Characterization of the Leucine/Lrp regulon in *Escherichia coli* K-12

RONGTUAN LIN

A Thesis

In

The Special Individual

Program

Presented in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy at

Concordia University

Montreal, Quebec, Canada

May 1992

ISBN 0-315-80974-4

ABSTRACT

Characterization of the Leucine/Lrp regulon in *Escherichia coli* K-12

Rongtuan Lin, PH.D.

Concordia University, 1992.

A highly pleiotropic *Escherichia coli* mutant, *lrp*::Tn*10*, was isolated and characterized. The *lrp* gene product was shown to be a regulator of a global response to L-leucine. 27 strains with λp*lac*Mu insertion in genes whose expression is affected by L-leucine were isolated and characterized. All of them are regulated by Lrp. Lrp and L-leucine influenced gene expression in a surprising variety of ways. Two genes that are regulated by Lrp and not affected by L-leucine were identified. It was demonstrated that Lrp positively regulated the expression of *serA*, *sdaB*, *leu*, and *gcv*; and negatively regulated the expression of *sdaA*, *tdh*, *ilvJ/K*, and *lrp*.

The gel retardation assays demonstrated that purified Lrp protein could bind to the upstream region of *lysU*, *serA*, *sdaA*, *gcv*, and *lrp*. *In vitro* footprinting experiments showed that Lrp protected against DNase I digestion over a long region. Primer extension analysis indicated that the overall positive effect by Lrp on the transcription of *serA* involved both activation and inhibition by Lrp on the dual *serA* promoters.

ACKNOWLEDGEMENTS

I would very much like to thank Dr. E. B. Newman for her guidance, financial support and providing me the best academic support throughout this study. I also feel indebted to her for motivating and encouraging me to accept the challenge of academic work and giving a lot of help since I met her in Beijing in 1985.

My sincere thanks go also to my committe members, Drs. R. K. Storms and M. B. Herrington, for their helpful advice ang suggestions.

Special thanks to Drs. D. A. Willins and J. M. Calvo for providing the purified Lrp protein.

Many thanks to Drs. R. D'Ari, M. Howe, J. M. Calvo, M. DuBow, R. G. Matthews, I. N. Hirshfield, E. McFall, G. A. Grant, G. V. Stauffer, C. Cupples, and P. Gulick for their insterest, advices and providing the strains and plasmids.

Thanks are extended to my colleagues in this lab, H, Su, J. Garnon, E. Brown, Z, Shao, Y. Shen, Y. Li, J. Moniakis, M. San Martano, J. Zhang, K. Shanmugan, G. Ambartsoumian, and A. Beeser for friendship, help, and discussions.

I would like to thank S. Beeser for reading and correction part of this thesis.

I must also thank Concordia University for providing me the International Student Fee Remission.

Finally, very special thanks goes to my wife and my parents for their understanding and support.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

In *Escherichia coli* and other enteric bacteria, synthesis of most proteins is under direct environmental control. Relatively few proteins are synthesized at a preset, environmentally independent rate, i.e. constitutively. The cell has a number of programmed responses to particular types of environmental changes resulting in the induction or repression of particular groups of genes, i.e. operons. There are many possible levels of control in the complex processes of transcribing genes and synthesizing appropriate amounts of gene products. However, regulation at the transcriptional level is the most studied, and perhaps the most direct way to selectively control gene expression.

In this thesis, I describe a new global regulatory system, the leucine regulon in *E. coli*, which responds mainly to the presence of L-leucine in the environment. Expression of the genes of the leucine regulon is controlled by the leucine-responsive regulatory protein, Lrp, encoded by the *lrp* gene. Lrp, a DNA-binding protein, regulates the expression of a number of genes- in most cases acting together with L-leucine.

This work originated from the finding that a mutant, *rbl*::Tn*10*, selected by virtue of its increased L-serine deaminase (L-SD) activity, also showed altered activity of a wide variety of enzymes involved in different metabolic pathways, the synthesis of most of these enzymes being

regulated by L-leucine. Previously isolated mutants from a variety of different laboratories, including *lhb*, *livR*, and *oppI*, have now been shown to be in the same gene. The researchers working on these mutants have agreed that the gene should be designated *lrp*.

In this work I show that the deficiency in *lrp* function in an insertion mutant increased L-serine deaminase and L-threonine dehydrogenase activities, decreased expression of *serA*, *gcv* and *leu* operons, and suppressed the slow growth of a *metK* mutant. The expression of *lrp* is subjected to negative autoregulation and is repressed by growing in rich medium. I then describe experiments concerned with the molecular basis of regulation by *lrp*. Gel retardation and DNase I footprint results show that the Lrp protein binds to specific sites in the upstream regions of *sdaA*, *serA* and *lysU* genes *in vitro*. This is consistent with the suggestion that Lrp is a transcriptional regulator of these genes.

Since this work began as a study of the genetic regulation of L-SD activity, Part I of this introduction will describe the genetic and environmental factors which affect the synthesis of L-SD, and shows that the *sdaA* gene is regulated by at least two global regulators. In Part II I therefore survey global regulatory mechanisms in general, discussing in Part III, the details of molecular mechanisms used in this regulation. In the final section of the introduction, part IV, I give an account of the genes now known to be controlled by Lrp.

I. GENETIC AND ENVIRONMENTAL FACTORS WHICH AFFECT L-SERINE DEAMINATION IN *E. coli*

L-serine deaminase (L-SD) catalyses the conversion of the amino acid L-serine to an Embden Meyerhof intermediate, pyruvic acid [Pardee and Prestidge, 1955]. As will be discussed later, *E. coli* has two different genes coding for enzymes with L-SD activity, *sdaA* and *sdaB*. These enzymes are subjected to a very complex regulation.

L-serine deaminase 1, the *sdaA* gene product, has been studied in considerable detail. It can be induced by glycine and/or L-leucine, but not by its substrate, L-serine [Pardee and Prestidge, 1955; Isenberg and Newman, 1974]. Wild type *E. coli* K-12 is unable to use L-serine as sole carbon source even though it synthesizes large amounts of L-serine deaminase while it grows in glucose minimal medium. However, it can grow on L-serine as sole carbon source when L-leucine and/or glycine are also present in the medium [Newman and Walker, 1982]. It has been possible to isolate mutants deficient in L-SD activity by screening for strains which are unable to grow on a combination of L-serine, glycine and L-leucine as carbon source [Newman *et al.*, 1985a; Newman *et al.*, 1985b] and to isolate mutants with overproduction of L-SD by selecting for strains which can use L-serine as sole carbon source [Newman *et al.*, 1981; Newman *et al.*, 1982a]. In this lab, seven genes have been found to affect L-SD activity.

I-1. L-SD Is Subject to Post-Translational Regulation

L-SD can be assayed in whole cells by the incubation of cells with a small amount of toluene to permeabilize cells, but it cannot be detected in cell extracts without prior activation by iron and dithiothreitol [Newman *et al.*, 1985a].

In the study of mutants which produce less L-SD activity, Newman and Su found that the strains carrying mutations in any of three different genes were deficient in L-SD activity *in vivo* (whole cell assay) and unable to use L-serine, glycine, and L-leucine as carbon source. However, the cell extracts from these three mutants (MEW128, MEW191, and MEW84) showed almost as much activity as the parent strain when activated with iron and dithiothreitol [Newman *et al.*, 1985a; Newman *et al.*, 1985b; Su, 1991]. It was surprising that all three mutants required thiamine for growth. It was suggested that L-SD is made in an inactive form, and enzymatically activated by an activating system *in vivo*, or chemically activated by iron and dithiothreitol *in vitro*. Three possible mechanisms of activation *in vivo* were suggested by Su [Su, 1991].

I-2. Two L-serine Deaminases in *E. coli*

While the first L-SD deficient mutants carried mutations affecting the post-translational activation system, recently mutants with defects in the structural gene have been isolated: strain MEW15, isolated by MudX insertion [Newman *et al.*, 1985a], and strain MEW22, a λP*lac*Mu9 insertion mutant [Su *et al.*, 1989].

The identification of these strains as carriers of structural gene mutations is suggested by the following two pieces of evidence. First neither strain MEW15 nor strain MEW22 showed L-SD activity, whether assayed *in vivo* or *in vitro*. Second, in both strains, the ß-galactosidase activity from the *lacZ* gene fusion was induced by glycine and L-leucine, the known inducers of L-SD 1.

The two mutations were mapped to the same locus of the *E. coli* chromosome indicating that they affect the same gene. The mutation in MEW22 was studied further and shown to be caused by an insertion in a gene named *sdaA* [Su *et al.*, 1989, Garnon, J. personal communication]. The *sdaA* gene has been cloned and sequenced [Su *et al.*, 1989], and shown to be the structural gene for L-SD1 [Su and Newman, 1991, Monlakis, J. personal communication].

