsorb for 20 min at 25°. Infected cells were spread on green indicator plates with ampicillin (50 μg/ml), kanamycin sulfate (20 μg/ml), and IPTG at the specified concentration. Plates were scored after incubation for 20 hr at 37°.

RESULTS

Selection for λ repressor binding: The assay we have used to detect specific DNA binding involves the repression of the temperate phage P22 *ant* (antirepressor) operon, which controls the decision between lytic and lysogenic development of P22 upon infection of a sensitive Salmonella host. If antirepressor is expressed at high levels during infection with P22, the infecting phage genome develops lytically and kills its host. If the *ant* gene is repressed during infection, P22 may opt for lysogenic development, and become established as an integrated provirus in surviving host cells [see SUSSKIND and YOUDERIAN (1983) for review].

A challenge phage is a derivative of P22 *arc-am-*

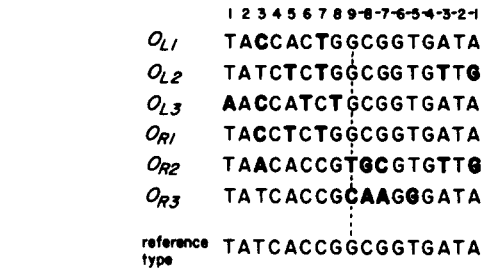

FIGURE 1.—Sequences of the natural λ operators. The six natural operator sequences are oriented in the same direction as the P_L and P_R promoters they regulate. Sequence differences from the reference-type (consensus) site are indicated in bold face for each operator.

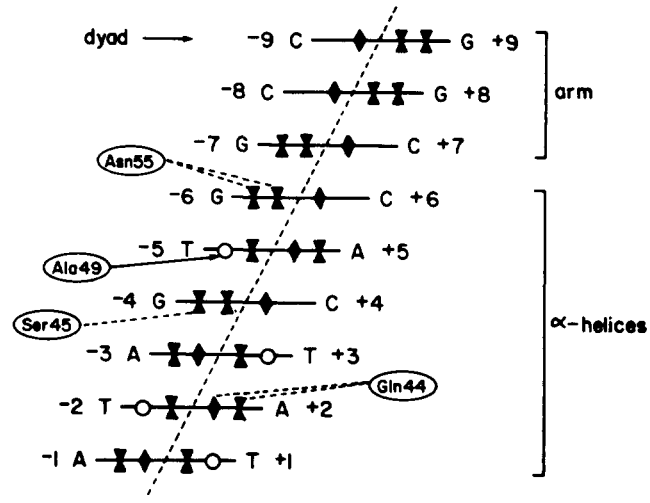

FIGURE 2.—Model for λ repressor binding. The functional groups of the base pairs in the major groove of the reference-type operator are shown in two dimensions, following the example of WOODBURY, HAGENBUCKLE and VON HIPPEL (1980). The symbols indicate functional groups as hydrogen-bond donors (♦), hydrogen-bond acceptors (⧗), and thymine methyl groups (O). Presumed hydrogen-bonds between repressor side chains and the bases that contribute to the specificity of binding are indicated as dotted lines; van der Waals contacts are indicated as arrows.

H1605 phage that carries a substitution of a synthetic DNA-binding site for the natural (*mnt*) operator that controls *ant* expression, and places antirepressor synthesis under the control of a specific DNA-binding protein that can occupy the substituted site. Upon infection by a challenge phage, a Salmonella host will survive infection only if it produces a protein that can bind the substituted site. The frequency of lysogeny of a challenge phage is a measure of how well the substituted DNA binding site is bound by a specific binding protein produced by the Salmonella host (BENSON *et al.* 1986).

Previously, we showed that challenge phages carrying the λ O_{L1} operator efficiently lysogenize cells that produce λ repressor, and efficiently kill cells that do not produce λ repressor (BENSON *et al.* 1986). A challenge phage with a symmetric, consensus (reference-type) λ operator (Figure 1) also allows us to select for λ repressor activity. As illustrated in Figure 3 and

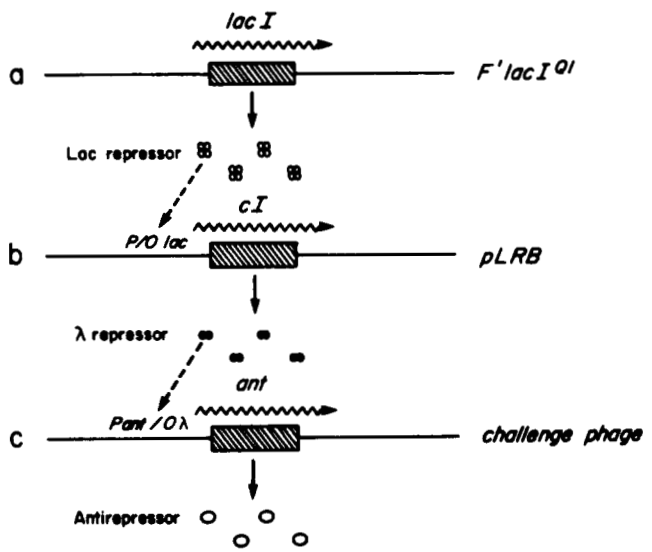

FIGURE 3.—The challenge phage selection. The figure shows the heirarchy of regulation that controls the lysogeny/lysis decision of a P22 challenge phage with a λ operator. (a) An F′ episome carrying the *lacI* gene under the control of the strong *Q1* promoter produces high, constitutive levels of Lac repressor. (b) In the absence of the inducer IPTG, Lac repressor turns down expression of the *cI* gene from plasmid pLRB; in the presence of IPTG, Lac represser cannot bind its operator on pLRB, and the plasmid produces λ repressor. (c) If λ repressor can bind a challenge-phage borne (λ) operator, it will prevent RNA polymerase from binding to the *ant* promoter, inhibit antirepressor production, permit the lysogenic development of the phage, and allow for the survival of the host and the formation of an antibiotic-resistant lysogen. Production of antirepressor leads to the lytic development of the challenge phage and the death of the host cell.

