J. Mol. Biol. (1993) 229, 306–318

Lrp, a Global Regulatory Protein of *Escherichia coli*, Binds Co-operatively to Multiple Sites and Activates Transcription of *ilvIH*

Qing Wang[1] and Joseph M. Calvo[2]†

Sections of [1]Genetics and Development and [2]Biochemistry, Molecular and Cell Biology Cornell University, Ithaca, NY 14853, U.S.A.

(Received 17 July 1992; accepted 2 October 1992)

Lrp (Leucine-responsive regulatory protein) has recently been recognized as a major regulatory protein that controls the expression of many operons in *Escherichia coli*. Footprinting and gel retardation experiments with DNA from *ilvIH*, one of the operons controlled positively by Lrp, indicate that Lrp binds to six sites over a 200 base-pair region upstream from the promoter. Binding of Lrp to some of these sites is highly co-operative. We suggest a consensus sequence for Lrp binding based upon a comparison of six binding sites. An analysis of mutants indicates that five out of six binding sites are important for transcription activation and that two or three adjacent Lrp binding sites act synergistically *in vivo*. The observed synergistic effects *in vivo* may result from co-operative binding of Lrp to adjacent sites. We propose a model in which multiple binding sites contribute to the formation of a nucleoprotein complex, but only a particular proximal site positions Lrp properly so that it interacts with RNA polymerase.

Keywords: Lrp; co-operative binding; *ilvIH*; transcription activation; *Escherichia coli*

1. Introduction

Lrp‡ has recently been recognized as a major regulatory protein in *Escherichia coli* that affects the expression of numerous operons, including *ilvIH* (Platko *et al.*, 1990), *serA* (Lin *et al.*, 1990; Rex *et al.*, 1991), *tdh* (Lin *et al.*, 1990; Rex *et al.*, 1991), *sdaA* (Lin *et al.*, 1990), *livJ* (Haney *et al.*, 1992), *livKHMGF* (Haney *et al.*, 1992), *oppABCDF* (Austin *et al.*, 1989), *fanABC* (Braaten *et al.*, 1992), *papBA* (Braaten *et al.*, 1992), *lysU* (Gazeau *et al.*, 1992; Lin *et al.*, 1992*b*), and *lrp* (J. V. Platko & J. M. Calvo, unpublished data; Lin *et al.*, 1992*a*). Other members of the Lrp regulon were identified by Ernsting *et al.* (1992) using two-dimensional gel electrophoresis. They found that the expression of at least 30 proteins, including OmpC, OmpF, glutamine synthetase, the small subunit of glutamate synthase, and W protein, were affected by Lrp. Furthermore, a number of other anonymous operons regulated by Lrp were identified by a *plac*Mu mutational analysis (Lin *et al.*, 1992*a*). For many of the examples cited, leucine also affects operon expression and it has been suggested that Lrp and leucine together constitute a global regulatory system that allows *E. coli* to determine whether it is in an intestinal environment (Ernsting *et al.*, 1992; Newman *et al.*, 1992).

An interesting and unique feature that has emerged from an analysis of all of these systems is the variety of ways that leucine and Lrp interact to regulate gene expression. Lrp acts negatively in some systems and positively in others. For each of these broad categories of action, three subcategories can be distinguished: leucine overcomes the effect of Lrp; leucine is required for the effect; and leucine has no effect. Thus, for example, Lrp activates the *ilvIH* operon and leucine reduces expression of the operon by reducing the extent of activation. Lrp acts negatively on *oppABCDF* expression (Austin *et al.*, 1989), on the other hand, and the inducing effect of leucine can again be explained by assuming that leucine interferes with Lrp action. For the *livJ* and *livKHMGF* operons, Lrp acts negatively, but in these cases, leucine is required for repression (Haney *et al.*, 1992). On the other hand, some anonymous operons identified by a *plac*Mu mutagenesis analysis show the opposite pattern: Lrp

† Author to whom all correspondence should be addressed.

‡ Abbreviations used: Lrp, leucine-responsive regulatory protein; AHAS, acetohydroxy acid synthase; bp, base-pair(s); PCR, polymerase chain reaction.

0022-2836/93/020306-13 $08.00/0

activates expression and that activation requires leucine (Lin *et al.*, 1992*a*). Finally, Lrp acts positively on the *papBA* operon, and negatively on *ompC*, but leucine does not affect the expression of either (Braaten *et al.*, 1992; Ernsting *et al.*, 1992). The molecular basis for these six different patterns of regulation involving Lrp and leucine are not known.

The *lrp* gene was cloned by several groups (Austin *et al.*, 1989; Platko *et al.*, 1990; Lin *et al.*, 1990) and sequenced (Austin *et al.*, 1989; Willins *et al.*, 1991). *lrp* is evolutionarily related to *asnC*, a gene that positively regulates the *asnA* operon (deWind *et al.*, 1985; Kolling & Lother, 1985). Both Lrp and AsnC are predicted to have a helix-turn-helix motif at their amino-terminal end (Dodd & Egan, 1987; Willins *et al.*, 1991). Lrp has been purified to near homogeneity. It exists in solution as a dimer composed of two identical subunits of predicted size 18·8 kDa (Willins *et al.*, 1991).

The work described here relates to *ilvIH*, one of the operons controlled by Lrp. The *ilvIH* operon encodes one of several acetohydroxy acid synthase (AHAS) isoenzymes catalyzing the biosynthesis of branched-chain amino acids (DeFelice *et al.*, 1982). Expression of the operon was repressed five- to ten-fold by addition of leucine to the growth medium (Haughn *et al.*, 1985). The *ilvIH* operon was cloned and sequenced (Haughn *et al.*, 1985; Squires *et al.*, 1981) and the promoter region defined by nuclease mapping and a mutational analysis (Haughn *et al.*, 1985). A 5′-deletion analysis showed that a 331 base pair region upstream from the start of transcription was sufficient for both transcription initiation and leucine-mediated repression of the operon, and that sequences more than 200 bp upstream from the site of transcription initiation were required for optimal expression (Haughn *et al.*, 1985). *In vitro*, Lrp bound upstream from the *ilvIH* promoter and the extent of binding was reduced by leucine, but not by isoleucine or valine (Ricca *et al.*, 1989). In a mutant lacking Lrp because of a Tn*10* insertion in *lrp*, expression from the *ilvIH* promoter was reduced more than 30-fold, and leucine no longer affected the residual promoter activity (Platko *et al.*, 1990). Consideration of these results plus others led to the following simple model: Lrp activates transcription from the *ilvIH* promoter and leucine reduces expression by interfering with the action of Lrp.

A major question concerning the *ilvIH* operon, and indeed all of the other operons regulated by Lrp, is whether Lrp acts directly on transcription initiation, or whether it acts only indirectly by affecting the expression of some other regulatory gene which in turn affects the target operon. The work described here demonstrates that Lrp acts directly to activate transcription of the *ilvIH* operon of *E. coli*. Using high resolution chemical footprinting employing MPE·Fe(II) [methidiumpropyl EDTA-Fe(II)] and purified Lrp, we show that Lrp binds specifically to at least six sites over a 200 bp region upstream from the *ilvIH* promoter. These binding sites are clustered into two regions, an upstream region containing sites 1 and 2 (−255 to −215)) and a downstream region containing sites 3, 4, 5 and 6 (−140 to −50). Mutations introduced separately into five of these binding sites reduced the affinity of these sites for Lrp *in vitro* and decreased the expression of the *ilvIH* operon *in vivo*. Our results support the idea that Lrp binding to multiple sites upstream from the promoter/regulatory region is required for the stimulatory effect that Lrp has on transcription initiated at the *ilvIH* promoter.

2. Materials and Methods

(a) *Bacterial strains*

Strains used are derivatives of *E. coli* K-12. Phage M13 was propagated in strain JM101 [*sup*E *thi* Δ(*lac proAB*)/F′ *traD*36 *proAB lacIqZ* M15]. Strain P90C [*ara*Δ(*lac-pro*) *thi*] was used as host for plasmid pRS415 and its derivatives (Simons & Kleckner, 1987), and for lysogenization by phage lambda and its derivatives.

Cells were grown in L-broth (LB) (Miller, 1972) or minimal medium (M9 minimal salts supplemented with 0·2% (w/v) glucose and micronutrients and 5 μg thiamine/ml (Rosenthal & Calvo, 1987)). Ampicillin (200 μg/ml) was added to media for selection of plasmid-containing strains.

(b) *DNA isolation, labeling and sequencing*

Standard techniques involving nucleic acids, including plasmid and lambda phage DNA isolation and labeling of DNA with ^{32}P by treatment with reverse transcriptase or with polynucleotide kinase, were performed by techniques described by Sambrook *et al.* (1989). DNA sequence analysis was performed by the dideoxy method of Sanger *et al.* (1977) using the Sequenase kit from United States Biochemical, Inc.

(c) *Synthesis of DNA fragments by PCR*

DNA fragments I, II, III and IV used in gel retardation and footprinting studies were made by the polymerase chain reaction (PCR) using as template the replicative form of a mp13 phage (mpQW156; origin described below) containing the promoter and regulatory region of the *ilvIH* operon. The sequences of the oligonucleotide primers (Fig. 1), synthesized by the Cornell University Oligonucleotide Synthesis Facility, were:

P1: 5′-ATTAGTCTAGATTTGCAAACGC-3′
P2: 5′-GTTGTGGGATTCAGCCGAT-3′
P3: 5′-ATCGGCTGAATCCCACAAC-3′
P4: 5′-GCTTAAGGAATTCGCAGCCAGA-3′
P5: 5′-GTGTAAAACGAATTCACCGTCCG-3′.