The fact that these L-SD 1-deficient *sdaA* mutants showed low level L-SD activity when grown in minimal medium and had a great deal of L-SD activity when grown in Luria broth suggested that a second L-SD structural gene might be found in the *E. coli* K-12 genome [Newman *et al.*, 1985b]. A mutant carrying a λp*lac*Mu insertion in a gene called *sdaB* and a *sdaA*::Cm^r inactivating *sdaA* showed no L-SD activity when grown in either medium, suggesting that there are two, and only two, L-SD enzymes in E. coli. This was further investigated by showing that ß-galactosidase activity from a *sdaB*::*lacZ* fusion was regulated as L-SD#2, i.e. induced in LB. The *sdaB* gene has since been sequenced, and shows a very high level of

nucleotide and deduced amino acid sequence identity between *sdaA* and *sdaB*, suggesting that *sdaB* is the structural gene of L-SD2 [Su and Newman, 1991; Shao, Z.Q., personal communication]. Su and Newman showed that L-SD2 is similar to L-SD1 in physical parameters and is also subject to post-translational activation [Su and Newman, 1991].

I-3. L-SD-1 Activity is Regulated by Global Regulatory System(s).

Mutants with a considerable increase in L-SD activity can be isolated by direct selection for growth on L-serine as sole carbon source. Some of these carry a mutation in a very pleiotropic gene, *ssd*, unlinked to either structural gene. An *ssd* mutation considerably increased L-SD activity and allowed the strain to use L-serine as sole carbon source [Newman *et al.*, 1981; Newman *et al.*, 1982b]. This mutation showed a broad effect on cell metabolism: the *ssd* mutant had a lower efficiency of glucose utilization; was unable to grow on succinate or related carbon sources, or to grow anaerobically; was resistant to aminoglycoside antibiotics, and sensitive to fluoride, and deficient in proline and arginine uptake [Morris and Newman, 1980; Newman *et al.*, 1981; Newman *et al.*, 1982].

The *ssd* gene was mapped near *metB*. Several mutants with phenotypes similar to that of the *ssd* mutant were mapped in this area. The genes involved have been named *ecfB* , *eup*, and *cpxA*, depending on their origin in different laboratories. [Plate, 1976; Plate and Suit, 1981; Thorbjarnardottir *et al.*, 1978; Silverman, 1985]. Based on the fact that the *cpxA*

mutation isolated from Silverman's laboratory also showed some of the phenotype associated with the *ssd/ecfB/eup* mutants (aminoglycoside antibiotics resistance and impaired proline uptake), they suggested that the *ssd*, *cpxA*, *ecfB* and *eup* mutations all identify the same gene [Rainwater and Silverman, 1990]. Results presented here show that a plasmid carrying *cpxA* can prevent overproduction of L-SD caused by an *ssd* mutation, supporting the suggestion that the *cpx* gene is the same gene as *ssd*.

Like the *ssd* mutant, the *lrp* mutant described in this work has been isolated by direct selection on L-serine, has considerably elevated levels of the enzyme L-SD, and can grow with L-serine as carbon source. This gene encodes a DNA-binding protein, which uses L-leucine as a co-effector and regulates the expression of a group of genes, among them *sdaA*. The lrp mutation is therefore extremely pleiotropic.

I-4. Environmental Stresses Affect L-SD Activity

Bacteria can respond to environment stress by modulating the expression of specific sets of genes. These are suggested to aid in defense against the stress. In *E. coli*, the synthesis of L-SD is affected by a variety of stresses, though the metabolic use of L-SD against the stress is not understood.

I-4-1. L-SD is Induced by Exposure of *E. coli* to DNA-Damaging Agents

More than 17 gene products are induced upon exposure of *E. coli* cells to agents or conditions that damage DNA or

interfere with DNA replication. These physiological responses have been termed the SOS responses and the genes whose expression are induced by the SOS system have been referred to as *din* (damage-inducible) genes [Kenyon and Walker, 1980; Walker 1987]. The expression of *din* genes is normally under the control of a repressor coded by the *lexA* gene.

Exposure of *E. coli* cells to a variety of DNA-damaging agents, such as UV irradiation, nalidixic acid. and mitomycin C, induced synthesis of L-SD [Newman *et al.*, 1982a], but it is not known whether the SOS is involved in this regulation. The increase in L-SD was shown to require protein synthesis, consistent with a regulation of *sdaA* transcription. A mutant with a hyperactive protease and a nonfunctional *lexA* repressor is constitutive for SOS functions, and also showed an elevated level of L-SD activity [Newman *et al.*, 1982a]. These facts suggested, but did not conclusively demonstrate, that the increase in L-SD in response to DNA-damaging agents might also be mediated via the SOS system [Newman *et al.*, 1982a].

The mechanism by which the *lexA* gene product regulates transcription is well understood. *In vitro* transcription studies showed that purified LexA protein inhibits transcription initiation. Purified LexA protein has been shown to bind to the control sequences of a number of *din* genes. Footprinting experiments revealed that genes repressed by LexA share a consensus sequence for LexA recognition and binding [Little *et al.*, 1981; Sancar *et al.*, 1982; Walker, 1984]. However, there is no sequence homology to the LexA repressor binding site in the region upstream of *sdaA* [Su *et*

al., 1989]. This might suggest that the expression of *sdaA* is indirectly regulated by LexA.

The molecular mechanism of the SOS response has been thoroughly studied. After exposure of *E. coli* cells to a DNA-damaging agent, the inducing signal reversibly activates a protease, the *recA* gene product. The activation of RecA protein results in the irreversible destruction of the LexA repressor by proteolysis. The removal of the LexA protein leads in turn to induction of *din* genes including *recA* gene and *lexA* gene itself [Walker, 1985; Little *et al.*, 1981; Gottesman, 1984].

The RecA protein has also been shown to act as a protease in other systems, cleaving the λ repressor and inducing λ prophage [Roberts and Roberts, 1975]. *In vitro* experiments indicated that the cleavage of the λ repressor proceeded more slowly than cleavage of LexA protein, so that the cellular *din* genes may be induced without induction of resident λ lysogenic phage [Little, 1983; Little and mount, 1982]. The SOS responses on the one hand allow the prophage to enter lytic growth and escape the damaged host. On the other hand, they allow nonlysogenic cells in which lysis is not induced to transcribe a set of genes whose products are necessary for the repair of the DNA damage.

1-4-2. High-Temperature and Alcohol Shock Induction of L-SD Activity

Cells or organisms possess the property of transiently inducing a set of heat shock proteins after a rise in

temperature. This heat-shock response is universal. It has been found in every organism in which it has been sought, from the prokaryotic, eucaryotic, and archaebacterial kingdoms [Lindquist, 1986; Neidhardt *et al.*, 1984]. In *E. coli* K12, 17 proteins have been identified as heat shock proteins by the analysis of two-dimensional gel electrophoresis patterns [Lemaux *et al.*, 1978; Herendeen *et al.*, 1979]. It has been shown that induction of heat shock proteins reflects an increased rate of synthesis of their respective mRNAs [Yamamori and Yura, 1980, 1982; Yamamori *et al.*, 1982] and increased transcription initiation at heat shock promoters [Taylor *et al.*, 1984]. The heat shock promoters are under the control of a heat shock regulatory protein encoded by the *htpR* gene [Neidhardt *et al.*, 1983]. The HtpR protein is an alternative σ factor for RNA polymerase, σ32, which recognizes a consensus sequence in the heat shock promoters and stimulates their transcription [Grossman *et al.*,1984].

L-SD activity has been shown to be markedly affected by growth temperature in *E. coli*. However the mechanism responsible for this is not yet clear. L-SD is induced by an increase in growth temperature. This induction requires protein synthesis, since the increase in L-SD activity was inhibited by incubation with chloramphenicol. A shift from a low temperature to a high temperature was followed by a rapid increase in L-SD activity, but the reverse shift from high temperature to low temperature only changed L-SD activity slightly. These results indicated that an increase in

temperature induces the synthesis of L-SD but does not alter enzyme activity [Newman *et al.*, 1982c].

It has been reported that the thermal induction of L-SD activity is under the control of the *htpR* gene product. This is shown by the abolishment of heat induction of L-SD in an *htpR* mutant. This is also suggested by the fact that both the heat shock regulon and sdaA are induced at normal temperatures by ethanol shock.

On the other hand the fact that the sequence upstream of *sdaA* lacks the heat shock promoter consensus sequence [Su *et al.*, 1989] makes it less likely that the *sdaA* gene is part of the *htpR* regulon. On the other hand, not all heat shock genes contain this consensus sequence. One which does not is *lysU*. Though two-dimensional gel patterns showed that the *lysU* gene product, lysyl-tRNA synthetase form II, is a heat shock protein [Hirshfield, *et al.*, 1981; Neidhardt and VanBogelen, 1981], the putative promoter sequence of *lysU* gene is similar to the consensus sequence of strong σ^{70} promoters and is not like that of heat shock promoters [Clark and Neidhardt, 1990]. This could be explained by the possibility that there is in fact a heat shock promoter in *lysU* but it was not seen because it lies outside the region of upstream DNA that was sequenced. Alternatively, it was suggested that the heat induction is an indirect effect of *htpR* function rather than a direct one for both *sdaA* and *lysU* [Su *et al.*, 1989; Clark and Neidhardt, 1990]. Several heat shock proteins are produced even in strains carrying loss-of-function insertion and deletion

mutations in the *htpR* gene, which supports the idea of the existence of additional mechanisms controlling the expression of some heat shock proteins [Zhou *et al.*, 1988].