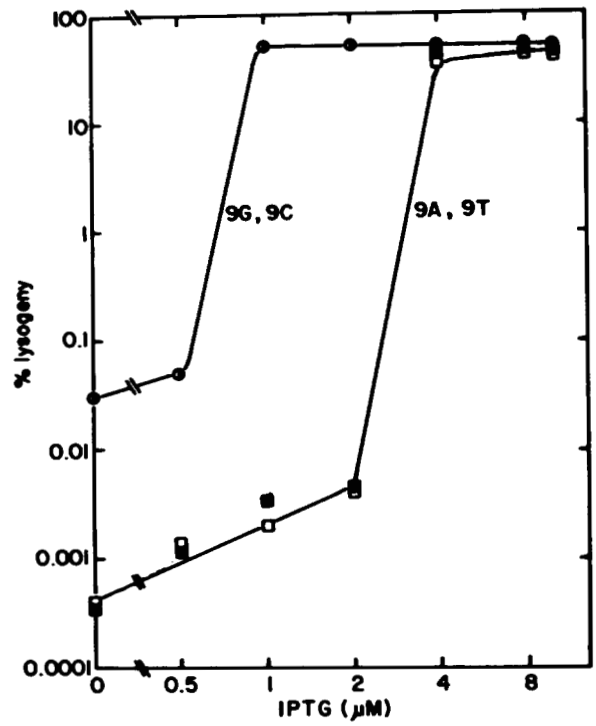

FIGURE 4.—The central, asymmetric base pair of the λ operator is important for operator function. The efficiency of lysogeny as a function of IPTG concentration is shown for repressor-producing cells (MS1868/F′*lacI*^{Q1} (pLRB)) infected with challenge phages carrying operators with changes in the central base pair (9). Challenge phage infections were performed as described in MATERIALS AND METHODS. In *E. coli*, pLRB, when fully induced, makes about 100-fold more repressor than a single copy λ lysogen (NELSON and SAUER 1985). A lower concentration of repressor is required to saturate operators with central G:C or C:G base pairs than operators with central T:A or A:T base pairs.

shown in Figure 4, a host cell carrying a plasmid-borne operon fusion of the *lacUV5* promoter to the λ repressor gene controlled by a single-copy *lacI*^{Q1} repressor gene fails to lysogenize a challenge phage with the reference-type operator efficiently in the absence of IPTG, an inducer of Lac repressor. This result shows that the ability of the challenge phage to lysogenize this host depends on the host's ability to produce active λ repressor. The reference-type operator is also bound by Cro protein, when Cro is produced from a similar operon fusion (N. BENSON, unpublished results).

The natural λ operators are bound in a pairwise cooperative manner by repressor dimers, and overlap many of the critical determinants of λ P_{RM} and P_R promoter function. We examined the binding of repressor dimers to a single copy of the operator on a challenge phage, so that our analysis would not be complicated by the cooperative interactions between repressor dimers bound at multiple, adjacent sites. We also placed the λ operator at the startpoint of transcription of the *ant* operon, a position that is not critical for *ant* promoter function, to avoid possible pleiotropic effects of operator mutations on promoter strength. In this way, we have designed our assay for repressor binding *in vivo* to reflect the interaction of repressor dimers with simple, single mutant operators.

HOCHSCHILD and PTASHNE (1986) and HOCHSCHILD, DOUHAN and PTASHNE (1986) have examined the effects of asymmetric, single base pair changes in the λ operator on the binding of λ repressor. In all cases, their results are complicated by the presence of more than one operator sequence on the templates they used. The differences between our results and theirs (see below) may be due to the cooperative interaction of binding sites in their experiments, a contingency we have explicitly avoided by using single operator sites on our templates.

Sequences flanking the operator are important for repressor binding: Several recent results have indicated that changes in neighboring base pairs that are not directly involved in specific protein/DNA interactions may alter the local structure of operator DNA, thereby affecting the formation of weak bonds at neighboring positions in an operator and influencing binding. This is the case for base pairs in the center of the coliphage 434 operator (KOUDELKA, HARRISON and PTASHNE 1987) and the *E. coli* trp operator (BASS *et al.* 1987). On the basis of these results, one might predict that the context (immediately flanking se-

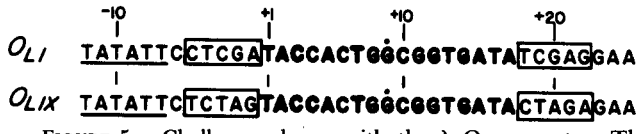

FIGURE 5.—Challenge phages with the λ O_{L1} operator. The sequences of the two contexts of cloned O_{L1} operator DNA are shown. The O_{L1} operator is indicated in bold face, the −10 region of the *ant* promoter is underlined, and the five base pair inverted repeats surrounding O_{L1} that differ in the two phages are in boxes. The operator called "O_{L1}" is stronger than "O_{L1x}" (see Figure 6).