A 100 μl reaction mixture contained 1 ng of template DNA (mpQW156), 1 μg of each primer, 50 mM-KCl, 10 mM-Tris·HCl (pH 8·8), 2·5 mM-$MgCl_2$, 0·01% (w/v) gelatin, 0·5 mM-each dNTPs and 5 units of *Taq* DNA polymerase. For each cycle, samples were denatured at 94 °C for 2 min, annealed at 37 °C for 2 min, and extended at 72 °C for 2 min. Thirty cycles were performed, with the extension part of the last cycle lengthened to 15 min. Products were purified by electrophoresis through 2% (w/v) low melting agarose. Fragments prepared by PCR were cut with a restriction enzyme and end labeled with reverse transcriptase and [^{32}P]dATP or [^{32}P]dCTP.

(d) *Preparation of synthetic Lrp binding sites*

Lrp binding sites 1, 2, 4 and 5 were synthesized chemically as complementary oligonucleotides. To form duplexes, oligonucleotides in 0·1 M-NaCl, 10 mM-Tris (pH 8), 1 mM-EDTA were heated at 95 °C for 5 min and cooled slowly to room temperature. Products were extracted with phenol/chloroform and precipitated with ethanol. Sites were labeled as follows: 1 and 2 with reverse transcriptase and [^{32}P]dCTP; 4 with polynucleotide kinase and [^{32}P]ATP; 5 with reverse transcriptase and [^{32}P]dTTP.

(e) In vitro *mutagenesis*

Site-directed mutagenesis was carried out following procedures described by Kunkel (1985) and Kunkel *et al.* (1987). Phage mp13GH7 contains region −332 to +29 of the *ilvIH* operon cloned into the *Sma*I site of phage mp8 (Haughn, 1985). The *ilvIH* fragment was excised with *Eco*RI and *Hin*dIII and subcloned into phage mp18, resulting in phage mpQW156. Using single-stranded DNA from phage mpQW156 as template and the indicated oligonucleotide primers, site-directed mutagenesis was used to create the following mutants: mpQW186 (primer 186), mpQW187 (primer 187), mpQW188 (primer 188), mpQW252 (primer 252), mpQW251 (primer 251), mpQW250 (primer 250), mpQW189 (primer 189), mpQW157 (primer 157) and mpQW185 (primer 185). The identity of the mutations (Fig. 1B) was verified by sequencing.

Nucleotide sequences of the primers are as follows (mutations in italics):

Primer

186 5′-CTACATTCAATAAAAAATGCAG*CTCGAG*CC-AGACAT-3′

187 5′-CATTCAGAATAAAATTC*CCATGGTCGACTC-GAG*TGCAGAATAAACCAGAC-3′

188 5′-ACATTCAG*CTCGAG*ATTCTACA-3′

252 5′-AATCCCACA*CTCGAG*AGCAATTTC-3′

251 5′-TAAATTGAT*CTCGAG*TCGGCTGAA-3′

250 5′-TTACAGAGG*CTCGAG*TTGATAATA-3′

189 5′-GATAAAATC*GATATCG*ACAGAGGAT-3′

157 5′-GTAATAAG*CTCGAG*CAGGAGGA-3′

185 5′-AGCCAGAG*CTCGAG*ACACGGGG-3′

Two other mutations were created starting from the mutation in phage mpQW157. DNA from the phage was cut at the unique *Xho*I site that was newly created by site-directed mutagenesis, trimmed with mung bean nuclease, and religated. Two mutants, mpQW158 (deletion from −76 to −69), and mpQW159 (T to C change at −76 and a deletion from −75 to −65) were identified by sequencing.

(f) *Construction of* ilvIH *promoter::*lacZYA *fusions*

*Eco*RI–*Bam*HI fragments from phage mpQW156 and its derivatives mpQW186, mpQW187, mpQW188, mpQW252, mpQW251, mpQW250, mpQW189, mpQW157, mpQW158, mpQW159, and mpQW185 were separately subcloned into plasmid pRS415 (Simons & Kleckner, 1987) cut with *Eco*RI and *Bam*HI, yielding plasmids pCV156, pCV186, pCV187, pCV188, pCV252, pCV251, pCV250, pCV189, pCV157, pCV158, pCV159 and pCV185, respectively. In each of these constructs, the *ilvIH* promoter/regulatory region is fused to a *lacZYA* reporter gene, resulting in a transcriptional fusion. These fusions were transfered to λRS45 by recombination (Simons & Kleckner, 1987), and the resulting lambda derivatives were used to form lysogens in strain P90C. Single copy lysogens were identified by assaying 5 colonies for β-galactosidase activity (most lysogens were single copy). The designation of these lysogens are: CV1042 (λQW156), CV1049 (λQW186), CV1050 (λQW187), CV1051 (λQW188), CV1320 (λQW252), CV1319 (λQW251), CV1318 (λQW250), CV1052 (λQW189), CV1043 (λQW157), CV1044 (λQW158), CV1045 (λQW159), CV1326 (λQW185). Note that the λ derivative carries a designation number that is the same as the plasmid from which it was derived.

Plasmid pCV223 was formed as follows. Plasmid pCV185 was cut with both *Eco*RI and *Xho*I, treated with the Klenow fragment of DNA polymerase I, and religated, resulting in plasmid pCV223. The corresponding λ lysogen is CV1327 (λQW223).

(g) *β-Galactosidase assays*

Stationary phase cultures were diluted 1:100 in M9 minimal medium containing 200 μg of proline, valine and isoleucine/ml, and either lacking or containing 200 μg of leucine/ml and grown at 37 °C. β-Galactosidase assays were performed as described by Platko *et al.* (1990). For each strain analyzed, cultures were prepared from 3 or more single colonies and specific activity was determined for each culture. Activity = units/min where 1 unit = $[A_{420} - (1{\cdot}65 \times A_{550})] \times 780$. Specific activity = units/min × A_{550}. In addition, differential rates of synthesis were measured for each strain using 5 to 7 samples taken from cultures in logarithmic phase having A_{550} from 0·1 to 0·7. Specific activity was calculated from the slope of the line resulting from a plot of units/ml against A_{550}. Values reported here are averages from all of the determinations.

(h) *Gel retardation experiments*

The Lrp used in these experiments, of about 98% purity, was a gift from D. A. Willins (Willins *et al.*, 1991). With the exception noted in the following sentence, reaction mixtures contained the following in a final volume of 20 μl: 1 ng labeled DNA, 20 mM-Tris·hydrochloride (pH 8), 0·4 mM-EDTA, 0·1 mM-dithiothreitol, 50 mM-NaCl, 1 mM-$MgCl_2$, 12·5% glycerol, 100 ng bovine serum albumin/ml, and 1 μg of sonicated calf thymus DNA. For the experiments employing chemically-synthesized binding sites, reaction mixtures were as above except that calf thymus DNA was omitted. Reactions, initiated by the addition of Lrp, were carried out at room temperature for 20 min. Incubation for times longer than 20 min did not alter the proportion of different complexes, indicating that equilibrium was attained (data not shown). Samples were fractionated by electrophoresis through 2% (w/v) agarose or 5% (w/v) polyacrylamide (acrylamide: bisacrylamide, 40 : 1) in 6·7 mM-Tris·hydrochloride (pH 7·6), 3·3 mM-sodium acetate, 1 mM-EDTA. Electrophoresis was carried out at 10 V/cm at room temperature. Gels were transferred to 3MM Whatman paper, dried, and autoradiographed. For quantitation, gels were analyzed using a Betascope Model 603 Blot Analyzer (Betagen Corporation).

(i) *MPE·Fe(II) footprinting analyses*

A 2-stage MPE·Fe(II) footprinting procedure described by Landolfi *et al.* (1989) was used with a few modifications. Reaction mixtures were prepared as described above for gel retardation experiments except that 10 ng of ^{32}P-labeled DNA was used. Reaction mixtures were then