I-4-3 Summary of Factors Affecting L-SD Activity

Apart from DNA damage, elevated temperature, and alcohol shock, anaerobic growth also stimulates the synthesis of L-SD. Why L-SD activity is affected by so many environmental factors is not known. Indeed the metabolic role of L-SD in *E. coli* growing in glucose minimal medium is as yet unknown. However, the fact that alterations in L-SD activity are incorporated into so many complex regulatory mechanisms suggest that it must be of considerable metabolic consequence, even if we cannot explain why this is so.

L-SD activity is also affected by several regulatory genes. Among these genes are two global regulators, *ssd* and *lrp*. Mutations in either the *ssd* or the *lrp* genes greatly increase L-SD activity. Since most of the work in this thesis is devoted to the characterization of the nature and mechanism of the *lrp* regulon, some of the already known global regulatory systems and their molecular mechanisms will now be reviewed.

II. A SURVEY OF GLOBAL REGULATORY SYSTEMS.

The accurate and efficient regulation of gene expression is an important aspect of biological phenomena such as cell growth, and the response to environmental conditions. *E. coli*

and other enteric bacteria can modulate the expression of their genes, resulting in rapid growth under favourable conditions and fostering survival under unfavourable conditions.

Regulation of transcription of single genes, or of a group of contiguous genes, by a transcription factor has been intensively studied, and the lactose operon particularly so. These 3 contiguous genes are regulated by a transcriptional regulator, the *lacI* gene product, such that the genes are transcribed only in the presence of lactose or other inducer (and the absence of glucose). By this system, ß-galactosidase can be synthesized in a vastly greater amount in the presence of lactose than in its absence. Indeed, there are about 3000 molecules of ß-galactosidase in each wild-type *E. coli* cell grown with lactose as carbon source, and less than one-thousandth of this number in cells growing in the presence of glucose, i.e. in lactose non-utilizing conditions.

As illustrated by the *lac* system, and many others, the cell has elaborated a sophisticated mechanism for the regulation of expression of genes involved in a single metabolic pathway. In recent years, it has been discovered that the cell also has developed controls which coordinate the expression of genes involved in several pathways, by way of a single regulator responding to internal or external stimuli. This coordinated expression of genes of several pathways is referred to as global regulation.

The term "regulon" has been coined to describe one or a network of individual transcription units (operons) under the

control of a common regulatory protein (global regulator). Usually the regulatory region of the DNA of each operon in a regulon has a particular sequence common to all members of the regulon, and that sequence is the recognition site for the regulator [Neidhardt, 1987; Maas and Clark, 1964].

Gottesman [Gottesman, 1984] characterizes the "global regulon" by several properties: it must have more than one subject operon; the genes in these operons should code for products involved in more than one metabolic pathway, and the expression of the genes should be regulated by a common regulator, rather than by a common stimulus. More than 20 global regulatory systems have been recognized in enteric bacteria [Neidhardt, 1987; Iuchi and Lin, 1988; Iuchi *et al.*, 1989].

Investigators have been particularly interested in the mechanism by which a set of genes may be selectively expressed in response to an external or internal stimulus. What signals transmit information from the exterior of the cell to the promoter of the gene? This involves sensors interacting with the environment, and regulatory proteins which can alter gene expression by binding to specific sites on DNA, and some communication between the two.

II-1. Global Regulatory Systems Controlled by Activators of Gene Transcription.

Regulatory proteins may either prevent or increase transcription by RNA polymerase, and often can exist in either

an active or an inactive form, depending on environmental conditions. Repressors may bind to the operator region of a gene and prevent the initiation of transcription. On the other hand, an activator binds to other DNA control sites and stimulates transcription initiation.

This interaction, between a positive regulatory protein and the DNA site to which it binds, resulting in activation or enhancement of transcription initiation by RNA polymerase, is termed positive control. The regulatory protein, the activator, is either made only when it is required, or made in an inactive form and activated when it is required. It may function as a new σ factor for the RNA polymerase holoenzyme, an accessory factor that allows RNA polymerase to initiate transcription at specific promoters, or a new RNA polymerase.

II-1-1 Regulation of Transcription by Alternative Sigma Factors

RNA polymerase purifies *in vitro* as a five subunit enzyme (ß'ßα2ω). However it functions *in vivo* as an assembly of 6 subunits, the sixth, sigma, being readily lost during purification. The first RNA polymerase preparations were complexed with a sigma known as σ^{70}. It was suggested at that time that different sigmas might stimulate transcription from different sets of genes of different classes of RNA molecules, which would constitute a mechanism for positive control [Burgess and Travers, 1970].

Specificity conferred by sigma factors has since been clearly demonstrated. Four bacterial alternate sigmas and two

phage-encoded sigmas have been found in *Bacillus subtilis*. These sigmas function as positive control factors in bacterial or phage developmental regulation [Losick and Pero, 1981, Johnson *et al.*, 1983]. A sigma factor from *E. coli* phage T4, has been shown to be involved in the transcription of phage late genes [Kassavetis and Geiduschek, 1984; Malik *et al.*, 1985].

Three alternative σ factors which govern specificity in *E. coli* global regulatory systems, have been intensively studied. The role of these factors, the products of the *htpR*, *rpoS* (*katF*), and *ntrA* (*glnF*) genes, is reviewed in the next sections [Grossman *et al.*, 1984; Hirschman *et al.*, 1985; Mulvey and Loewen, 1989; Lange and Hengge-Aronis, 1991].

II-1-1a. HEAT SHOCK/sigma 32

Seventeen heat shock proteins are coded by genes whose induction by heat is under the transcriptional control of the HtpR protein, σ^{32}. It has been suggested that heat shock response is a result of both activation of this sigma from a preexisting inactive form, and increased synthesis of σ^{32} from newly transcribed mRNA [Grossman *et al.*, 1984]. The heat shock proteins also can be induced by UV irradiation and nalidixic acid [Krueger and Walker, 1984], ethanol [Neidhardt *et al.*, 1984], infection with λ phage [Drahos and Hendrix, 1982; Kochan and Murialdo, 1982], inactivation of 4.5S RNA [Bourgaize *et al.*, 1990], and carbon starvation [Jenkins *et al.*, 1991]. Induction by these stimuli is also mediated by

increased σ^{32} levels [Bahl *et al.*, 1987; Bourgaize *et al.*, 1990; Gross *et al.*, 1990; Jenkins *et al.*, 1991].

II-1-1b. STATIONARY PHASE GENE EXPRESSION/sigma KatF.

In *E. coli*, the physiology of stationary-phase cells is different from that of rapidly growing cells. Stationary phase more closely resembles the conditions that *E. coli* encounters in its mostly nutrient-limited natural environments. Stationary phase cells are smaller, more spherical, and more resistant to high temperature, high concentrations of H_2O_2 and high osmolarity than exponentially growing ones [Jenkins *et al.*, 1988; 1990; Lange and Hengge-Aronis, 1991].

Studies in several laboratories have shown that transcription of a number of genes increases during stationary phase, suggesting the possibility that they may be co-regulated by a global regulator. The genes which are transcribed to an increased extent in stationary phase include the genes responsible for H_2O_2 resistance, the gene for resistance to near-ultraviolet radiation damage, *xthA*; the genes involved in cell division, *ftsQ*, *ftsA*, and *ftsZ*; the morphogene *bolA*; the gene encoding acid phosphatase, *appA*; the microcin operon, *mcbABCDEFG* and *mcc*; and some *glg* genes involved in glycogen production [Jenkins *et al.*, 1988; 1990; Lange and Hengge-Aronis, 1991; Mulvey and Loewen, 1989].

The gene product of *katF*, seems to be the global regulator of stationary-phase gene expression. From the nucleotide sequence of *katF*, Mulvey and Loewen [1989] suggested that KatF

protein is a novel σ transcription factor, σ^S. It has been shown to be required for the expression of several of the stationary phase genes: *katE* and *xthA* [Sak *et al.*, 1989], and also for $bolAp_1$ transcription [Lange and Hengge-Aronis, 1991]. This putative σ factor is also required for the expression of proteins involved in thermotolerance, production of glycogen, long term starvation survival, the expression of acidic phosphatase (*appA*), and synthesis of a number of proteins identified only by two-dimensional gel electrophoresis [Lange and Hengge-Aronis, 1991a].

The expression of *katF* itself is induced during transition into stationary phase. Several criteria, in addition to sequence comparison mentioned above, strengthen the proposal that KatF protein is a sigma factor. Based on the comparison of the *xthA*, *katE*, and $bolAp_1$ promoters, potential -35 and -10 consensus sequences have been suggested for σ^S recognition. However, there is still no direct biochemical evidence showing that KatF is indeed a σ factor [von Ossowski *et al.*, 1991; Lange and Hengge-Aronis, 1991].

II-1-1c. USE OF ALTERNATIVE NITROGEN SOURCES/sigma 54.

E. coli and related bacteria have a global regulatory system termed Ntr for regulation in response to nitrogen limitation. When ammonia or other inorganic nitrogen sources are limited in the growth medium, enzymes necessary for the utilization of various alternative nitrogen sources are induced [Magasanik and Neidhardt, 1987]. The regulation of the

activity and level of glutamine synthetase by the Ntr system has been extensively studied, in *E. coli*. However the system also regulates expression of many other genes including those that encode proteins that are involved in the degradation of arginine, ornithine, agmatine, putrescine, and γ-aminobutyrate, and in the transport of glutamine and γ-aminobutyrate [reviewed in Reitzer and Magasanik, 1987]. In *Klebsiella pneumoniae*, the operons of Nif system, responsible for dinitrogen fixation, are also under Ntr control [Merrick, 1983].