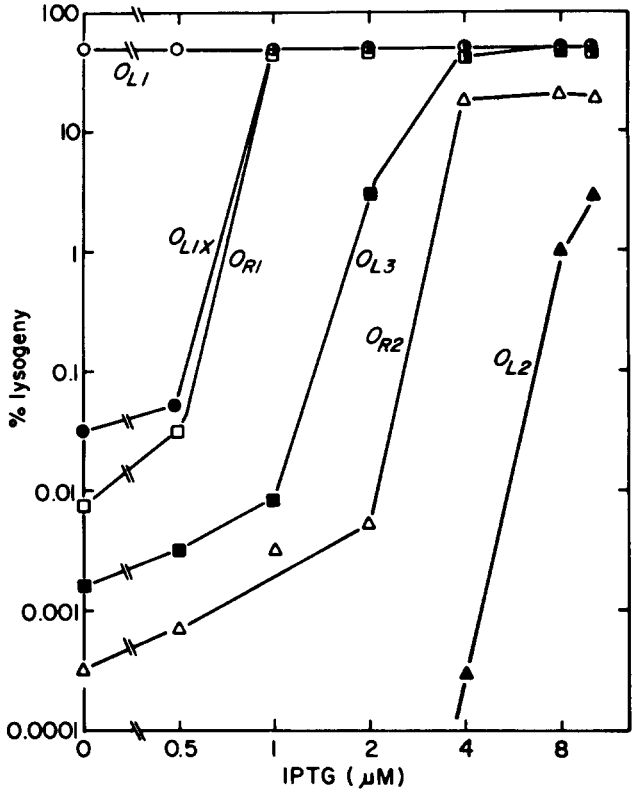

FIGURE 6.—Order of affinities of the natural λ operators. Legend is as in Figure 3.

quences) around an operator also might influence its affinity.

To examine the possible effects operator context might have on the binding of λ repressor, we compared the binding of wild-type repressor to the λ O_{L1} operator in two different contexts, as diagrammed in Figure 5. Figure 6 shows that when the O_{L1} operator is surrounded by inverted repeats of the sequence, 5′ TCTAG 3′ (O_{L1x}), it is notably weaker than when it is surrounded by inverted repeats of the sequence, 5′ CTCGA 3′ (O_{L1}). These changes affect base pairs −5 to −1 and +18 to +22 (with respect to the startpoint of transcription) of the regulated *ant* promoter. Neither of these substitutions reduces *ant* promoter strength significantly; challenge phages with each substitution retain the Arc$^-$ lethal phenotype (SUSSKIND 1980), indicating that they direct the synthesis of high levels of antirepressor. This is the case both in the absence of λ repressor, and in the presence of a mutant λ repressor defective in binding (Asn55→Lys) (HECHT, NELSON and SAUER 1983) made from a plasmid otherwise isogenic with pLRB (data not shown). In fact, both of these phages have plaque morphologies under nonpermissive conditions (on a *sup*° host at 37°) similar to that of their parent, P22 *Kn9 arc-amH1605* phage, indicating that their relative promoter strengths are comparable (YOUDERIAN, BOUVIER and SUSSKIND 1982; GRANA, YOUDERIAN and SUSSKIND 1985).

We conclude that changes in the sequences flanking an operator may result in significant changes in operator strength *in vivo*. PRENTKI, CHANDLER and GALAS (1987) have shown that this is also the case for *E. coli* IHF (integrative host factor) binding sites. This result emphasizes the *critical* importance of using otherwise isogenic sets of operators in making binding comparisons. In the experiments described below, all of the operators are in the context that results in stronger binding to O_{L1} (derived from synthetic operator sequences cloned in plasmid pPY140; see MATERIALS AND METHODS).

Symmetry requirements for λ repressor binding: The 17-bp reference-type operator bound by repressor is nearly symmetric (Figure 1). An axis of dyad symmetry passes through the central G:C base pair and relates the remaining 16 bp perfectly. Presumably, each operator half-site is recognized by a symmetrically related subunit of a repressor dimer in an identical way. Chemical and enzymatic analyses of the bound repressor/operator complex reveals that many of the symmetrically positioned atoms in the operator behave in similar ways upon the binding of repressor (JOHNSON, MEYER and PTASHNE 1979; JOHNSON 1980; JOHNSON, PABO and SAUER 1980), consistent with this idea.

E. coli Lac repressor also binds a nearly symmetric operator with a central, asymmetric base pair. SADLER, SASMOR and BETZ (1983) have shown that a synthetic, symmetric version of the *lac* operator missing the central base pair binds Lac repressor more tightly, suggesting that Lac repressor prefers an operator with a center of perfect symmetry. Consistent with this idea, H. SASMOR and J. L. BETZ (unpublished results) have found that upon the binding of repressor, symmetric functional groups in this symmetric site are protected from chemical or enzymatic attack to similar extents. Moreover, sites cloned from other organisms that function as Lac operators have half-sites with a symmetric disposition (SIMONS *et al.* 1984).

To test whether λ repressor prefers to bind a perfectly symmetric operator, we synthesized two variants of the reference-type operator, one 16 bp long (with a deletion of the central, asymmetric G:C base pair) and one 18 bp long (with an insertion of a C:G base pair). As shown in Table 1, neither the shorter nor

TABLE 1

Frequencies of lysogeny of λ repressor-producing cells infected with challenge phages carrying symmetric variants of the consensus λ operator

Operator	Sequence	IPTG 0 μM	IPTG 2 μM	IPTG 10 μM
O_{L1}	TACCACTGGCGGTGATA	0.3	0.3	0.4
18 bp	TATCACCGGCCGGTGATA	$<10^{-7}$	$<10^{-7}$	$<10^{-7}$
16 bp	TATCACCG..CGGTGATA	$<10^{-7}$	$<10^{-7}$	$<10^{-7}$
ref (9G)	TATCACCGGCGGTGATA	2×10^{-4}	0.2	0.2
9C	TATCACCGCCGGTGATA	2×10^{-4}	0.2	0.2
9A	TATCACCGACGGTGATA	3×10^{-6}	4×10^{-5}	0.2
9T	TATCACCGTCGGTGATA	2×10^{-6}	4×10^{-5}	0.2

Cells producing different, increasing amounts of λ repressor, due to growth in the presence of increasing amounts of IPTG, were infected with challenge phages carrying the listed operator sequences. Numbers are the frequencies of lysogeny (the number of surviving cells divided by the number of infected cells) from challenge phage infections performed as described in MATERIALS AND METHODS.