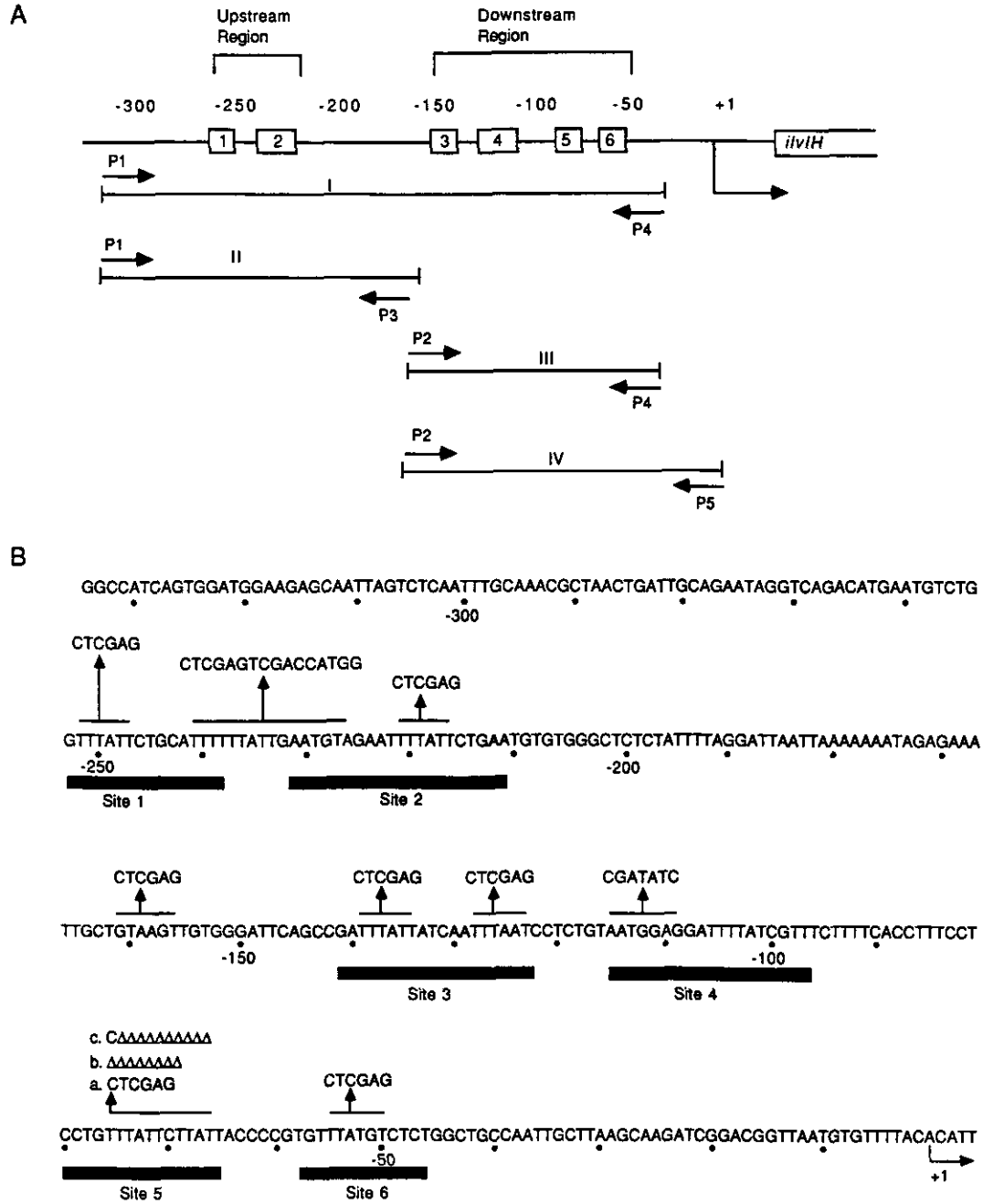


Figure 1. The region upstream from the *ilvIH* promoter. A, Schematic representation. Rectangles, Lrp binding sites; +1, transcription start site; I, II, III and IV, DNA fragments used for gel retardation and footprinting assays; P1, P2, P3, P4 and P5, primers used to prepare fragments I, II, III and IV by PCR. B, Nucleotide sequence. Filled rectangle, Lrp binding sites; vertical arrows, mutations; Δ, deletion of base-pair.

treated with MPE by adding 2 μl of a solution containing 1 mM-MPE, 1 mM-$Fe(NH_4)2(SO_4)_2 \cdot 6H_2O$ and 40 mM-DTT and incubated for 15 min at room temperature. The samples were immediately fractionated through a 5% polyacrylamide gel (40:1, acrylamide:bis acrylamide). Lrp–DNA complexes and free DNA, identified by autoradiography, were excised from the gel, eluted, and then analyzed on a 8% polyacrylamide/7·5 M-urea sequencing gel. The products of the G+A cleavage reactions (Maxam & Gilbert, 1980) were co-electrophoresed with samples to identify protected nucleotides. Quantitation was performed with a Betascope Blot Analyzer.

3. Results

(a) *Lrp binds with different affinities to widely separated regions upstream from the* ilvIH *promoter*

Ricca *et al.* (1989) showed that Lrp binds to at least two regions upstream from the *ilvIH* promoter, between −260 and −190 (hereafter called the upstream region) and between −100 and −40 (the downstream region). We measured the relative affinity with which purified Lrp binds to these two regions. PCR was used to prepare DNA fragments corresponding to the entire *ilvIH*

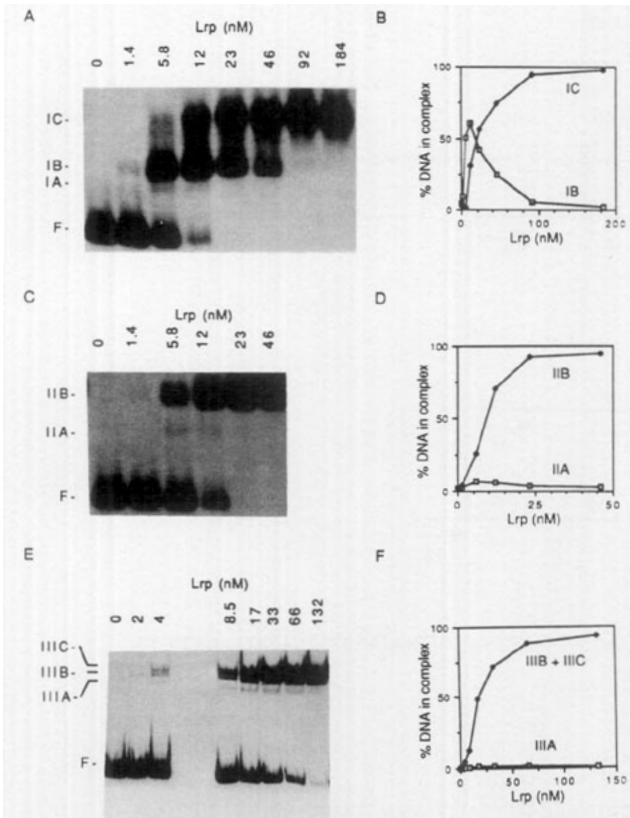


Figure 2. Gel retardation analysis of Lrp binding to DNA fragments I (A, B), II (C, D) and III (E,F). Purified Lrp was incubated with end-labeled DNA fragments under conditions defined in Materials and Methods. The concentration of Lrp shown is that calculated for the dimer. Samples were electrophoresed on 2% (w/v) agarose gels (A and C), or a 5% polyacrylamide gel (E), dried, and autoradiographed. Lrp–DNA complexes are identified at the left of panels A, C and E. For example, IIA and IIB are different complexes formed by DNA fragments II. F, free DNA. The amount of DNA in each complex in panels A, C and E was quantified with a Betascope Blot Analyzer and plotted in panels B, D and F, respectively, as a function of Lrp concentration.

promoter/regulatory region (Fig. 1A, fragment I), the upstream part of the region (fragment II), and the downstream part of the region (fragments III and IV). These PCR fragments were labeled and used in gel retardation experiments. Figure 2A shows the results of such experiments performed with fragment I and different concentrations of purified Lrp. Three Lrp–DNA complexes were observed, two present in large amounts (1C and 1B) and one present in small amount (1A). Footprinting studies established that complexes 1A and 1B contain Lrp bound at the upstream region only *whereas complex 1C contains Lrp bound to both* upstream and downstream regions (data not shown).

The concentration of Lrp that caused half of the DNA to become complexed was taken as the dissociation constant of Lrp–DNA complexes (molar concentration of Lrp 10-20 fold higher than DNA). The amount of label in individual complexes was quantified with a Betascope Blot Analyzer and plotted against the concentration of Lrp (Fig. 2B). The apparent dissociation constants for complex 1B (upstream region) and complex 1C (downstream region) were $5{\cdot}9\times10^{-9}$ M and $2{\cdot}2\times10^{-8}$ M, respectively. Thus, Lrp binds to the upstream region about four times more strongly than to the downstream region under these *in vitro* conditions. The gel retardation experiments described above were performed in the presence of 1 μg of calf thymus competitor DNA, an amount that is in 1000-fold excess over the labeled DNA fragment. Apparent dissociation constants determined for reaction mixtures lacking calf thymus DNA, were $3{\cdot}2\times10^{-9}$ M and $7{\cdot}9\times10^{-9}$ M for the upstream and downstream regions, respectively (data not shown). Thus, a 1000-fold excess of competitor DNA reduced Lrp binding only two to threefold, indicating that Lrp binds to specific sequences.

Tsai *et al.* (1989) derived an equation that can be used to estimate the degree of co-operativity for binding of the type:

$$\text{DNA} + \text{protein} \xrightarrow{K_{\text{I}}} \text{complex I}$$

$$\text{complex I} + \text{protein} \xrightarrow{K_{\text{II}}} \text{complex II.}$$

The degree of co-operativity $= K_{\text{I}}/K_{\text{II}} = [1/f_{\text{maxI}} - 1]^2$, where K_{I} and K_{II} are equilibrium dissociation constants and f_{maxI} is the maximum amount of complex I that is formed over the range of protein added. Applying the data in Figure 2B to this equation, $K_{1\text{B}}/K_{1\text{C}} = 0{\cdot}5$, a value that indicates that Lrp does not bind co-operatively to upstream and downstream binding regions.