There are five regulatory genes involved in the Ntr regulon, *glnG*, *glnL*, *glnB*, *glnD*, *glnF*, the last of which encodes a new sigma factor. The first of these, *glnG* (*ntrC*) codes for nitrogen regulator I (NR_I). The active form of nitrogen regulator I (NR_I-P) is required for activation of transcription of the genes of this regulon. The inactive form of NR_I, is phosphorylated by the NR_{II} kinase, which converts it to the active phosphorylated form, NR_I-P. NR_{II}, the product of *glnL* (*ntrB*) gene, is an NR_I-kinase/NR_I-P phosphatase.

What then controls whether NR_I is active? This depends on the state of P_{II}, the product of the *glnB* gene. When the glutamine/2-ketoglutarate ratio is low, as might be expected when nitrogen is limiting, P_{II} is converted to P_{II}-UMP by a uridylyltransferase, the product of *glnD* gene. Free NR_{II} protein then acts as a kinase and activates NR_I.

This is reversed when nitrogen is available and the glutamine/2-ketoglutarate ratio is high. Then, the *glnD* gene

product acts as a uridylyl-removing enzyme, removing the uridylyl group from P_{II}. The interaction of the deuridylylated form of P_{II} with NR_{II}, converts the latter into a phosphatase which dephosphorylates NR_I, and inactivates transcription initiation at_Ntr promoters [Magasanik and Neidhardt, 1987; Stock *et al.*, 1989]. NR_I also acts as a repressor; transcription of *glnA* from one of the two *glnA* promoters, $glnAp_1$, is repressed by NR_I, whereas transcription from $glnAp_2$ is activated by NR_I [Reitzer and Magasanik, 1985].

In addition to the complex system summarized above, another alternative sigma factor, coded for by the fifth regulatory gene *ntrA* (*glnF*), is also involved in this regulation. The active form of NR_I works in conjunction with σ^{54} to initiate transcription at Ntr promoters recognised by RNA polymerase associated with the alternative σ^{54}, [Hunt and Magasanik, 1985; Hirschman *et al.*, 1985].

Some details of the molecular mechanisms involved have been elucidated. The σ^{54} RNA polymerase recognizes the nitrogen-regulation/nitrogen fixation promoter consensus sequence CTGGYAYR-N_4-TTGCA. It was shown that σ^{54} RNA polymerase interacts with the $glnAp_2$ promoter independently of NR_I or NR_I-p and forms a closed RNA polymerase-promoter complex. Then NR_I-p catalyses isomerization of the closed form complex between σ^{54}-RNA polymerase ($E\sigma^{54}$) and the promoter to form transcriptionally active open complexes in an ATP-dependent reaction [Popham *et al.*, 1989; Ninfa *et al.*, 1987].

In *K. pneumoniae*, a nitrogen-fixing organism, NR_I-P activates transcription of the *nifLA* operon as well as of Ntr

promoters. In the absence of NR_I-p, a closed $E\sigma^{54}$-promoter complex was formed at $glnAP_2$ promoter, but close contacts between $E\sigma^{54}$ and -24, -12 region (promoter) of *nifL* promoter were not observed. These results suggest that NR_I-p may play an important role in stabilising the closed complex at the *nifL* promoter as well as its role in catalysing open complex formation. It has been suggested that σ^{54}-dependent promoters are important in activation by NR_I and that the sequence of the σ^{54} RNA polymerase binding site is an important determinant of the efficiency of activation [Ray *et al.*, 1990].

II-1-2 Positive Global Regulation by Accessory Transcription Factors.

Transcription is positively regulated by sigma factors, as reviewed above, and also by a large variety of proteins, operating by a considerable variety of mechanisms. Among the global regulatory systems of this type are the two-component regulatory systems. In these systems, one component (the sensor component) of each system acts as an environmental sensor that responds to an environmental signal and transmits the signal to a second regulatory component (activator), which in turn regulates the transcription of certain genes [Ronson *et al.*, 1987].

In *E. coli*, two-component systems are involved in nitrogen regulation (*ntrB/ntrC*), oxygen regulation (*cpxA/arcA*, *narX/narL*), response to osmolarity (*envZ/ompR*), phosphate starvation (*phoR/phoB*), and chemotaxis (*cheA/cheB*). In these systems all members of the sensor class regulate their

partners by a covalent modification (e.g. phosphorylation) [Stock *et al.*, 1989]. Some of these systems will be reviewed in the next sections.

II-1-2a. REGULATION OF THE USE OF ALTERNATIVE PHOSPHATE SOURCES

When *E. coli* is starved for inorganic phosphate (p_i), the expression of at least 18 genes is induced [Wanner and Mcsharry, 1982]. This involves at least 3 regulatory genes (*phoB*, *phoR*, *phoM*) [Guan *et al.*, 1983; Shingawa *et al.*, 1983; Tommassen *et al.*, 1982]. They regulate expression of genes involved in phosphate uptake and metabolism. These include genes which code for a high-affinity phosphate-specific transport (pst) system [Rosenberg *et al.*, 1977; Willsky and Malamy, 1980] which is involved in the scavenging of low amounts of inorganic phosphate, and for several sugar phosphate transport systems such as Gpt, Pgt, and Uhp, as well as for PhoA, PhoE, which are involved in the assimilation of phosphorus-containing organic compounds [Wanner, 1983; Stock *et al.*, 1989].

The mechanism for regulation of the p_i (Pho regulon) is very complex. The regulator PhoB is essential for the transcriptional activation of the Pho regulon, including *phoA*, *phoE*, the *pst* operon, and the *phoBR* operon itself [Stock *et al.*, 1989]. In its active, phosphorylated form, the *phoB* gene product, P-PhoB, activates the transcription of all operons in the Pho regulon.

Whether PhoB is phosphorylated or not depends on the activity of both the *phoR* and *phoM* gene products. These are both histidine protein kinases. PhoR is the major histidine kinase of the Pho regulon [Stock *et al.*, 1989]. but PhoM may replace PhoR to activate (phosphorylate) PhoB [Stock *et al.*, 1989].

The activity of PhoR depends on the environment of the cell. It may be regulated either directly by phosphate in the periplasm or indirectly by an interaction of the receptor with periplasmic phosphate-binding protein. *In vitro*, PhoR can undergo autophosphorylation when incubated with ATP. The autophosphorylated PhoR can transfer phosphate to PhoB. This suggests that when PhoR is phosphorylated *in vivo*, it can transfer phosphate to PhoB, thus activating it.

II-1-2b THE CATABOLITE ACTIVATOR REGULON

In the cAMP-CAP system, adenylate cyclase acts as a sensor protein, but it does not modify CAP covalently. It senses the available carbon source and makes cAMP according to the information it sees, and the cAMP noncovalently binds to CAP and alters its function.

The preferred carbon and energy source of *E. coli* and related bacteria is glucose. When glucose is available in its surroundings, the synthesis of many of the enzymes responsible for the utilization of other carbon and energy sources is decreased (the glucose effect). Several components are involved in this inhibitory effect, which is also known as

catabolite repression [Magasanik, 1961; 1970; Ullmann and Danchin, 1983].

Catabolite repression is under the control of a global regulatory system, the cAMP-CAP regulon, acting at a wide variety of promoters, in combination with individual controls specific for each catabolic system. Both CAP, the *crp* gene product, and cAMP, produced by adenylylcyclase, the product of *cya* gene, are essential for the expression of catabolite-sensitive operons.

Binding of cAMP to CAP is a reversible reaction [Botsford, 1981]. When a carbon source such as glucose is present, the intracellular cAMP concentration is very low, and the cAMP/CRP complex does not form to any significant extent. However, when glucose is absent, the intracellular cAMP levels are increased, resulting in the formation of the cAMP-CAP complex [Ullmann and Danchin,1983]. Upon binding cAMP, the CAP dimer undergoes a comformational change and binds to DNA target sites of numerous regulated promoters, activating transcription initiation from some promoters such as $lacP_1$, $galP_1$, and *araBAD*, and inhibiting transcription from other promoters such as $galP_2$ and $lacP_2/P_3$ [Musso *et al.*, 1977; Xiong *et al.*, 1991]. The mechanism by which cAMP-CAP regulates transcription initiation has been extensively studied (see next section).

II-1-2c. REGULATION BY OxyR: Helix-turn-helix motifs.

Many transcriptional regulators in bacteria contain a "helix-turn-helix" domain by which the proteins recognize and

bind to specific DNA sequences [Pabo and Sauer, 1984; Gehring, 1987]. Among these are the LysR family of transcription activators, whose DNA sequences as translated into protein, appear to contain a "helix-turn helix" DNA binding motif near their N-termini [Henikoff *et al.*, 1988; Christman *et al.*, 1989].