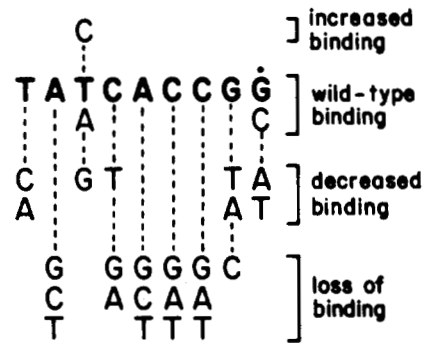

FIGURE 7.—Symmetric base pair substitutions at every position within the λ operator impair repressor binding. Cells producing high levels of repressor (in the presence of 4×10^{-6} M IPTG) were infected with challenge phages differing from the reference-type operator by symmetric pairs of base pairs. The results of these infections are represented in two dimensions. The sequence of the top strand of the left, reference-type half-site is represented by bold letters; the dyad axis passes through position 9 of the operator, indicated by the dot. Immediately below the reference sequence are listed the two changes that do not change operator strength (3A and 9C). Below the reference sequence are listed the changes that reduce the ability of repressor to bind. These mutations are grouped into two categories, changes that reduce survival less than 10^4-fold (decreased binding) and more than 10^4-fold (loss of binding). The 3C change is shown above the consensus half-site, since it results in an operator with higher affinity for repressor. This analysis was complicated by the fact that one of the substitutions (1, T:A→G:C; operator 1G, not shown) results in the modification of three adjacent base pairs, since it creates a site recognized by Salmonella Dam methylase (5′ GATC 3′), in which adenine C6-amino groups are methylated (GOMEZ-EICHELMANN 1979). In a *dam*$^+$ host, this change not only replaces the T:A base pair at position 1 with G:C, but also results in the methylation of the C6 amino groups of adenine in base pairs 2 and 3. When methylated, this operator sequence does not bind λ repressor.

longer operators bind λ repressor. These mutations change both the distance between operator half-sites and the rotational disposition of the half-sites on the surface of the B-form helix. Presumably, changes in the relative positions of the two operator half-sites in each of these variant operators prevents λ repressor from binding.

How critical is the central, asymmetric base pair for λ repressor binding? Figure 4 shows that λ repressor binds operators with central G:C and C:G base pairs equally well, but higher concentrations of repressor are required to saturate operators with central A:T and T:A base pairs.

Operator determinants of λ repressor binding: If each half-site of the λ operator is recognized in a similar way by a symmetrically related monomer of a repressor dimer, then operators derived from the reference type with a symmetric pair of base pair changes should present the same complementary DNA surface to each repressor monomer in a dimer. Therefore, we synthesized a complete set of variant λ operators differing from the reference type by pairs of symmetric single base pair substitutions, and constructed challenge phages with each operator. Each of the challenge phages retains the Arc$^-$ lethal phenotype (SUSSKIND 1980) and has an *ant* promoter comparable in strength to its *Kn9 arc-amH1605* parent *in vivo*, as evidenced by plaque morphology under a variety of conditions (data not shown) (YOUDERIAN, BOUVIER and SUSSKIND 1982; GRANA, YOUDERIAN and SUSSKIND 1985). To determine the ability of wild-type repressor to bind each of these sequences *in vivo*, the efficiency of survival of our host that produces λ repressor, MS1868/F′*lacI*^{Q1} (pLRB), was measured after infection with challenge phages carrying each of the mutant operators. We varied the level of λ repressor produced by the infected host, by varying the level of induction of the hybrid operon on plasmid pLRB that produces λ repressor with IPTG. At higher concentrations of IPTG, this host produces higher levels of λ repressor.

As summarized in Figure 7, 6 of 8 bp (2, 4, 5, 6, 7 and 8) in the operator half-site are critical for repressor binding, since at least one of the three possible substitutions at each of these positions abolishes binding. Eight mutant operators with changes at positions 1 through 9 retain at least some ability to bind repressor. The symmetric variant operators bound by repressor may be subdivided into six classes, depending on how much repressor is required to saturate each operator (Figure 8). Operator 3C carries symmetric transitions at position 3, and binds repressor as well as O_{L1} and better than the reference type operator. One altered operator, 3A, binds repressor as well as the reference type (as does the 9C operator; see Figure 4, above). The remaining seven altered operators bound by repressor are bound less well than the reference type. These operators may be ranked in the order of affinity: 3C = O_{L1} > reference type = 3A > 3G = 4T > 1A = 1C = 8T > 8A > 8C.