(b) *The upstream region contains at least two sites that bind Lrp co-operatively*

MPE·Fe(II) footprinting was used to define Lrp binding sites within the upstream region. MPE is a small DNA-cleaving agent that intercalates in the minor groove of DNA and in the presence of ferrous ions and oxygen produces single-stranded breaks (Hertzberg & Dervan, 1984). MPE·Fe(II) cleaves DNA relatively non-specifically and, owing to its small size, gives relatively detailed footprints (Sawadogo & Roeder, 1985). In our experiments, fragment II (Fig. 1A), 5′ end-labeled on either the coding or non-coding strands, was incubated with Lrp and the sample was treated with MPE·Fe(II) before being subjected to electrophoresis through polyacrylamide. Bands corresponding to the two complexes were excised, and DNA within them denatured and fractionated on a DNA sequencing gel.

Such experiments were performed with the complexes formed between Lrp and fragment II labeled at the 5′ end of the coding or non-coding strand. Both in the absence (Fig. 2C) and the presence of MPE·Fe(II) (data not shown), two complexes were formed, called IIA and IIB. For complex IIB labeled at the 5′ end of the non-coding

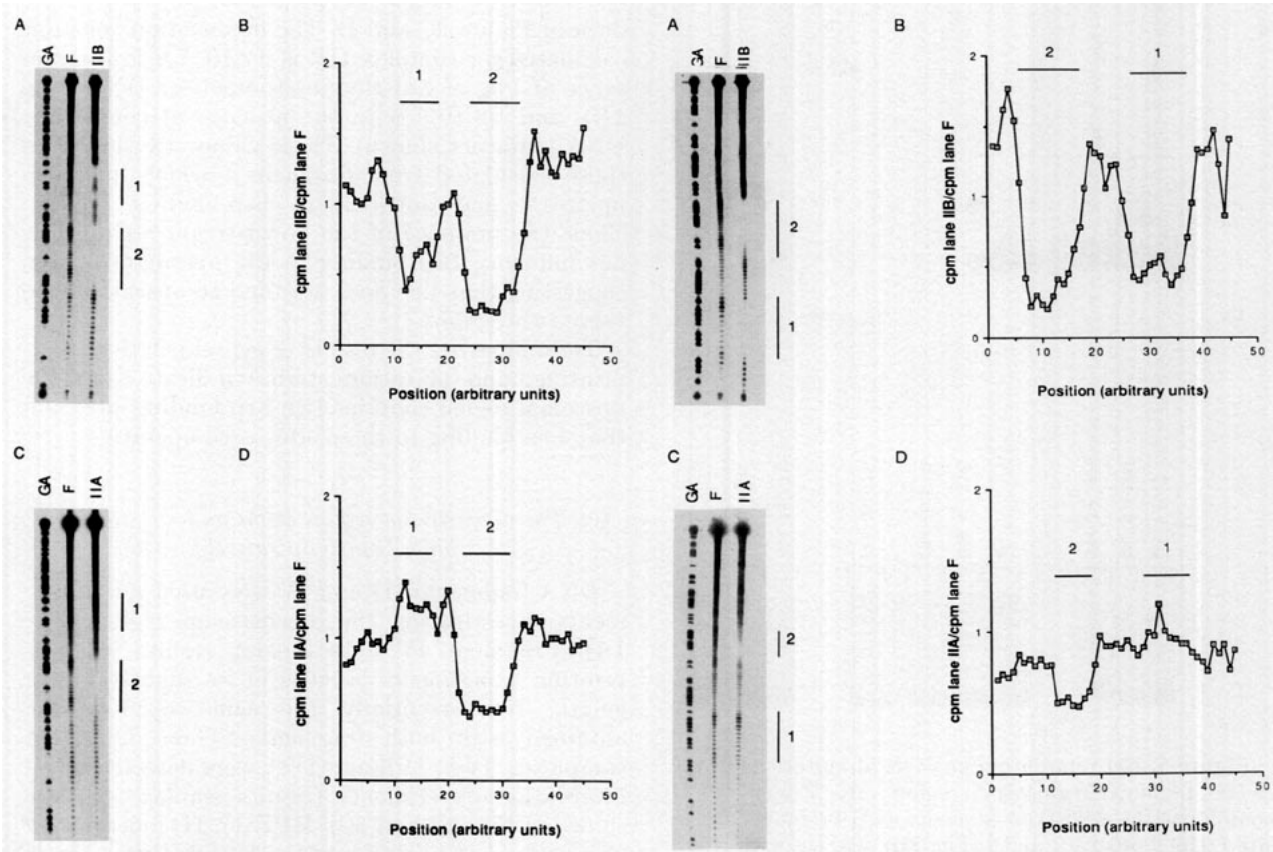


Figure 3. MPE·Fe(II) footprinting analysis of the non-coding strand of DNA fragment II. As used here, the non-coding strand has the same sequence as the mRNA. Fragment II was end-labeled at the *Hin*fI site (see sequence of PCR primer P3) with [^{32}P]dATP and reverse transcriptase. Lrp–DNA complexes were formed and then treated with MPE·Fe(II) as described in Materials and Methods and fractionated on a polyacrylamide gel. Complex IIB and free DNA (F) (see Fig. 2C for an example) were excised and fractionated on a denaturing polyacrylamide gel alongside a G+A sequencing ladder from the same fragment (GA), giving the pattern shown in A. C is the same as A except that complex IIA was analyzed. Short lines denote the positions of Lrp binding sites 1 and 2. Radioactivity in an Lrp–DNA complex and in free DNA was quantified with a Betascope Blot Analyzer and the ratio plotted against the position in the gel. The top of the gel in A and C relates to position 1 in B and D, respectively.

Figure 4. MPE·Fe(II) footprinting analysis of the coding strand of DNA fragment II. The coding strand is the DNA strand that is the template for transcription. Fragment II was end-labeled at the *Xba*I site (see sequence of PCR primer P1) with [^{32}P]dCTP and reverse transcriptase and analyzed as described in the legend to Fig. 3.

strand, two distinct protected regions were observed, called binding site 1 (−253 to −238) and 2 (−228 to −212) (Fig. 3A, B). Nucleotide sequences of these and other binding sites are in Figure 1B and Figure 10. Similar results were obtained with fragment II labeled at the 5′ end of the coding strand. Two protected regions, corresponding to sites 1 (−253 to −243) and 2 (−231 to −213), were detected in complex IIB (Fig. 4A and B).

The corresponding footprints for complex IIA are shown in Figure 3C and D for the non-coding strand and in Figure 4C and D for the coding strand. Only one region, corresponding to binding site 2, was protected. Cleavage by MPE·Fe(II) within the region of site 1, seemed, if anything, slightly enhanced. Thus, complex IIA has Lrp bound at site 2 whereas complex IIB has Lrp bound at both sites 1 and 2.

Lrp binding to sites 1 and 2 is highly co-operative. As shown in Figure 2C and D, complex IIA was visible at low concentrations of Lrp, but its concentration as a fraction of the total never exceeded 7%. Using equation $Kd_{IIA}/Kd_{IIB} = [1/f_{maxIIA} - 1]^2$ (Tsai *et al.*, 1989) and the data in Figure 2D, we determined that the relative dissociation constant of complex IIB is more than 175 times lower than that of complex IIA. Thus, Lrp binds with higher affinity to binding site 1 when neighboring binding site 2 is occupied.

To further explore the nature of the upstream region, separate mutations were introduced into sites 1 and 2. TTTATT within site 1 was changed to CTCGAG (Fig. 1B), and fragment II containing this mutation was prepared and used in a gel retardation experiment. At a concentration of Lrp that caused 62% of wild type fragment II to form complex IIB, only 12% of the mutant DNA formed this complex (Fig. 5A). Thus, the mutation introduced into site 1

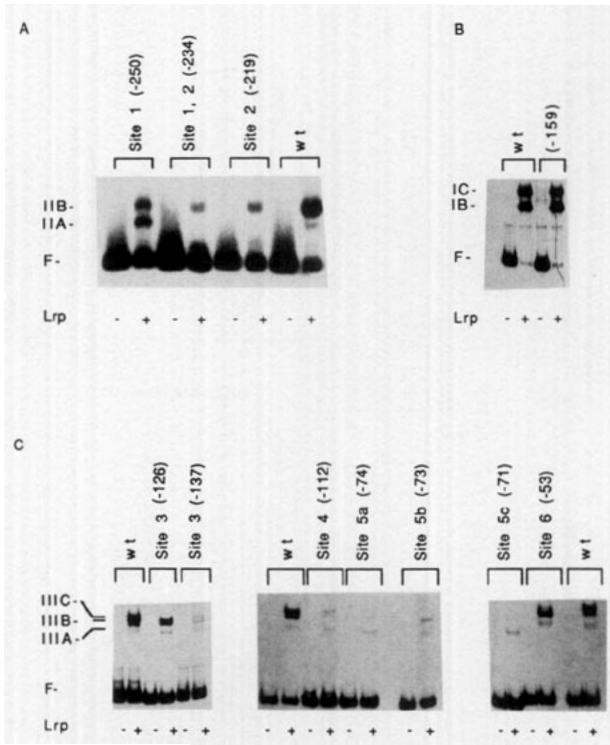


Figure 5. Gel retardation analysis of mutations in the *ilvIH* promoter/regulatory region. A, Fragment II containing the wild-type sequence (wt) or mutations in site 1, site 2 or sites 1 and 2 (Fig. 1B), was cut with *Xba*I, end labeled with [^{32}P]dCTP and reverse transcriptase, incubated without (−) or with (+) 25 ng of Lrp, and analyzed by gel retardation as described in Materials and Methods. Numbers in parentheses identify the locations of mutations (Fig. 1B). The positions of free DNA (F) and of Lrp–DNA complexes are indicated on the left. B, As in A except that fragment I containing a mutation located between sites 2 and 3 (Fig. 1B) was cut with *Eco*RI, end labeled with [^{32}P]dATP, and incubated with 50 ng of Lrp. (C) As in A except that fragment III containing mutations in sites 3, 4, 5, or 6 (Fig. 1B) was cut with *Eco*RI, end-labeled with [^{32}P]dATP, and incubated with 50 ng of Lrp.

markedly reduced the ability of Lrp to bind to site 1. Moreover, the amount of complex IIA formed by mutant DNA was relatively high, suggesting that this site 1 mutation substantially reduced co-operativity between site 1 and site 2.