One member of this family, the *oxyR* gene product, OxyR, is responsible for a global response to oxidative stress. When *E. coli* and *S. typhimurium* cells are treated with low doses of H_2O_2, at least 30 proteins are induced in response to this oxidative stress and the cells become resistant to subsequent lethal doses of hydrogen peroxide and other peroxidizing agents [Christman *et al.*, 1985; Demple and Halbrook, 1983; Morgan *et al.*, 1986]. The *oxyR* gene regulates transcription of one of the (at least) two regulons involved in this response to oxidative stress, controlling expression of at least nine of the hydrogen peroxide-inducible genes [Christman *et al.*, 1985].

The OxyR protein is homologous to the LysR family of bacterial regulatory proteins. All regulators in this family are between 30 and 35 kDa in size and all act as positive regulators.

The oxyR protein senses the oxidizing conditions and regulates gene expression accordingly. *In vitro* studies suggest that direct oxidation of the OxyR protein results in a conformational change by which OxyR transduces the oxidative stress signal to RNA polymerase [Storz *et al.*, 1990]. Like other members of this family (LysR, NodD, IlvY and MetR), OxyR

also negatively regulates its own synthesis from the gene that codes for it [Henikoff *et al.*, 1988; Christman *et al.*, 1989; Wek and Hatfield, 1988].

Studies of two OxyR regulated genes, *katG* and *ahpC*, show that OxyR activates at the level of transcription [Morgan *et al.*, 1986; Tartaglia *et al.*, 1989]. This might imply that the promoters of these genes have a common sequence recognized by OxyR, and in fact DNaseI protection studies of the *katG*, *ahpC* and *oxyR* promoters showed OxyR-dependent footprinting in the upstream regions of the genes, but there were no obvious sequence homologies in the protected sites. To account for this lack of homology, it has been suggested that DNA secondary structure may be important in the binding of OxyR [Tartaglia *et al.*, 1989]. Unlike the case for other regulatory proteins, both the active (oxidized) form and inactive (reduced) form of OxyR protein can bind to these three diverse sequences upstream of OxyR-regulated promoters. The lengths of DNA sequence protected by OxyR under oxidizing conditions are different from those protected under reducing conditions suggesting a distinct conformational change of OxyR between the two states. The *in vitro* studies shown that oxidized but not reduced OxyR activates transcription. The results suggest that OxyR protein undergoes a conformational change in response to the oxidative stress signal, and that while both forms bind, only the oxidized form acts as an activator [Storz *et al.*, 1990].

Another regulon involved in oxygen stress, controlled by *soxR*, governs at least part of the global response to

superoxide in *E. coli* [Tsaneva and Weiss, 1990]. When *E. coli* is exposed to superoxide anion-generating compounds, the expression of about 40 proteins is increased [Greenberg and Demple, 1989; Walkup and Kogoma, 1989]. Nine of these proteins are under the control of SoxR, the regulatory protein of the superoxide response regulon [Greenberg *et al.*, 1990; Tsaneva and Weiss, 1990]. This works together with another gene product, SoxS, to activate this regulon [Wu and Weiss, 1991]. Their sequences suggest that both SoxR and SoxS contain helix-turn-helix DNA binding domains. The C-terminal region of SoxS protein shows homology with members of the AraC family of positive regulators, suggesting that SoxS may be a regulator. The SoxR, which contains four closely spaced cysteines, might then act as a sensor [Wu and Weiss, 1991].

II-2. Negative Control of Transcription.

Repressors that selectively bind to operators and block the initiation of transcription are a widely used component in the regulation of gene expression in all cells [Brennan and Matthews, 1989]. Of the 107 sigma 70 promoters of *E. coli* and *S. typhimurium* that have been analyzed, 76 are repressible promoters [Collado-Vides *et al.*, 1991]. Repressors may control single operons, or global regulons. Indeed, some global regulators act as both repressor and activator.

II-2-a. THE SOS REGULON.

As mentioned earlier, LexA protein is a repressor of the SOS system. *In vitro* studies show it to be a repressor of

transcription initiation of the SOS genes [Brent and Ptashne, 1981; Little *et al.*, 1981; Mizusawa *et al.*, 1983; Sancar, *et al.*, 1982]. Operon fusion and mRNA studies show that DNA damaging agents induce the transcription of SOS genes by interfering with LexA repression [Kenyon and Walker,1980, Kenyon *et al.*, 1982].

II-2-b. REGULONS INVOLVED IN THE ADAPTATION OF *E. coli* TO DIFFERENT RESPIRATORY CONDITIONS: Fnr, Nar and Cpx.

E. coli and other enterobacteria can use a variety of electron acceptors such as oxygen, nitrate, fumarate, dimethyl sulfoxide,and trimethylamine-N-oxide for respiration [Ingledew and Poole, 1984; Iuchi and Lin, 1987]. In the absence of available exogenous electron acceptors, these bacteria can switch from respiration to fermentation, which yields less energy since in this pathway the energy is derived only from substrate-level phosphorylation. The efficient use of available compounds involves both global and specific regulatory systems which regulate the use of the alternate respiratory pathways [Iuchi and Lin, 1991]. Three global regulatory systems have been shown to be involved, and some operons are subject to overlapping control [Iuchi and Lin, 1991].

The regulatory protein Fnr (fumarate and nitrate reductases) is necessary for the control of the expression of genes that encode not only anaerobic respiratory proteins functioning as terminal reductases and primary dehydrogenases

but also aerobic respiratory proteins functioning as terminal oxidases.

Fnr acts both positively and negatively. It acts as a transcriptional activator required for the anaerobic expression of *frdABCD*, *fdnGHI*, *narGHJI*, and *dmsABC*, the genes encoding fumarate reductase, formate dehydrogenase-N nitrate reductase, and DMSO reductase, respectively [Chippaux, *et al.*, 1981; Lambden and Guest, 1976; Stewart,1982; Ingledew and Poole, 1984]. On the other hand, Fnr acts as a negative autoregulator [Spiro and Guest, 1987] and also represses the expression of *ndh*, *cyoABCDE* and *cydAB*, the genes encoding NADH dehydrogenase II, cytochrome o, and cytochrome d, respectively [Spiro *et al.*, 1989; Cotter *et al.*, 1990]. Under microaerobic conditions, Fnr is required for the peak expression of the *cyd* operon [Fu *et al.*, 1991]. Some conflict in results from different laboratories has been ascribed to the different growth conditions used [Fu *et al.*, 1991]. The amino acid sequence of Fnr shows a high degree of homology to Cap. The target promoters regulated by Fnr contain a 5bp inverted repeat: TTGAT----ATCAA [Spiro and guest, 1990].

The presence of nitrate can repress the transcription of genes involved in anaerobic respiration and fermentation, a response mediated by the two-component regulatory system, NarX/NarL, [Cotter and Gunsalus, 1989; Iuchi and Lin,1987; Kalman and Gunsalus, 1988; 1989; Sawers and Bock, 1988; Stewart and Berg, 1988; Stewart and Parales, 1988]. This response represses transcription of genes such as *frdABCD*, and

dmsABC; and induces transcription of genes (*narGHJI*, *fdnGHI*) involved in the anaerobic respiratory chain, formate-nitrate oxidoreductase, and the *pfl* gene encoding pyruvate formate-lyase.

Nitrate regulation requires molybdate [Iuchi and Lin, 1987b]. NarX and NarL have been suggested to comprise a nitrate-responsive sensor-regulator pair. NarX acts as a histidine protein kinase, which senses the availability of nitrate and molybdenum, and activates NarL. The activated form of NarL can bind to nitrate-controlled promoters and regulate transcription from these promoters [Kalman and Gunsalus, 1990]. It was found that the effects of *narX* insertion mutations on the nitrate response are due to decreased *narL* expression, as a result of polarity which led to the suggestion that NarX is not essential for normal nitrate regulation [Egan and Stewart, 1990]. Although all nitrate regulated genes are controlled by Fnr, the promoters controlled by both Fnr and NarL have different recognition sites for both regulators. Several NarL-regulated promoters contain two conserved sequences, CTCCTT and TACCCAT [Iuchi and Lin, 1991].

II-2-c. REGULATION OF OXIDATIVE ENZYMES: ArcA and ArcB.

In *E. coli*, the global regulatory genes *arcA* and *arcB* mediate the repression of genes encoding for TCA cycle enzymes, glyoxylate shunt enzymes, primary dehydrogenases, pyruvate dehydrogenase, cytochrome o, enzymes of ß-oxidation of fatty acids and the activation of the cytochrome d operon

[Iuchi and Lin, 1988; 1991; Iuchi *et al.*, 1990]. Based on their amino acid sequences, it has been suggested that ArcA acts as a regulator of a two-component regulatory system and ArcB as a histidine kinase [Ronson *et al.*, 1987; Iuchi and Lin, 1991]. ArcA also is required for the expression of the F plasmid DNA transfer (*tra*) genes. In different laboratories this gene has been called *dye/msp/fex/sfrA/cpxC* respectively [Beutin and Achtman, 1979; Buxton and Drury, 1983; Lerner and Zinder, 1979; Silverman *et al.*, 1980]. In the case of the regulation of F plasmid DNA transfer genes, ArcA acts as a transcriptional activator, and another transmembrane protein, CpxA, acts as a sensor protein [Ronson *et al.*, 1987; Iuchi and Lin, 1991]. It was shown that the Arc (aerobic respiration control) and Sfr (sex factor regulation) functions of the *arcA* gene product are separately expressed and regulated [Silverman *et al.*, 1991].