The most striking result of this comparison is that

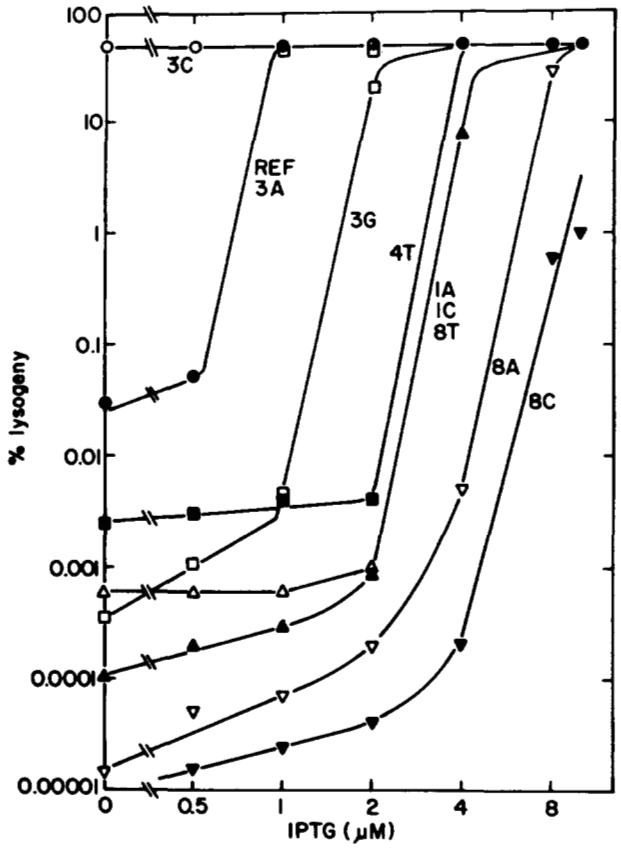

FIGURE 8.—Symmetric mutant operators that bind repressor differ in relative strength. Legend is as in Figure 4. Challenge phages carry the indicated symmetric base pair changes in the reference-type (ref) sequence.

operator 3C binds repressor better than the reference-type operator; the 3C change is an example of an 0°, or "superoperator," mutation (JACOB *et al.* 1960). Not only is less repressor required to saturate operator 3C, but challenge phages with the 3C operator cannot grow lytically in the absence of IPTG upon infection of a host that produces P22 c2 repressor and carries the plasmid producer of λ repressor, whereas challenge phages with the reference-type operator can (data not shown).

To obtain independent genetic evidence that the superoperator change, 3C, increases the affinity of the operator for wild-type repressor, we asked whether the superoperator change could suppress other changes in the operator that impair binding. We synthesized 12 variants of the reference-type operator, each of which carries two pairs of mutations, the 3C pair of changes and an additional pair of changes at position 2, 4, 6, or 7. Table 2 shows that 4 of these 12 changes that are detrimental to binding, 4G, 4A, 7G and 7T, are (at least partially) suppressed when paired with the symmetric 3C change. We conclude that the 3C change is somewhat allele-specific as a suppressor, since it suppresses only a narrow spectrum of operator mutations that decrease binding. It is not position-specific, however, since substitutions at both positions 4 and 7 are suppressed.

TABLE 2

Frequencies of lysogeny of λ repressor-producing cells infected with challenge phages carrying λ operators with the superoperator (3C) change

	Alone (−3C) with IPTG at:		With the superoperator change (+3C) with IPTG at:	
Operator	0 μM	10 μM	0 μM	10 μM
2G	$<10^{-7}$	$<10^{-7}$	$<10^{-7}$	$<10^{-7}$
2C	$<10^{-7}$	$<10^{-7}$	$<10^{-7}$	$<10^{-7}$
2T	$<10^{-7}$	$<10^{-7}$	$<10^{-7}$	$<10^{-7}$
4G	$<10^{-7}$	2×10^{-7}	$<10^{-7}$	4×10^{-4}*
4A	$<10^{-7}$	$<10^{-7}$	1×10^{-5}	0.2
4T	1×10^{-5}	0.3	2×10^{-6}	0.1
6G	$<10^{-7}$	$<10^{-7}$	$<10^{-7}$	$<10^{-7}$
6A	$<10^{-7}$	$<10^{-7}$	$<10^{-7}$	$<10^{-7}$
6T	$<10^{-7}$	$<10^{-7}$	$<10^{-7}$	$<10^{-7}$
7G	$<10^{-7}$	$<10^{-7}$	$<10^{-7}$	4×10^{-3}*
7A	$<10^{-7}$	$<10^{-7}$	$<10^{-7}$	$<10^{-7}$
7T	$<10^{-7}$	$<10^{-7}$	$<10^{-7}$	0.2*

Challenge phages carry operators with the indicated change detrimental to repressor binding. This change is either unpaired (−3C) or paired (+3C) with the 3C superoperator change; infections were performed at the two indicated IPTG concentrations. An asterisk indicates that colonies appear after 40 hr of incubation.

TABLE 3

Frequencies of lysogeny of λ repressor-producing cells infected with challenge phages carrying the natural λ operators

		IPTG		
Operator	Sequence	0 μM	2 μM	10 μM
ref	TATCACCGGCGGTGATA	2×10^{-4}	0.2	0.4
O_{L1}	TACCACTGGCGGTGATA	0.3	0.3	0.4
O_{L2}	TATCTCTGGCGGTGTTG	$<10^{-7}$	2×10^{-7}	0.03*
O_{L3}	AACCATCTGCGGTGATA	2×10^{-5}	0.02*	0.4
O_{R1}	TACCTCTGGCGGTGATA	8×10^{-5}	0.2	0.3
O_{R2}	TAACACCGTGCGTGTTG	2×10^{-6}	4×10^{-5}	0.2
O_{R3}	TATCACCGCAAGGGATA	$<10^{-7}$	$<10^{-7}$	$<10^{-7}$

Repressor-producing cells were infected with challenge phages carrying one of the natural λ operators. Numbers are frequencies of lysogeny as in Table 1. An asterisk indicates that colonies appear after 40 hr of incubation. Since these survivors are unstable lysogens carrying the challenge phage as prophage (and form dark green colonies on green indicator plates), we interpret their slow growth to mean that the amount of repressor produced under these conditions is near the threshold required to saturate a particular operator.