In a parallel experiment in which a TTTATT within site 2 was changed to CTCGAG (Fig. 1B) about 12% of mutant DNA formed complexes (Fig. 5A), indicating that site 2 is essential for tight binding of Lrp to the entire upstream region. This result again suggests that there is a high degree of co-operativity between sites 1 and 2.

A third mutation was created which overlaps sites 1 and 2, spanning the region −241 to −226 (Fig. 1B, TTTTTTATTGAATGTA to CTCGAGTCGACCATGG). This mutation, like the site 2 mutation, dramatically reduced Lrp binding to the upstream region (Fig. 5A).

The apparent dissociation constant of Lrp binding to fragment II was determined from the data in Figure 2C and D. The dissociation constant calculated for complex IIB is 9×10^{-9} M in the presence of 1 μg of calf thymus competitor DNA (Fig. 2D), and 6×10^{-9} M in the absence of competitor DNA (data not shown). These values are similar to those calculated from fragment I which has both upstream and downstream Lrp-binding regions. Thus, the presence of the downstream region does not influence Lrp binding to the upstream region, suggesting that Lrp does not bind co-operatively to these two regions.

To summarize, the results of gel retardation, footprinting, and mutation studies indicate that the upstream region contains two Lrp binding sites and that Lrp binding to these sites is co-operative.

(c) *The downstream region contains four sites that bind Lrp co-operatively*

DNA fragments III and IV were used for experiments investigating the downstream region (Fig. 1A). Fragment IV gave clearer results in footprinting experiments because of its slightly longer length, but essentially the same results were obtained with both fragments. Three Lrp-DNA complexes, IVA, IVB and IVC, were detected in gel retardation experiments (results similar to those shown in Figs 2E and 5C). MPE·Fe(II) footprinting was performed using fragment IV labeled at the 5′ end of either the coding or the non-coding strand. Complexes IVB and IVC were too close to one another to be analyzed separately and therefore were analyzed together. For complexes IVB and C prepared from fragment IV labeled at the 5′ end of the non-coding strand, four distinct binding sites were detected, called sites 3 (−141 to −122), site 4 (−115 to −96), site 5 (−80 to −68), and site 6 (−57 to −46) (Figs 6A,B and 10). For complex IVA, on the other hand, only sites 3 and 4 were protected against MPE·Fe(II) cleavage (Fig. 6A, C). These results indicate that complex IVA contains Lrp bound to sites 3 and 4. Presumably, complex IVB has Lrp bound to sites 3, 4 and 5 or 3, 4 and 6 and complex IVC has Lrp bound to all four sites. Essentially the same results were obtained using fragment IV labeled at the 5′ end of the coding strand (Fig. 7). A summary of these results with fragment IV is shown in Figure 10.

The binding of Lrp to sites within the downstream region is highly co-operative (Fig. 2E, F). Complexes IIIBC were detectable at low concentrations of Lrp and increased in amount at progressively higher concentrations. Complex IIIA, on the other hand, remained at a low level, never exceeding 1·4% of the total. From the equation $Kd_{IIIA}/Kd_{IIIBC} = [1/f_{maxIIIA} - 1]^2$ (Tsai *et al.*, 1989) and the data in Figure 2F, the relative dissociation constant of complex IIIBC is about 5000-fold less than that of complex IIIA. Thus, Lrp binds with very high affinity to a pair of downstream binding sites when a pair of neighboring sites are occupied.

To further characterize the nature of Lrp binding in the downstream region, mutations were intro-

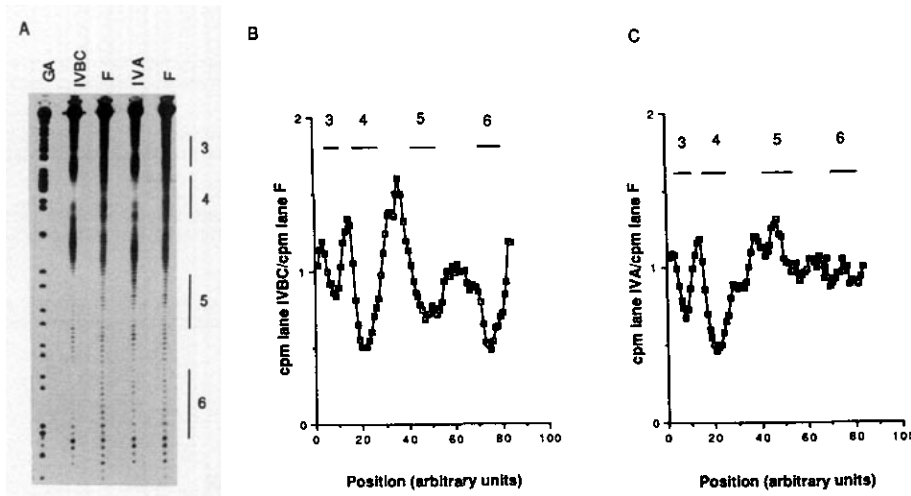


Figure 6. MPE·Fe(II) footprinting analysis of the non-coding strand of DNA fragment IV. Fragment IV was cut with *Eco*RI (see sequence of PCR primer P5) and end-labeled with [^{32}P]dATP and reverse transcriptase. Lrp–DNA complexes were formed and then treated with MPE·Fe(II) as described in Materials and Methods and fractionated on an acrylamide gel. Complexes IVA and IVBC and free DNA (see Fig. 2E for an example) were excised and fractionated on a denaturing polyacrylamide gel (A). Short lines denote the positions of Lrp binding sites 3, 4, 5 and 6. The quantification shown in B and C follows that described in the legend to Fig. 3.

duced into individual binding sites by site-directed mutagenesis. Binding site 3 has two sequences related to a putative consensus and therefore two mutations were made in it. For one mutation (termed site 3, −137), TTTATT at −139 to −134 is replaced by CTCGAG, and in the second (termed site 3, −126), TTTAAT at −128 to −123 is replaced by CTCGAG (Fig. 1B). Fragment III with each of these mutations was prepared and gel retardation experiments were performed to test the effect of the mutations on binding. As shown in Figure 5C, at a concentration of Lrp that caused 46% of wild type fragment III to form complexes, 15% of fragment III having the site 3 (−137) mutation and 33% the site 3 (−126) mutation formed complexes. These results suggest that sequences within the left half of binding site 3 are more important for Lrp recognition than sequences within the right half.

A mutation in site 4 changed AATGGAG to CGATATC at −115 to −109 (Fig. 1B). For binding site 5, three mutations were created: 5a (TTTATT to CTCGAG from −76 to −71), 5b (deletion from −76 to −69), and 5c (deletion from −75 to −65 and T to C change −76) (Fig. 1B). Fragment III containing these mutations was analyzed by gel retardation. As shown in Figure 5C, at a concentration of Lrp that caused 50% of wild type fragment III to form complexes, only relatively low percentages of complexes were formed from fragment III having mutations: 6%, site 4 mutation; 7%, site 5a mutation; 11%, site 5b mutation; and 4%, site 5c mutation.

To summarize, all of these results strongly suggest that Lrp binds with high co-operativity to sites 3, 4 and 5.

As opposed to the results above, a mutation in

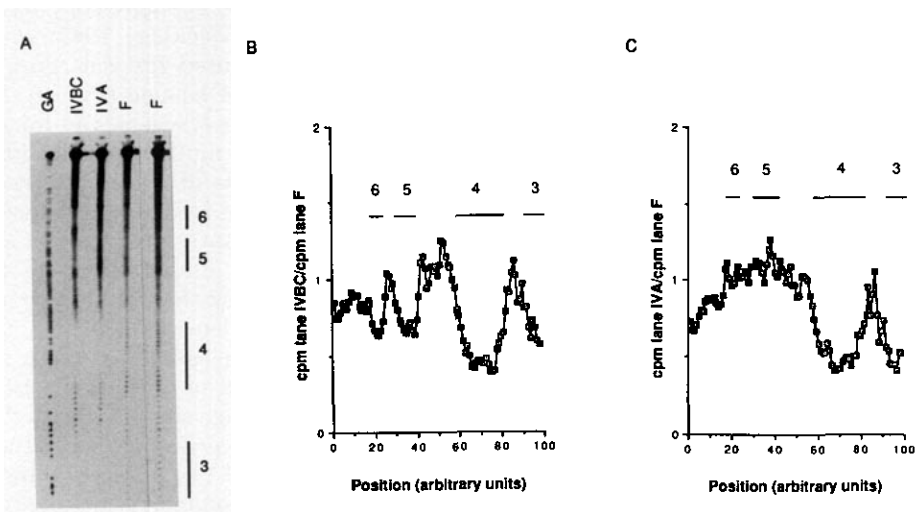


Figure 7. MPE·Fe(II) footprinting analysis of the coding strand of DNA fragment IV. Fragment IV was cut with *Hin*fI (see sequence of PCR primer P2) and end-labeled with [^{32}P]dATP and reverse transcriptase, and analyzed as described in the legend to Fig. 5.

site 6, a TTATGT to CTCGAG change at −55 to −50, did not affect the fraction of fragment III that was complexed with Lrp (Fig. 5C). However, the proportion of the mutant fragment III in complex IIIC was greatly reduced. These results are consistent with the idea that in complex IIIB, Lrp is bound to sites 3, 4 and 5 and that in complex IIIC, Lrp is bound, in addition, to site 6.