III. A SURVEY OF SOME POSSIBLE MECHANISMS OF REGULATION OF TRANSCRIPTION INITIATION

The nature of the DNA sequence of a given promoter largely determines the frequency of transcription initiation at that promoter. There seem to be two kinds of promoters in *E. coli*: one type acts constitutively unless repressed and the other works poorly unless activated [Ptashne, 1991].

In the classic model of transcription initiation [Chamberlin, 1974], the formation of an open complex between DNA and RNA polymerase is considered to involve a two-step

interaction of RNA polymerase with the promoter. In the first step RNA polymerase binds to the promoter and forms a competitor-sensitive closed complex. This step is described by an equilibrium constant, K_b. The DNA near the transcription start site then unwinds and forms the tighter-binding open complex [Kirkegaard *et al.*, 1983]. The second step, with a rate constant K_f, is a slow step which has been termed the isomerization step. Promoter mutations which increase or decrease promoter function result in increasing or decreasing either K_b, K_f, or both, respectively [McClure, 1985]. The regulatory proteins, activators or repressors, can also affect the formation of the open complex by affecting K_b or K_f (Hoopes and McClure, 1987].

Recent studies suggest that repression and activation in *E. coli* may function by common mechanisms. The regulatory proteins may act by directly facilitating or blocking RNA polymerase binding, or by affecting promoter structure resulting in an increased or decreased affinity for RNA polymerase.

A comparison of DNA sequences of *E. coli* $E\sigma^{70}$ promoters identifies two hexanucleotide consensus sequences, TATAAT and TTGACA, centered approximately 10 and 35 bp, respectively, upstream from the transcription startpoint [Hawley and McClure, 1983; Raibaud and Schwartz, 1984].

Any activator protein can act as a repressor if it binds in a position that interferes with RNA polymerase-promoter interactions [Hoops and McClure, 1987]. Many bacterial

regulatory proteins can function either as activators or repressors depending on the location of the sites to which they bind and the nature of the promoter. Regulatory proteins such as Cap and Fnr activate most of the promoters they regulate by binding to a position near -40 and repress some of the promoters when their sites are downstream of -20 [Collado-Vides *et al.*, 1991]. On the other hand, OmpR regulatory protein can repress and activate from the same position [Tsung *et al.*, 1990]. Perhaps the same OmpR-polymerase interaction which can help stabilize binding to weak basal elements can also inhibit the binding when strong basal elements have already allowed stable polymerase binding [Gralla, 1991].

III-1. The Mechanisms of Negative Regulation

The function of a repressor is to decrease transcription initiation. Different repression mechanisms are found in bacteria. In most cases the operators (the binding sites of repressor) overlap with the RNA polymerase binding site. The repressor functions either by competing with RNA polymerase for a specific DNA sequence, thus blocking RNA polymerase binding, or by interfering with the function of bound RNA polymerase, as by blocking isomerization or inhibiting the actual initiation steps. [Collado-Vides, *et al.*, 1991]. In some cases, repression also involves DNA looping [Adhya, 1989], or depends on interactions with other DNA-binding proteins [Gerlach, *et al.*, 1991].

III-1-1. Repressor Directly Competes or Interacts with RNA Polymerase

To initiate transcription, RNA polymerase has to bind to a promoter to form a closed complex. If the repressor binding site overlaps with the site of the closed complex formation, the binding of a repressor will prevent the binding of RNA polymerase. Among 76 repressible promoters that have been analyzed, 70 contain at least one operator in a proximal position, where the repressor can contact either the polymerase or the polymerase recognition elements. The position of the operator can influence repressibility by determining when it is cleared and available for repressor binding and also by determining how effectively the bound repressor can block the polymerase binding [Collado-Vides, *et al.*,1991]. Studies have shown that the cI repressor of bacteriophage λ decreased the binding constant (K_b) of RNA polymerase for the P_R promoter [Hawley *et al.*, 1985].

During the formation of the open complex, the DNA helix is opened and the conformation of RNA polymerase is changed. DNaseI footprinting shows that there is also a change in the size of the DNA sequence which contacts RNA polymerase [Mecsas *et al.*, 1991]. Some repressors can block open complex formation either by binding to the DNA sequence necessary for the formation of a stable complex with RNA polymerase or by changing the DNA conformation in such a way that RNA polymerase is unable to open the helix for isomerization [Adhya, 1989]. The Arc repressor of bacteriophage P22 of

Salmonella typhimurium can bind between the -10 and -35 region of the P_{ant} promoter and repress through an effect on K_f [Vershon *et al.*, 1987].

III-1-2. Repression by DNA Looping

In both prokaryotic and eukaryotic systems, the direct interaction between proteins bound to separated sites on the DNA to form DNA loops is one of the important mechanisms for gene regulation [Adhya, 1989; Gralla, 1989]. Loop formation is also important for other biological systems such as initiation of DNA replication and site-specific recombination [Echol, 1986; Moltoso de Vargas *et al.*, 1989]. A protein-mediated DNA looped complex may serve to increase the relative concentration of regulatory proteins at the control site [Mossing and Record, 1986], providing a way for multiple proteins to affect transcription from a promoter [Schleif, 1987], or bring regulatory protein into contact with RNA polymerase [Wedel *et al.*, 1990; Ptashne, 1986], and facilitate transfer of a protein from one region of DNA to another [von Hippel and Berg, 1989]. In some *E. coli* operons, distant duplicated operator sites are involved in repression [Collado-Vides *et al.*, 1991].

In the *lac* operon, there are three locations for Lac repressor binding: a strong repressor binding site O_1 (primary operator) and two weak binding sites (pseudo-operators), O_2 (402 bp downstream from O_1, within the *lacZ* coding region) and O_3 (93 bp upstream from O_1) [Adhya, 1989; Gralla, 1989]. *In vivo* footprinting and expression studies show that O_1 and O_2

cooperate to form a repression complex *in vivo*, and repress transcription by both strengthening repressor binding at the initiation site within O_1 and blocking elongation [Flashner and Gralla, 1988; Sasse-Dwight and Gralla, 1988]. O_1 and O_2 may form a looped complex which can inhibit both the initiation and elongation of transcription, though the primary effect is at the initiation step [Gralla, 1989]. The ability of the Lac repressor to mediate formation of DNA loops was shown both *in vivo* and *in vitro* [reviewed in Brenowitz *et al.*, 1991]. *In vivo* studies show that full repression of the *lac* operon requires all three binding sites. The destruction of just one pseudo-operator, O_2 or O_3, reduces repression by wild-type tetrameric Lac repressor 2- to 3-fold; inactivation of both pseudo-operators decreases repression more than 50-fold. Binding to O_1 alone represses only 20-fold. A dimeric active ac repressor represses the wild-type *lac* operon to about the same level as wild-type binding to O_1 alone. These observations indicate that cooperative interactions between three operator sites is due to DNA loop formation mediated by a tetrameric Lac repressor. [Oehler *et al.*, 1990]. Studies with the *lac* repressor-operator system indicate that DNA looping can affect local DNA supercoiling during transcription [Wu and Liu, 1991]. Whether altered supercoiling affects the physiological functions of DNA remains unknown.

The two divergent *ara* promoters, P_{BAD} and P_C, are regulated both positively and negatively by AraC. In the absence of arabinose, both promoters are repressed by AraC. When AraC binds to arabinose, it activates P_{BAD} but still represses P_C.

AraC can bind to three operator sites upstream from P_{BAD}: I and O_1 are located between P_{BAD} and P_C, O_2 is within the coding region of *araC*. The *In vivo* DNA protection experiments show cooperative binding between I-O_2 and O_1-O_2. Unliganded AraC cooperatively binds to I and O_2 and does not bind to O_1, while liganded AraC cooperatively binds to O_1 and O_2, and also binds independently to I [Martin *et al.*, 1986]. These observations suggest that AraC mediated DNA-loop formation is involved in gene regulation. Binding of unliganded AraC to I and O_2 forms a DNA loop between two sites separated by 210 bp and represses both P_{BAD} and P_C. Binding of liganded AraC to I will activate the transcription from P_{BAD}, and the DNA looped complex between liganded AraC bound at O_1 and O_2 still represses P_C [Lobell and Schleif, 1991]. Both P_{BAD} and P_C are activated by CAP. In vitro studies have shown that CAP activates by breaking the O_2-I repression looped complex. The CAP activity is dependent on the orientation and distance of the CAP binding site relative to I [Lobell and Schleif, 1991].

Two *gal* promoters are subjected to the dual control of CAP and Gal repressor, CAP activates transcription initiation from P_1 but inhibits initiation from P_2. Gal repressor represses transcription at both promoters [Musso *et al.*, 1977]. The two operator sites, O_E and O_I, do not overlap the promoters. The *In vivo* transcription studies show that the repression requires cooperative binding of the repressor to both sites, Binding to either O_E or O_I is not enough for repression [Mandal *et al.*, 1990]. Electron microscopy observations indicate that repressor binding to both sites can form a DNA loop [Adhya,

1989]. It has been suggested that a protein-mediated looped complex between O_E and O_1 is required for the repression of transcription from the *gal* promoter [Adhya, 1989].