Order of affinities of the natural operators: As shown in Table 3 and Figure 6, we have determined the order of affinities of the natural operators by our assay. The λ repressor binds O_{L1} and O_{R1} even at low concentrations, O_{L3} at moderate concentrations, and O_{R2} and O_{L2} at high concentrations. Operator O_{R3} is not bound even with high concentrations of repressor present in the host cell.

DISCUSSION

Sequence determinants of λ repressor binding: We have developed an assay for λ repressor binding *in*

vivo to a single, symmetric binding site positioned at the startpoint of transcription of the *ant* operon of a P22 challenge phage. In this assay, the ability of λ repressor to bind its operator is measured as the frequency of lysogeny of an operator-bearing challenge phage upon infection of a host that produces λ repressor. To define the operator determinants of binding *in vivo*, we have made all possible symmetric variants of this operator and tested the ability of wild-type repressor to bind each of the mutant operators. We find that at least one mutation at every operator position impairs binding, indicating that every operator position contributes to the formation or stability of the repressor/operator complex. Mutations in 6 bp (2, 4, 5, 6, 7 and 8) of the operator half-site eliminate binding, and show that these base pairs are critical for binding (Figure 7).

One variant operator (3C) binds repressor better than the reference-type and is a superoperator; an operator with a single, asymmetric 3C change also has this phenotype (data not shown). In addition, the 3C change suppresses a limited spectrum of symmetric pairs of mutations in the operator that impair binding (Table 2). HOCHSCHILD and PTASHNE (1986) have claimed that this change has no effect on repressor binding; we find that this is not the case. This discrepancy may be due to the fact that we have examined the binding of repressor to a single operator site in isolation, whereas HOCHSCHILD and PTASHNE (1986) studied binding to sites on templates carrying more than one operator, which may be complicated by pairwise cooperative interactions.

Role of the helix-turn-helix secondary substructure in repressor binding: The model proposed by LEWIS *et al.* (1983) for λ repressor binding to the reference-type sequence is summarized in Figure 2. This model predicts that four amino acids in the conserved helix-turn-helix secondary substructure of repressor should be involved in specific recognition, Gln44, Ala49, Ser45, and Asn55. For each position in the operator we presume is contacted by protein side chains, we can predict what base pair substitutions should have slight or negligible effects on binding, and what substitutions should have more severe effects on binding.

Gln44 should both donate a proton to the N7 group and accept a proton from the N6 group of adenine +2, to make two hydrogen bonds with the operator. If we assume that the loss of either or both of these bonds is sufficient to disrupt binding, then one would predict that all three single bp changes from consensus at position 2 would impair binding, the result we observe.

The hydroxyl group of Ser45 is predicted to donate a proton to the N7 group of guanine −4 to form a hydrogen bond. We find that, consistent with this hypothesis, adenine, but not thymine or cytosine, can substitute for guanine at this position. The change 4T, which places adenine at −4, results in an operator with slightly reduced affinity for repressor (Figure 7). This might be attributed to the lower intrinsic potential of the adenine N7 group to accept a proton, due to the slight difference in electronegativities of the 6-oxo group of guanine and 6-amino group of adenine; alternatively, this may be due to the slightly different positions of the purine N7 groups in guanine and adenine. Neither pyrimidine has a hydrogen bond acceptor in a position corresponding to the purine N7 group.

The methyl group of Ala49 is proposed to make van der Waals contacts with thymine methyl groups at +3 and −5. Our evidence indicates that of thymines at +3 and −5, only thymine −5 is critical for binding. At +3, adenine can substitute for thymine without loss of binding, and the substitution of cytosine for thymine increases operator strength. If we assume that the loss of a van der Waals contact can disrupt binding, the proposed interaction of the Ala49 methyl group with the thymine −5 methyl group is sufficient to account for the recognition specificity observed at position 5 in the operator.

Finally, the model predicts that the side chain of Asn55 denotes two protons to the O6 and N7 groups of guanine −6. Consistent with this, only guanine at this position leads to efficient binding. All other base pair substitutions are predicted to disrupt one or both of these hydrogen bonds, and are found to result in a loss of binding (Figure 6).

Given the simple assumption that loss of a pair of hydrogen bonds or van der Waals contacts between a repressor dimer and a pair of operator half-sites is sufficient to result in a loss of binding in our assay, the detailed predictions for how repressor might interact with operator derived from model-building (LEWIS *et al.* 1983) coincide with the results of our assay *in vivo*, with several minor exceptions. We suspect that the methyl group of Ala49 does not participate in an energetically favorable van der Waals interaction with the +3 thymine methyl group. This is the case, since the change of a T:A to a C:G base pair at this position results in the loss of the thymine methyl from the pyrimidine ring, yet improves the binding of repressor. In addition, it appears that repressor shows some preference for a T:A base pair at position 1; the reason for this preference is yet unclear.