In addition, a mutation was created between the upstream and downstream region. This mutation, a GTAAGT to CTCGAG change between −161 and −156 (Fig. 1B), did not affect Lrp binding to the entire *ilvIH* regulatory region (Fig. 5B).

The apparent dissociation constant of the Lrp-DNA fragment III complex was 2×10^{-8} M in the presence of calf thymus competitor DNA (Fig. 2F) and $7{\cdot}9 \times 10^{-9}$ M in the absence of competitor DNA. Both values are close to those obtained using fragment I, which has both upstream and downstream regions. Thus, the presence of the upstream region does not affect the dissociation constant for the downstream region, again suggesting that there is no co-operativity between upstream and downstream regions.

(d) *Lrp binding to chemically-synthesized binding sites*

Binding sites 1, 2, 4 and 5 were chemically synthesized as pairs of complementary single strands (Fig. 8A). These synthetic binding sites were labeled and used in gel retardation experiments. As shown in Figure 8B, synthetic site 2 bound relatively strongly to Lrp in the absence of calf thymus competitor DNA. The addition of 1 μg of competitor DNA reduced Lrp binding by 20-fold, but Lrp binding was still detectable under these conditions (data not shown). Weak binding to site 4 was also detected (Fig. 8B), but this binding disappeared in the presence of competitor DNA (data not shown). A synthetic DNA fragment, similar to site 4 (Fig. 8A) except that the middle 20 base pairs had a random DNA sequence, did not bind Lrp, even in the absence of competitor DNA (data not shown). No Lrp binding was detected to synthetic sites 1 or 5 (Fig. 8B) or to a 31-mer synthetic fragment having a sequence unrelated to Lrp binding sites (data not shown). These results indicate that the observed binding to synthetic sites 2 and 4 is not the result of non-specific binding.

In summary, single binding sites show relatively weak binding or no binding to Lrp. The strong binding associated with the wild type sequence seems to be due to sites being organized in tandem. These results are consistent with the co-operative binding results described above.

(e) *Effects of mutations on* ilvIH *promoter expression* in vivo

To estimate the contribution of Lrp binding sites to *in vivo* expression of the *ilvIH* operon, wild type

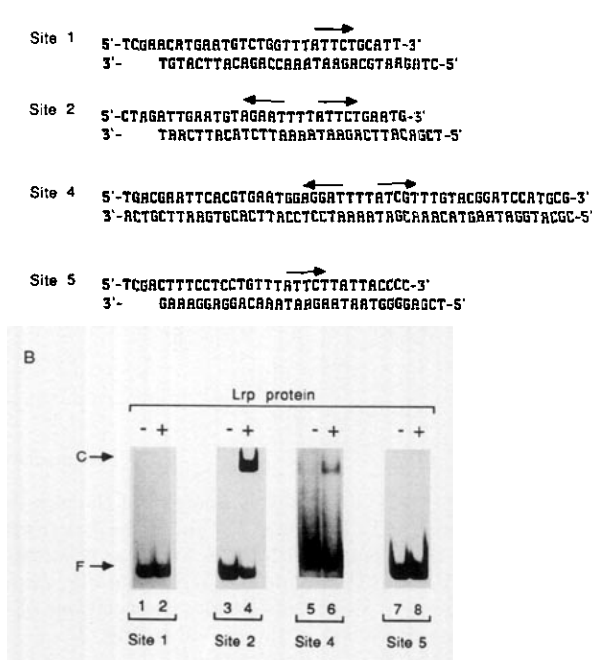
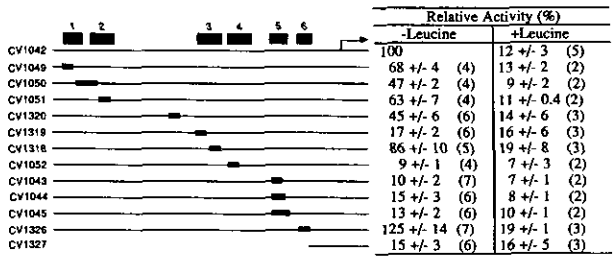


Figure 8. Lrp binding to chemically-synthesized sites. A, Nucleotide sequences of chemically-synthesized binding sites. Half-palindromic motifs are indicated by arrows. B, Gel retardation analysis of Lrp–DNA complexes with chemically-synthesized Lrp binding sites. ^{32}P-labeled synthetic binding sites (2 ng) were incubated without (−) or with (+) 50 ng of purified Lrp in a total volume of 20 μl under the conditions described in Materials and Methods except that no calf thymus DNA was added. The samples were fractionated through 5% polyacrylamide gels, dried and autoradiographed. F, free DNA; C, Lrp–DNA complex.

and mutant promoter/regulatory regions were attached to a *lacZ* reporter gene as transcriptional fusions on a plasmid. These constructs were transferred to a derivative of phage lambda by homologous recombination and then introduced into the *E. coli* chromosome in single copy by forming lysogens (Simons & Kleckner, 1987). β-Galactosidase activity was measured for cells grown in the presence or absence of leucine (Fig. 9). Measurements were made both for lysogens and for strains having the *lacZ* fusions on multicopy plasmids. The conclusions from both sets of measurements are the same and only data from lysogens are shown.

Strains having mutations in either site 1 or site 2, or spanning both sites 1 and site 2, had about twofold less activity than wild type and they were still regulated by leucine. The fact that the more extensive mutation affecting both sites had the same phenotype as mutations in either single site suggests that neither site 1 nor site 2 alone contributed to *ilvIH* expression but rather the simultaneous presence of two adjacent sites was required for Lrp action. In summary, the upstream region is not absolutely essential but it contributes to maximal expression from the promoter.

The mutation in site 3 centered at −137, which



	Relative Activity (%)	
	-Leucine	+Leucine
CV1042	100	12 +/- 3 (5)
CV1049	68 +/- 4 (4)	13 +/- 2 (2)
CV1050	47 +/- 2 (4)	9 +/- 2 (2)
CV1051	63 +/- 7 (4)	11 +/- 0.4 (2)
CV1320	45 +/- 6 (6)	14 +/- 6 (3)
CV1319	17 +/- 2 (6)	16 +/- 6 (3)
CV1318	86 +/- 10 (5)	19 +/- 8 (3)
CV1052	9 +/- 1 (4)	7 +/- 3 (2)
CV1043	10 +/- 2 (7)	7 +/- 1 (2)
CV1044	15 +/- 3 (6)	8 +/- 1 (2)
CV1045	13 +/- 2 (6)	10 +/- 1 (2)
CV1326	125 +/- 14 (7)	19 +/- 1 (3)
CV1327	15 +/- 3 (6)	16 +/- 5 (3)

Figure 9. Effect of mutations in the *ilvIH* promoter/regulatory region upon operon expression *in vivo*. Mutations defined in Fig. 1B were transferred to phage λRS45 in front of a *lacZ* gene and single copy lysogens were isolated in *E. coli* strain P90C. The strain designations of the resulting lysogens are shown at the left. The positions of the mutations (filled rectangles) are shown relative to the Lrp binding sites (shaded rectangles). The arrow defines the initiation site of transcription. Cells were grown in a minimal medium either lacking (−Leucine) or containing (+Leucine) 200 μg of leucine/ml. Operon expression was measured as β-galactosidase activity and values are normalized to the wild-type grown in the absence of leucine. 100% corresponds to about 485 Miller units. Values shown are averages followed by standard deviations and the number of experiments in parentheses.

reduced Lrp binding to the downstream region significantly, and reduced *ilvIH* expression about sixfold (Fig. 9, CV1319). The other mutation in site 3, which reduced Lrp binding only slightly, decreased *ilvIH* expression by 14% (Fig. 9, CV1318). Thus, for this case there is a correlation between Lrp binding affinity and *in vivo* expression of the *ilvIH* operon. The mutation in site 4, and the three mutations in site 5, all of which strongly reduced Lrp binding, reduced *ilvIH* expression 7 to 11-fold.

As opposed to the other mutations described, the mutation in site 6 did not affect *ilvIH* expression, indicating that Lrp binding to site 6 is not required for *ilvIH* expression. To summarize, three of the four downstream Lrp binding sites seem to be crucial for *ilvIH* promoter expression.

The results described above can be compared to those for the mutation in strain CV1327, in which the entire region containing the six Lrp binding sites was deleted (Fig. 9). This mutation reduced expression from the *ilvIH* promoter about sixfold, both for cells grown with and without leucine. Thus, mutations in either sites 3, 4 or 5 have about the same effect on expression as a deletion that removes the entire region, suggesting that Lrp binding to sites 3, 4 and 5 act synergistically *in vivo*. Furthermore, the fact that the residual activity in strain CV1327 was no longer affected by leucine indicates that the same region of DNA is involved in both Lrp binding and leucine-mediated repression.