In *E. coli*, the *deo* operon encoding four ribonucleoside and deoxyribonucleoside catabolizing enzymes is regulated at the transcriptional level. Initiation of transcription occurs from two promoters (P_1 and P_2) which are located 599 bp apart [Dandanell and Hammer, 1985]. Transcription from P_1 is repressed by DeoR, Whereas P_2 is subject to a double negative control by both DeoR and CytR, and is activated by CRP [Valentin-Hansen, 1985]. There are three operator sites in this operon for DeoR binding; O_1 and O_2 completely overlap P_1 and P_2, respectively. Another site O_E is 279 bp upstream of O_1. Full repression requires all three binding sites. The distance between O_1 and O_2 is 599 bp, but repression of P_2 can also be achieved by moving the O_1 1 to 5 Kb downstream of O_2 in a manner independent of orientation [Dandanell *et al.*, 1987]. *In vitro* studies show that the DeoR repressor cooperatively binds to two operator sites [Mortensen, *et al.*, 1989]. The electron microscopy observations show that when the DeoR repressor binds to its natural operator sites, single and double loops are formed [Amouyal *et al.*, 1989]. It has been suggested that a looped complex between O_1 and O_2 is required for the repression of transcription, and that binding to the O_E site may help repression by forming an insurance loop with O_1 or O_2 [Adhya, 1989].

Most promoters with distant duplications of operators require CRP for full activation. It was suggested that remote

operator duplications are usually associated with multiple promoters subject to more than one system of regulation [Collado-Vides *et al.*, 1991].

III-1-3. Repression involving interaction of Repressor with Other DNA Binding Proteins

In the classical model for negative regulation in prokaryotes, it is suggested that repressors block access of RNA polymerase to promoters [Majors, 1975]. Other mechanisms may also be involved in repression of transcription initiation. If the promoter depends on an activator for initiation of transcription, blocking the interaction between the activators and RNA polymerase also can lead to repression of transcription. Examples of such repression are the cI/Cro system of bacteriophage λ [Ptashne, 1986a], and some of the promoters regulated by both CytR and CAP [Martinussen *et al.*, 1989].

The CytR repressor regulates initiation of transcription from at least seven promoters which are involved in the expression of genes encoding proteins for the transport and catabolism of nucleosides and deoxynucleosides [Hammer-Jespersen, 1983; Munch-Petersen and Mygind, 1983; Valentin-Hansen *et al.*, 1986]. All promoters regulated by CytR are also activated by CAP.

As mentioned above, $deoP_2$ is under negative control by both DeoR and CytR repressors and activation by CAP. The studies of CytR regulation and CAP activation indicate that the CytR repressor inhibits CAP-dependent transcription

initiation and that tandem, intact CAP binding sites are required for CytR repression [Sogaard-Andersen *et al.*, 1990]. It is proposed that CytR and CAP may compete for the binding to some DNA sequences and that the binding of each excludes the binding of the other. Alternatively, CytR could bind to the operator site without precluding CAP binding to DNA and antagonize the activation of CAP either by changing DNA conformation or by directly contacting CAP [Sogaard-Andersen *et al.*, 1990]. *In vitro* studies show that CytR alone interacts weakly with the $deoP_2$ promoter, whereas with cAMP-CAP, the binding affinity of CytR increases 1000-fold [Gerlach *et al.*, 1991]. It seems that CytR relies on CAP for strong binding to the $deoP_2$ promoter, and CAP acts both as an activator and a corepressor [Sogaard-Andersen *et al.*, 1991]. The studies of another CytR regulated promoter, $tsxP_2$, indicate that the CytR also relies on the presence of cAMP-CAP to regulate $tsxP_2$ transcription and that the formation of an active repression complex requires CAP as a corepressor [Gerlach *et al.*, 1991].

III-2. The Mechanisms of Positive Regulation

In *E. coli* $E\sigma^{70}$ promoters, the mechanisms of positive regulation are similar to those used for negative regulation. Comparison of the sequences of positively controlled promoters with those of other typical promoters show that sequences (-35 region and the first T of the -10 region) of positively controlled promoters are often different from the consensus sequence of typical promoters [Raibaud and Schwartz, 1984]. It is believed that sequences of positively controlled promoters

are deficient in information for one or more steps of transcription initiation by RNA polymerase [Adhya and Garges, 1990]. Indeed, activator proteins may act to replace the -35 region and provide a substitute signal for RNA polymerase [Makino *et al.*, 1988; Collado-Vides *et al.*, 1991]. An activator may function in transcription initiation by increasing K_b and/or K_f; either it acts by directly contacting RNA polymerase or by changing DNA conformation.

III-2-1. Activator Interacting with RNA Polymerase

Studies of activatable promoters indicate that most activator proteins bind near the -35 region and increase the rate or extent of RNA polymerase open complex formation at the promoter [Raibaud and Schwartz, 1984; Collado-Vides *et al.*, 1991]. Binding of one protein to a strong DNA site can help a second molecule bind to a weak affinity site [Hochschild and Ptashne, 1986]. Cooperative binding of an activator protein and RNA polymerase has also been shown in both CAP and λ cI-P_{RM} positive control systems [Li and Krakow, 1988; Ren *et al.*, 1988; Hwang and Gussin, 1988].

It is believed that direct protein-protein interaction between activators and their corresponding RNA polymerases is one of the important mechanisms for transcription activation. This mechanism has been suggested for activators that bind close to the RNA polymerase binding site, such as OmpR [Tsung *et al.*, 1990], AraC [Lee *et al.*, 1987], and CAP [Ren *et al.*, 1988; Straney *et al.*, 1989; Gaston *et al.*, 1990]. Since most of the activators of $E\sigma^{70}$ promoters in fact bind close to the

RNA polymerase site, we would expect direct interaction between activator and RNA polymerase to be the main activation mechanism for these promoters.

This interaction between the activator and RNA polymerase can lead to increase of either K_b or K_f or both. At the *lac* operon promoter, CAP has been shown to increase K_b by about 20-fold [Malan *et al.*, 1984]. The cI protein of bacteriophage λ on the other hand acts at the λ P_{RM} promoter by increasing the isomerization rate; the preserce of cI increases K_f by approximately 10-fold [Hawley and McClure, 1983]. Another λ activator, cII, affects both the binding affinity and the isomerization rate of RNA polymerase at the promoters it activates [Shih and Gussin, 1984].

III-2-2. DNA Bending and Activation

Sequence-directed and protein-induced DNA bending has been described in a variety of promoter regions and has been thought to be important for the regulation of gene expression, DNA recombination, and DNA replication [Liu-Johnson *et al.*, 1986; Goodman and Nash, 1989;]. DNA bending may increase the transcription initiation rate [Lamond and Travers, 1983; Bossi and Smith, 1984; Gourse *et al.*, 1986; McAllister and Achberger, 1988]. Activator-mediated DNA bending or other conformational changes may affect neighbouring promoters, making them more adequate for transcription initiation by RNA polymerase [Adhya and Garges, 1990].

Upon binding to DNA, CAP protein in fact induces a substantial bending of the DNA more than 90^0 [Wu and Crothers,

1984; Liu-Johnson *et al.*, 1986]. This may be related to regulation since a synthetic bent DNA sequence can replace the CAP binding site and activate transcription from the CAP-dependent *gal* promoter and *lac* promoter even in the absence of CAP. These results suggest that bending induced by CAP may constitute its mechanism of transcription activation [Bracco *et al.*, 1989; Gartenberg and Crothers, 1991].

Two models have been proposed to explain the role of the CAP-induced bend. The first explanation is steric: the bend makes protein-DNA and/or protein-protein contacts possible which are not sterically feasible with a straight promoter and these are essential for the transcription initiation [Wu and Crothers, 1984]. The second explanation is energetic: the energy stored in the bend may be used to facilitate the transcription initiation [Zinkel and Crothers, 1991].

Two other general regulatory proteins, FIS and IHF, also can bind and bend DNA. The DNA bending mediated by these two proteins is also thought to play an important role in their biological functions. FIS (Factor for Inversion Stimulation) is a heat-stable protein that was first identified as a host factor required for site specific DNA inversion reactions such as the Cin system of phage P_1 [Huber *et al.*, 1985], the Gin system of phage Mu [Kahmann *et al.*, 1985], and the Hin system of *Salmonella typhimurium* [Johnson and Simon, 1985]. These site specific recombination systems are stimulated by FIS binding to a *cis*-acting recombinational enhancer sequence. It is believed that binding of FIS to two specific sites separated by a well defined intervening sequence induces DNA

bending. This conformational change is required for the recombinational enhancer activity [Huber *et al.*, 1989].

FIS was also found to be involved in the transcriptional regulation of the rRNA operon *rrnB* [Gourse *et al.*, 1986; Ross *et al.*, 1990] and some tRNA operons such as *tyrT* [Lamond and Travers, 1983], *leuV* [Bauer *et al.*, 1988], and *thrU* (*tufB*) [van Delft *et al.*, 1987; Nilsson *et al.*, 1990]. The expression of these operons is highly coordinated. All of these operons have an upstream activating sequence (UAS) located between positions -154 and -48 [Ross *et al.*, 1990; Nilsson *et al.*, 1990]. These UASs are AT-rich and display DNA bending [Gourse *et al.*, 1986]. FIS binding to the UAS may enhance the bending of the UAS DNA and activate transcription [Nilsson *et al.*, 1990]. At the *rrnB* P_1 promoter of *E. coli*, binding of FIS to the UAS increased transcription 20- to 30-fold both *in vivo* and *in vitro*. Even in the absence of FIS, the UAS activated transcription 2- to 4-fold *in vivo*, presumably due to its conformational effect [Leirmo and Gourse, 1991].