Role of the amino-terminal arm in repressor binding: The model of specific repressor/operator interactions proposed by LEWIS *et al.* (1983) does not account for the recognition specificity we observe at positions 7, 8, and 9 of the operator. The first eight amino acids of λ repressor are disordered in the crystal structure (PABO and LEWIS 1982) and are thought to

assume an extended configuration, or "arm," in solution that wraps around the center of the operator upon binding. Biochemical analyses of the binding of a fragment of λ repressor missing the first three residues of its amino-terminal arm show that the arm makes contacts with at least the three central base pairs (positions 8 and 9) of the operator (PABO *et al.* 1982). The first five amino acids of the arm (H_2N-ser-thr-lys-lys-lys) are polar, and can make either nonspecific (phosphate backbone) or specific contacts with DNA. It appears that a C:G base pair at position 7 is critical for repressor binding, whereas the base pairs at positions 8 and 9 are not so important. This result suggests that the arm may make a specific contact with guanine −7, and additional contacts with guanines +8 and +9.

Relative affinities of the natural operators: From our analysis of symmetric operators, we can rank changes in the natural operators as being severely detrimental, mildly detrimental, neutral, or beneficial for the binding of repressor (Figure 7). Of the natural sites (Figure 1), the site recognized best by repressor, O_{L1}, differs least from our reference type. It has two differences in its left half-site (Figure 1), one of which increases operator affinity (3C), and one of which decreases affinity (7T).

In the O_{L1} operator, the 3C change is epistatic in its effect to the 7T change (Table 2), since O_{L1} binds wild-type repressor with a higher affinity than consensus, as does an operator with a single, asymmetric 3C change. A symmetric 3C change also partially suppresses a symmetric 7T change (Table 2). If, however, we consider that the λ O_{L1} operator is composed of two half-sites, one variant and one consensus, we find an extraordinary result. Although an operator with *one* variant O_{L1} half-site is *stronger* than consensus, an operator with *two* variant O_{L1} half-sites is *weaker* than consensus (Table 2)! Therefore, in the symmetric operator with a pair of 3C and a pair of 7T changes, the 7T change is epistatic to the 3C change. The same relative order of affinities is found for an otherwise isogenic set of operators with a C:G base pair substituted for the central (+9) G:C base pair of the operator (data not shown). These results show that, although the effects of symmetric changes in the operator on the binding of wild-type repressor are simply interpretable, the effects of asymmetric changes in the operator (*i.e.*, the interactions between half-sites differing by multiple changes) are not simply interpretable. Therefore, these results call into question the assumption of HOCHSCHILD and PTASHNE (1986) and HOCHSCHILD, DOUHAN and PTASHNE (1986) that the effects of asymmetric changes in the operator will be independent of the effects of other sequence changes from consensus in the same operator (what they term immediate context).

Nonetheless, several naive arguments that presume that the effects of asymmetric mutations in the operator will be somewhat additive are consistent with the order of affinities of the natural operators we observe. O_{L1} has the consensus base pair in its left half-site at position 5, whereas O_{R1} has a T:A base pair at this position, which results in a loss of binding when present as a symmetric change from consensus (Figure 4). Therefore, we might expect when this change is added to the O_{L1} operator to create O_{R1}, the resulting operator would be weaker, the result we find. The other natural operators differ from consensus by larger numbers of mutations that, when present as symmetric changes, are detrimental to binding. Thus, we might predict naively that O_{R1} would be stronger than the other four natural operators. The order of affinities we observe, $O_{L1} > O_{R1} > O_{L3} > O_{R2} > O_{L2} > O_{R3}$, is consistent with this.

Genetic analysis of binding specificity *in vivo*: Using the challenge phage selection, we have identified the operator determinants critical for the binding of λ repressor. Our data, taken together with the assumption that the loss of a symmetric pair of specific contacts is sufficient to prevent binding in our assay, support the model proposed from crystallographic data for the specific weak bonds that underlie this interaction (LEWIS *et al.* 1983). In a subsequent paper, we show that neither of the two models proposed for the specific contacts between Cro and the λ operator are supported so well by a parallel analysis of Cro binding (N. BENSON, unpublished results).

We thank BILL METZLER and PONZY LU for the oligonucleotides representing the natural operators, and MIMI SUSSKIND for her helpful discussions. This research was supported by National Institutes of Health grant GM34150 to P.Y. P.S. was the recipient of a 1984 summer fellowship from the Student Research Associates Program of the California Heart Association.

LITERATURE CITED

ANDERSON, J. E., M. PTASHNE and S. C. HARRISON, 1987 Structure of the repressor-operator complex of bacteriophage 434. Nature **326:** 8746–852.

BASS, S., P. SUGIONO, D. N. ARVIDSON, R. P. GUNSALUS and P. YOUDERIAN, 1987 DNA specificity determinants of *Escherichia coli* tryptophan repressor binding. Genes Dev. **1:** 565–572.

BENSON, N., P. SUGIONO, S. BASS, L. V. MENDELMAN and P. YOUDERIAN, 1986 General selection for specific DNA-binding activities. Genetics **114:** 1–14.

GOMEZ-EICHELMANN, M. C., 1979 Deoxyribonucleic acid adenine and cytosine methylation in Salmonella typhimurium and Salmonella typhi. J. Bacteriol. **140:** 574–579.

GRANA, D., P. YOUDERIAN and M. M. SUSSKIND, 1985 Mutations that improve the *ant* promoter of Salmonella phage P22. Genetics **110:** 1–16.

GUSSIN, G. N., A. D. JOHNSON, C. O. PABO and R. T. SAUER, 1983 Repressor and Cro proteins: structure, function, and role in lysogenization. pp. 93–123. In *Lambda II*, Edited by R. W. HENDRIX, J. W. ROBERTS, F. W. STAHL and R. WEISBERG. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

HANAHAN, D., 1983 Studies on transformation of *E. coli* with plasmids. J. Mol. Biol. **166:** 557–580.