Surprisingly, a mutation located between the upstream and downstream regions reduced *ilvIH* expression twofold, an effect comparable to that caused by mutations within sites 1 or 2 (Fig. 9, CV1320). The mutation, a substitution of GTAAGT by CTCGAG at positions −161 to −156 (Fig. 1B), did not affect binding of Lrp *in vitro* to either upstream or downstream regions (Fig. 5B) and thus must affect Lrp–DNA or Lrp–Lrp interactions in some subtle way. Indeed, DNase I footprinting showed some protection within this region (data not shown).

4. Discussion

(a) *The effect of Lrp on* ilvIH *expression is direct*

Altogether, some 16 operons are known to be under control of Lrp, not including a sizable number of anonymous genes identified by a two-dimensional gel electrophoresis study and by a study involving *plac*Mu mutagenesis (Ernsting *et al.*, 1992; Lin *et al.*, 1992*a*). An important question for each of these cases is whether Lrp acts directly as an activator or a repressor, or whether its effects are only indirect, as would be the case if it acted on another regulatory gene which in turn affected target genes. Here we show that at least for the *ilvIH* operon, Lrp acts directly. Lrp binds to six different sites upstream from the *ilvIH* promoter, and mutations within five of those sites both reduce binding to the respective sites and reduce expression from the *ilvIH* promoter. This conclusion that Lrp acts directly is corroborated by studies showing that *in vitro*, purified Lrp increases transcription from the *ilvIH* promoter two- to fivefold (D. A. Willins & J. M. Calvo, unpublished data).

(b) *Determinants of Lrp binding specificity*

How does Lrp recognize specific sites upstream from the *ilvIH* promoter? In Figure 10, the sequences of six Lrp binding sites are shown, aligned in a way that emphasizes a common sequence element. Sites 2 and 4, the strongest Lrp binding sites as measured by gel retardation assays (Fig. 8), both contain dyad symmetry. Site 2 has a perfect palindrome, 5′-AGAATtttATTCT-3′, with the two inverted repeats being separated by a spacer of three T residues. Site 4 is similar to site 2 but contains an imperfect palindrome, 5′-AGGATtt-tATCGT-3′, with again two inverted repeats being separated by a spacer of three T residues (an underline denotes a difference from a putative consensus). The fact that site 4 is a weaker Lrp binding site than site 2 may be related to the fact that it contains an imperfect palindrome. We did not detect Lrp binding to synthetic sites 1 or 5 (Fig. 8). However, footprinting and gel retardation experiments demonstrated that Lrp binds to sites 1 or 5 when they are located about 30 bp from sites 2 or 4, respectively. Neither site 1 nor site 5 has a palindromic sequence of the type described above but they both contain TTT followed by a perfect half-palindromic motif, 5′-ATTCT-3′. Site 3 also contains TTT followed by an imperfect half-palindromic motif 5′-ATTAT-3′, and it is located about 30 bp from a site that has two palindromic motifs (site 4).

Figure 10. *Alignment of the 6 Lrp binding sites identified* by MPE·Fe(II) footprinting analyses. Sequences protected by Lrp against MPE·Fe(II) cleavage and their positions are shown together with a consensus sequence derived from the alignment. The arrow indicates a half-palindromic motif and asterisks denote differences from the consensus.

Thus, the requirement for tight binding may be a primary binding site, containing two palindromic motifs, together with half palindromic sites located about 30 base pairs away. With these considerations in mind, site 6, the site that does not seem to be important for Lrp action, can be compared to the other sites. Site 6, TTTATGTC, has only 5/8 bp matching the consensus and it is located 20 base pairs from site 5 and 50 base pairs from site 4.

The tentative consensus 5′-AGAATtttATTCT-3′, derived from an analysis of six binding sites upstream from the *ilvIH* promoter, will be tested when other Lrp binding sites are determined. The fact that MPE·Fe(II) footprints of individual Lrp binding sites are larger than this core sequence suggests that additional, less well-conserved residues on both sides of the consensus sequences contribute to the specificity of Lrp binding. Furthermore, the fact that a mutant binding site 1 having only a relatively poor match to the consensus still bound Lrp at that site suggests that key aspects of the consensus are not defined by this analysis of six binding sites. It should also be kept in mind that Lrp, because it shows so many different regulatory patterns, may have a particularly complex consensus, or that it may have more than a single consensus. Clearly, additional studies involving Lrp binding sites within other operons, extensive mutational analyses of a single binding site, and "Selex" type selection experiments (Tuerk & Gold, 1990) will be required to define the nature of Lrp binding sites.

(c) *Co-operative binding of Lrp may contribute to the mechanism of activation*

One conclusion from this study is that adjacent Lrp binding sites interact synergistically *in vivo*. By synergism we mean that two or more sites co-operate to give an effect that is greater than the sum of the isolated sites. Alteration of either site 1 or 2 abolished the contribution of the entire upstream region to the activation of transcription from the *ilvIH* promoter. Similarly, alteration of site 3, or site 4, or site 5 completely abolished the contribution of the downstream region to the activation of transcription. These synergistic interactions observed *in vivo* are likely mediated by the high degree of co-operative binding of Lrp to neighboring sites demonstrated *in vitro*. The degree of co-operativity was estimated to be more than 100-fold for binding to sites 1 and 2 and about 5000-fold for binding to sites 3, 4 and 5.

Co-operative binding of a regulatory protein to multiple sites can increase the overall size and therefore specificity of the effective binding site. This may be important for a site as small as tttATTCT, which can occur by chance in genes unassociated with Lrp. Similar situations exist for other systems, including the simian virus 40 early promoter (six GC motifs) (Barrera-Saldana *et al.*, 1985), *Drosophila* heat shock promoters (multiple AGAAA units) (Xiao *et al.*, 1991), and for promoters in yeast that are subject to general amino acid control (multiple TGACTC sites) (Arndt & Fink, 1986; Hope & Struhl, 1985).

We suggest here another role for the co-operative binding of Lrp. Lrp bound at sites 3 or 4 (positions −136 and −103, respectively) would not be expected to directly contact RNA polymerase bound at the *ilvIH* promoter unless the intervening DNA were severely bent. However, because of co-operativity, Lrp molecules bound at sites 3 and 4 enhance binding of Lrp to site 5, located at −72, and Lrp bound at −72 can potentially interact directly with an adjacent RNA polymerase. Thus, an important role for co-operativity in this system may be to efficiently deliver Lrp to a critical binding site, that we propose here to be site 5.

It is noteworthy that site 6, centered at −52, is positioned closest to the RNA polymerase binding site but it is not required for *ilvIH* promoter expression. The relationship between the position of CRP binding and the extent of activation was determined for several promoters. For the *melR* promoter, CRP activated transcription strongly when its binding site was located at −41·5 and −61·5, but only weakly when it was at −51·5 or −52·5 (Gaston *et al.*, 1990). In the case of the *lac* promoter, CRP activated transcription with a binding site positioned at −61·5, −71·5, or −82·5, but not at −51·5 (Ushida & Aiba, 1990). It is interesting that two different activators, Lrp and

CRP, show a similar position dependence for functional binding sites.

(d) *The role of upstream binding sites in transcriptional activation of the* ilvIH *operon*

Mutations in either site 1 or site 2 reduced expression from the *ilvIH* promoter about twofold. How do sites 1 and 2, located at −250 and −215, respectively, affect transcription initiation? One possibility is that through looping of the intervening DNA, Lrp bound at sites 1 and 2 directly contacts RNA polymerase. In considering this possibility, it is instructive to refer to the analysis of bacterial promoters by Collado-Vides *et al.* (1991) and Gralla (1991). They suggested that *E. coli* and *Salmonella typhimurium* promoters that were positively controlled could be divided into two groups, one of them being a large collection of sigma 70 promoters in which activation can be understood in terms of direct interaction between neighboring activator and RNA polymerase (Collado-Vides *et al.*, 1991; Gralla, 1991). The other group was made up of a much smaller number of sigma 54 promoters. For this latter group, binding sites for activator proteins were located far upstream from the promoter, and could be moved even further away without destroying activator function (Collado-Vides *et al.*, 1991; Gralla, 1991). Activation of these sigma 54 promoters was imagined to occur through a looping mechanism bringing the activator in contact with RNA polymerase. Some recent *in vitro* transcription experiments employing purified RNA polymerase strongly suggest that the *ilvIH* promoter is recognized by a sigma 70 promoter (D. A. Willins & J. M. Calvo, unpublished data). Furthermore, experiments performed by E. Ricca & M. Sacco and their colleagues and confirmed by us have shown that moving the block of six Lrp binding sites further upstream destroys the ability of Lrp to activate *ilvIH* expression *in vivo* (personal communication from E. Ricca & M. Sacco; Q. Wang & J. M. Calvo, unpublished data). Taken together, these results suggest that Lrp bound at sites 1 and 2 does not directly interact with RNA polymerase.

Another possibility is that Lrp bound at sites 1 and 2 interacts with Lrp bound at sites 3, 4 and 5. This idea is consistent with the result that mutations in sites 3, 4 and 5 have about the same effect on promoter expression as a deletion that removes the entire *ilvIH* regulatory region. This result suggests that the function of the upstream region requires the presence of the downstream region. Formation of a complex involving the interaction of Lrp bound at upstream and downstream sites requires that the DNA between sites 2 and 3 form a loop. The fact that the DNA upstream from *ilvIH* is naturally bent (center of bending at −120), and that binding of Lrp further induces bending (Q. Wang & J. M. Calvo, unpublished data), may contribute to loop formation. Furthermore, the results of DNase I footprinting studies suggest that Lrp perturbs the structure of the DNA that lies between sites 2 and 3 (Q. Wang & J. M. Calvo, unpublished data). Finally, the fact that a cluster of five base pair substitutions located between −161 and −156 affected promoter expression is most easily explained by assuming some interaction between upstream and downstream binding regions.

Assuming that Lrp bound to sites 1 and 2 can become part of a larger complex, how does that affect transcription initiation? One possibility, that sites 1 and 2 interact co-operatively with critical site 5, appears unlikely, given the fact that no co-operativity was observed between the binding of Lrp to upstream and downstream regions. Another possibility is that a large complex can form, but that it has only marginal stability (but sufficient to affect cleavage by DNase I). Within the larger complex, Lrp bound to the critical site 5 may be better positioned as an activator, thus accounting for the twofold increase in activation attributed to sites 1 and 2.

This work was supported by grant GM39496 from the National Institutes of Health. We are especially indebted to D. A. Willins for purified Lrp. We thank R. W. Simons for plasmids, phage and strains, P. B. Dervan for MPE, and J. D. Helmann and J. T. Lis for a critical reading of the manuscript.

References

Arndt, K. & Fink, G. R. (1986). GCN-4 protein, a positive transcription factor in yeast, binds general control promoters at all 5′ TGACTC 3′ sequences. *Proc. Nat. Acad. Sci., U.S.A.* **83**, 8516–8520.

Austin, E. A., Andrews, J. C. & Short, S. A. (1989). *Molecular Genetics of Bacteria and Phages*, p. 153, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Barrera-Saldana, H., Takahashi, K., Vigneron, M., Wildeman, A., Davidson, I. & Chambon, P. (1985). All six GC-motifs of the SV40 early upstream element contribute to promoter activity *in vivo* and *in vitro*. *EMBO J.* **4**, 3839–3849.

Braaten, B. A., Platko, J. V., Van Der Woude, M. W., Simons, B. H., De Graaf, F. K., Calvo, J. M. & Low, D. A. (1992). Leucine-responsive regulatory protein controls the expression of both the *pap* and *fan* pili operons in *Escherichia coli*. *Proc. Nat. Acad. Sci., U.S.A.* **89**, 4250–4254.

Collado-Vides, J., Magasanik, B. & Gralla, J. D. (1991). Control site location and transcriptional regulation in *Escherichia coli*. *Microbiol. Rev.* **55**, 371–394.

DeFelice, M., Lago, C. T., Squires, C. H. & Calvo, J. M. (1982). Acetohydroxy acid synthase isoenzymes of *Escherichia coli* K12 and *Salmonella typhimurium*. *Annal. Microbiol. (Inst. Pasteur)*, **133A**, 251–256.

deWind, N., deJong, M., Meijer, M. & Stuitje, A. R. (1985). Site-directed mutagenesis of the *Escherichia coli* chromosome near *oriC*: identification and characterization of *asnC*, a regulatory element in *E. coli* asparagine metabolism. *Nucl. Acids Res.* **13**, 8797–8811.

Dodd, I. B. & Egan, J. B. (1987). Systematic method for the detection of potential lambda cro-like DNA-binding regions in proteins. *J. Mol. Biol.* **194**, 557–564.

Ernsting, B. R., Atkinson, M. R., Ninf, A. J. & Matthews, R. G. (1992). Characterization of the regulon controlled by the leucine-responsive regulatory protein in *Escherichia coli*. *J. Bacteriol.* **174**, 1109–1118.

Gaston, K., Bell, A., Kolb, A., Buc, H. & Busby, S. (1990). Stringent spacing requirements for transcription activation by CRP. *Cell*, **66**, 733–743.

Gazeau, M., Delort, F., Dessen, P., Blanquet, S. & Plateau, P. (1992). *Escherichia coli* leucine-responsive regulatory protein Lrp controls lysyl-tRNA synthetase expression. *FEBS Letters*, **300**, 254–258.

Gralla, J. D. (1991). Transcriptional control-lessons from an *E. coli* promoter data base. *Cell*, **66**, 415–418.

Haney, S. A., Platko, J. V., Oxender, D. L. & Calvo, J. M. (1992). Lrp, a leucine-responsive protein, regulates branched-chain amino acid transport genes in *Escherichia coli*. *J. Bacteriol.* **174**, 108–115.

Haughn, G. W. (1985). An analysis of the region between the *leuABCD* and *ilvIH* operons in *Escherichia coli*. Ph.D. thesis, Cornell University, Ithaca, NY.

Haughn, G. W., Squires, C. H., DeFelice, M., Lago, C. T. & Calvo, J. M. (1985). Unusual organization of the *ilvIH* promoter of *Escherichia coli*. *J. Bacteriol.* **163**, 186–198.

Hertzberg, R. P. & Dervan, P. B. (1984). Cleavage of DNA with methidiumpropyl-EDTA-Iron (II): reaction conditions and product analysis. *Biochemistry*, **23**, 3934–3945.

Hope, I. A. & Struhl, K. (1985). GCN-4 protein, synthesized *in vitro*. binds HIS3 regulatory sequences: Implications for general control of amino acids biosynthetic genes in yeast. *Cell*, **43**, 177–188.

Kolling, R. & Lother, H. (1985). AsnC: an autogenously regulated activator of asparagine synthetase A transcription in *Escherichia coli*. *J. Bacteriol.* **164**, 310–315.

Kunkel, T. A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Nat. Acad. Sci., U.S.A.* **82**, 488–492.

Kunkel, T. A., Roberts, J. D. & Zakour, R. A. (1987). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* **154**, 367–382.

Landolfi, N. F., Yin, X., Capra, J. D. & Tucker, P. (1989). Protection analysis (or "footprinting") of specific protein-DNA complexes in crude nuclear extracts using methidiumpropyl-EDTA-Iron (II). *BioTechniques*, **7**, 500–504.

Lin, R., D'Ari, R. & Newman, E. B. (1990). The leucine regulon of *Escherichia coli* K-12: a mutation in *rblA* alters expression of L-leucine-dependent metabolic operons. *J. Bacteriol.* **172**, 4529–4535.

Lin, R., D'Ari, R. & Newman, E. B. (1992*a*). λ *plac*Mu insertions in genes of the leucine regulon: extension of the regulon to genes not regulated by leucine. *J. Bacteriol.* **174**, 1948–1955.

Lin, R., Ernsting, B., Hirshfield, I. N., Matthews, R. G., Neidhardt, F. C., Clark, R. L. & Newman, E. B. (1992*b*). The *lrp* gene product regulates expression of *lysU* in *Escherichia coli* K-12. *J. Bacteriol.* **174**, 2779–2784.

Maxam, A. M. & Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**, 499–560.

Miller, J. H. (1972). *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Newman, E. B., D'Ari, R. & Lin, R. T. (1992). The leucine-Lrp regulon in *E. coli*: a global response in search of a raison d'Etre. *Cell*, **68**, 18–20.

Platko, J. V., Willins, D. A. & Calvo, J. M. (1990). The *ilvIH* operon of *Escherichia coli* is positively regulated. *J. Bacteriol.* **172**, 4563–4570.

Rex, J. H., Aronson, B. D. & Somerville, R. L. (1991). The *tdh* and *serA* operons of *E. coli*: mutational analysis of the regulatory elements of leucine-responsive genes. *J. Bacteriol.* **173**, 5944–5953.

Ricca, E., Aker, D. A. & Calvo, J. M. (1989). A protein that binds to the regulatory region of the *ilvIH* operon of *Escherichia coli*. *J. Bacteriol.* **171**, 1658–1664.

Rosenthal, E. R. & Calvo, J. M. (1987). Transcription termination sites at the distal end of the *leu* operon of *Salmonella typhimurium*. *J. Mol. Biol.* **194**, 443–452.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5463–5467.

Sawadogo, M. & Roeder, R. G. (1985). Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell*, **43**, 165–175.

Simons, R. W. & Kleckner, N. (1987). Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. *Gene*, **53**, 85–96.

Squires, C. H., DeFelice, M., Wessler, S. R. & Calvo, J. M. (1981). Physical characterization of the *ilvHI* operon of *Escherichia coli* K12. *J. Bacteriol.* **147**, 797–804.

Tsai, S. Y., Tsai, M. & O'Malley, B. W. (1989). Cooperative binding of steroid hormone receptors contributes to transcriptional synergism at target enhancer elements. *Cell*, **57**, 443–448.

Tuerk, C. & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, **249**, 505–510.

Ushida, C. & Aiba, H. (1990). Helical phase dependent action of CRP: effect of the distance between the CRP site and the −35 region on promoter activity. *Nucl. Acids Res.* **18**, 6325–6330.

Willins, D. A., Ryan, C. W., Platko, J. V. & Calvo, J. M. (1991). Characterization of Lrp, an *Escherichia coli* regulatory protein that mediates a global response to leucine. *J. Biol. Chem.* **266**, 10768–10774.

Xiao, H., Perisic, O. & Lis, J. T. (1991). Cooperative binding of *Drosophila* heat shock factor to arrays of a conserved 5 bp unit. *Cell*, **64**, 585–593.

Edited by M. Gottesman