HECHT, M. H., H. C. M. NELSON and R. T. SAUER, 1983 Mutations in λ repressor's amino-terminal domain: implications for protein stability and DNA binding. Proc. Natl. Acad. Sci. USA **80:** 2676–2680.

HOCHSCHILD, A., and M. PTASHNE, 1986 Homologous interactions of λ repressor and λ Cro with the λ operator. Cell **44:** 925–933.

HOCHSCHILD, A., J. DOUHAN III and M. PTASHNE, 1986 How λ repressor and λ Cro distinguish between O_{R1} and O_{R3}. Cell **47:** 807–816.

JACOB, F., D. PERRIN, C. SANCHEZ and J. MONOD, 1960 L'operon: groupe de genes a expression coordinee par un operateur. C. R. Acad. Sci. **250:** 1727–1729.

JOHNSON, A. D., 1980 Ph.D. thesis, Harvard University, Cambridge, Mass.

JOHNSON, A. D., B. J. MEYER and M. PTASHNE, 1979 Interactions between DNA-bound repressors govern regulation by the λ repressor. Proc. Natl. Acad. Sci. USA **76:** 5061–5065.

JOHNSON, A. D., C. O. PABO and R. T. SAUER, 1980 Bacteriophage λ repressor and Cro proteins: interactions with operator DNA. Methods Enzymol. **65:** 839–856.

KOUDELKA, G. B., S. C. HARRISON and M. PTASHNE, 1987 Effect of non-contacted bases on the affinity of 434 operator for 434 repressor and Cro. Nature **326:** 886–888.

LEWIS, M., A. JEFFREY, J. WANG, R. LADNER, M. PTASHNE and C. O. PABO, 1983 Structure of the operator-binding domain of λ repressor: implications for DNA recognition and gene regulation. Cold Spring Harbor Symp. Quant. Biol. **47:** 435–440.

MESELSON, M., and R. YUAN, 1968 DNA restriction enzyme from *E. coli.* Nature **217:** 1110–1114.

NELSON, H. C. M., and R. T. SAUER, 1985 λ Repressor mutations that increase the affinity and specificity of operator binding. Cell **42:** 549–558.

NELSON, H. C. M., M. H. HECHT and R. T. SAUER, 1983 Mutations defining the operator-binding sites of bacteriophage λ repressor. Cold Spring Harbor Symp. Quant. Biol. **47:** 441–449.

OHLENDORF, D. H., W. F. ANDERSON, R. G. FISHER, Y. TAKEDA and B. W. MATTHEWS, 1982 The molecular basis of DNA-protein recognition inferred from the structure of Cro repressor. Nature **298:** 718–723.

OHLENDORF, D. H., W. F. ANDERSON, M. LEWIS, C. O. PABO and B. W. MATTHEWS, 1983 Comparison of the structure of Cro and repressor proteins from bacteriophage λ. J. Mol. Biol. **169:** 757–769.

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 **298:** 441–443.

PABO, C. O., and M. LEWIS, 1982 The operator-binding domain of λ repressor: structure and DNA recognition. Nature **298:** 443–447.

PABO, C. O., and R. T. SAUER, 1984 Protein-DNA recognition. Annu. Rev. Biochem. **53:** 293–321.

PRENTKI, P., M. CHANDLER and D. J. GALAS, 1987 *E. coli* integration host factor bends the DNA at the ends of IS*1* and in an insertion hotspot with multiple IHF binding sites. EMBO J. **6:** 2479–2487.

PTASHNE, M., A. JEFFREY, A. D. JOHNSON, R. MAURER, B. J. MEYER, C. O. PABO, T. M. ROBERTS and R. T. SAUER, 1980 How the λ repressor and Cro work. Cell **19:** 1–11.

SADLER, J. R., H. SASMOR and J. L. BETZ, 1983 A perfectly symmetric *lac* operator binds the *lac* repressor very tightly. Proc. Natl. Acad. Sci. USA **80:** 6785–6789.

SANGER, F., S. NICKLEN and A. COULSON, 1977 DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:** 5463–5476.

SCHEVITZ, R. W., Z. OTWINOWSKI, A. JOACHIMIAK, C. L. LAWSON and P. B. SIGLER, 1985 The three-dimensional structure of Trp repressor. Nature **317:** 782–786.

SIMONS, A., D. TILS, B. VON WILCKEN-BERGMANN and B. MULLER-HILL, 1984 Possible ideal *lac* operator: *Escherichia coli lac* operator-like sequences from eukaryotic genomes lack the central G:C pair. Proc. Natl. Acad. Sci. USA **81:** 1624–1628.

SUSSKIND, M. M., 1980 A new gene of bacteriophage P22 which regulates synthesis of antirepressor. J. Mol. Biol. **138:** 685–713.

SUSSKIND, M. M., and P. YOUDERIAN, 1983 Bacteriophage P22 antirepressor and its control. pp. 347–363. In: *Lambda II*, Edited by R. W. HENDRIX, J. W. ROBERTS, F. W. STAHL and R. WEISBERG. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

WOODBURY, C. P., JR., O. HAGENBUCKLE and P. H. VON HIPPEL, 1980. DNA site recognition and reduced specificity of the *Eco*RI endonuclease. J. Biol. Chem. **255:** 11534–11546.

YOUDERIAN, P., S. BOUVIER and M. M. SUSSKIND, 1982 Sequence determinants of promoter activity. Cell **30:** 843–853.

ZAGURSKY, J. R., K. BAUMEISTER, N. LOMAX and M. L. BERMAN, 1985 Rapid and easy sequencing of large linear double stranded DNA and supercoiled plasmid DNA. Gene Anal. Techn. **2:** 89–94.

Communicating editor: G. MOSIG