Another regulator which is thought to work by DNA bending is the integration host factor (IHF), a heterodimeric protein of *E. coli* encoded by the *himA* and *hip* genes [Nash and Robertson, 1981]. This sequence-specific DNA binding protein [Craig and Nash, 1984] participates in a wide variety of biological functions in bacteriophages, bacteria and plasmids, including a variety of types of site-specific recombination, phage packaging and partition, and DNA replication [Drlica and Rouviere-Yaniv, 1987; Stenzel *et al.*, 1987; Friedman, 1988]. In phage λ site-specific recombination system, protein-induced

bending and protein-mediated looping cooperate to form a functional multiprotein-DNA complex. In this complex, the DNA bending induced by IHF can stimulate the binding of integrase [Moitoso de Varges *et al.*, 1989]; and binding sites for IHF can be successfully replaced by a sequence-directed bend fragment [Goodman and Nash, 1989]. It is suggested that the function of IHF is to bend the DNA [Moitoso de Varges *et al.*, 1989].

IHF is also directly involved in the regulation of gene expression of many other genes in phage λ and Mu, *E. coli*, and other gram-negative bacteria [Drlica and Rouviere-Yaniv, 1987; Friedman, 1988; Gober and Shapiro, 1990; Hoover *et al.*, 1990]. In *E. coli*, IHF was shown to be involved in activating the expression of *gyrA*, *IlvB* and *IlvGEDA* genes [reviewed in Drlica and Rouviere-Yaniv, 1987], and in inhibiting the transcription of *ompC*, *ompF* and *ompB* genes [Huang *et al.*, 1990; Tsui *et al.*, 1991]. The gene products of *himA* and *hip* also negatively regulate their own expression as well as the expression from the other [Miller *et al.*, 1981].

Where IHF appears as an activator, it does not itself activate transcription, but rather increases the activity of a another required activator. The IHF binding sites are always located between the polymerase and activator binding sites. IHF then helps the activation by bending the DNA. This IHF-mediated DNA conformation change would bring the bound activator into a position favourable for contact with the RNA polymerase-promoter complex [Hoover *et al.*, 1990; Collado-Vides *et al.*, 1991]. In this mechanism, IHF would resemble

certain eukaryotic general regulatory proteins, which regulate the expression of a particular gene by interacting with other regulatory proteins [Hoover *et al.*, 1990; Mitchell and Tijan, 1989].

Other mechanisms are also involved in the activation of transcription. In the ara_{BAD} operon, CAP may activate transcription initiation by breaking the $araO_2$-$araI_1$ repression loop [Lobell and Schleif, 1991]. In the activation of transcription initiation at divergent promoters *malEp* and *malKp*, a complex nucleoprotein structure formed at the regulatory region is believed to involve wrapping of DNA around a core formed by the two bound activator proteins, MalT and CAP. This nucleoprotein complex may be involved in activation of transcription [Raibaud *et al.*, 1989] and supercoiling plays a crucial role in the formation and stability of this initiation complex [Richet and Raibaud, 1991]. In this system, MalT is absolutely required for activation, while CAP acts as an accessory factor [Raibaud *et al.*, 1989]. Recent studies suggest that the activation of *malKp* depends on the repositioning of MalT induced by CAP [Riched *et al.*, 1991].

III-2-3. The Activation of σ^{54} Promoters

As mentioned above, σ^{54} is a relatively minor σ factor found in *E. coli* and related bacteria. $E\sigma^{54}$ recognizes different basal elements located at -12 and -24 region. All σ^{54} promoters studied are regulated solely by activation rather

than by repression; in most cases, the activator binding site is located far from the promoter [Collado-Vides *et al.*, 1991].

In the absence of an activator, $E\sigma^{54}$ is thought to bind to the promoter and forms stable, but inactive, closed complexes [Sasse-Dwight and Gralla, 1988; Popham *et al.*, 1989]. The subsequent binding of an activator activates transcription by catalyzing isomerization of this closed complex to an open complex [Niafa *et al.*, 1989; Wedel *et al.*, 1990]. By contrast, $E\sigma^{70}$ can bind to a promoter and melt the promoter DNA in the absence of an activator [Gralla, 1990].

Another mechanistic difference between $E\sigma^{70}$ and σ^{54} is that in σ^{54} promoters, activator binding sites are centred farther from the transcription site, usually near position -110. Indeed, the NR_I binding sites of *glnAp2* can be moved more than 1 Kb upstream or downstream and still have 50% activity [Reitzer and Magasanik, 1986]. The binding sites of two other σ^{54} promoter activators, FhlaB and NIfA, can be moved 2 Kb (from *FdhF* and *nIfH*, respectively) and still retain more than 10% activity [Birkmann and Bock, 1989; Buck *et al.*, 1986]. As shown by electron microscopy, the activation of σ^{54} promoters also involves the activator touching RNA polymerase at the promoter by a looping out of the intervening DNA [Su *et al.*, 1990]. Some of σ^{54} promoters also use IHF as a coactivator. IHF can stimulate activation by bending DNA and stabilizing the formation of a looped complex which facilitate productive contacts between the activator and $E\sigma^{54}$ [Hoover *et al.*, 1990].

The observations indicate that σ^{54} promoters resemble eukaryotic polymerase II promoters in that σ^{54} system contains

an enhancer-like element, makes use of activators with eukaryotic-type functional domains [Sasse-Dwight and Gralla, 1990], requires ATP hydrolysis for initiation [Popham *et al.*, 1989], uses coactivators, and uses a phosphorylated enhancer protein [Weiss and Magasanik, 1988].

Some activators can also act as repressors. Most of these dual-function regulators mediate negative autoregulation. However several repress transcription of other genes. Thus, CAP which can repress several promoters [Collado-Vides *et al.*, 1991], in addition to the many it activates. Similarly, ArcA can repress many promoters and activate some, but the mechanism for its regulation is still unknown.

Whether a protein activates or repressed may depend on the properties of the protein, the DNA site where the protein binds, and other properties of the promoter. One important factor is the position of the binding site. A data base analysis shows that 60% of activatable promoters have sites overlapping a -40 sequence, whereas 49% of repressible promoters have operator sites overlapping the transcription start site [Collado-Vides *et al.*, 1991]. In most cases, when an activator functions as a repressor, it binds at or near the transcription start site. These observations suggest that the position on the DNA to which a regulatory protein binds- in particular its distance from the transcription start site, may determine whether activation or repression occurs [Collado-Vides *et al.*, 1991].

The nature of the promoter itself is also a determining factor in the repression/activation balance. When the same

ompR binding site was fused to a weak and a strong promoter, both from *E. coli*, OmpR enhanced RNA polymerase binding to the weak promoter and activated transcription, but repressed transcription from the strong promoter [Tsung *et al.*, 1990]. It is suggested that the enhancement of RNA polymerase binding to the *ompF* promoter by OmpR can stabilize the closed complex formation and inhibit the isomerization step [Tsung *et al.*, 1990].

IV. GENES KNOWN TO BE REGULATED BY L-LEUCINE

The presence of L-leucine in the growth medium can affect the activity of many enzymes in *E. coli*. One would certainly expect that the activity of enzymes which are involved in L-leucine metabolism may be affected by L-leucine. But many enzymes whose activity is affected by L-leucine are physiologically not related to branched-chain amino acid metabolism. L-leucine may act as an effector for a global response in *E. coli* [Fraser and Newman, 1975; Newman *et al.*, 1976]. In this work, we found that in most cases, the effect of L-leucine is mediated by the Lrp protein. Some of the genes regulated by L-leucine will be reviewed in this section.

IV-1. Genes Related to Branched-Chain Amino Acid Metabolism

The expression of genes whose products participate in L-isoleucine and L-valine biosynthesis is derepressed when L-leucine is limiting [Umbarger, 1987]. This includes *IlvBN*, *IlvIH*, *IlvC*, and *IlvGMEAD*. The expression of most of these genes is subject to multivalent repression requiring all three

branched-chain amino acids [Freundlich *et al.*, 1962]. Regulation of *IlvGMEAD* and *IlvBN* operons is also mediated by transcription attenuation control. In L-leucine limiting conditions, the restricted formation of leucyl-tRNA leads to the derepression of these two operons [Umbarger, 1987]. Expression of *IlvC* gene is controlled by substrate induction. Limiting L-leucine leads to an increase of acetohydroxy acid synthase II activity, and this in turn induces the expression of the *IlvC* gene [Umbarger,1987].

The *IlvIH* operon encodes the two subunits of acetohydroxy acid synthase III (AHAS III), one of the three isoenzymes which catalyse the first biosynthetic step common to branched-chain amino acid synthesis [Squires *et al.*, 1981]. Northern hybridization experiments indicate that transcription of the *IlvIH* operon is repressed by L-leucine, but not by L-isoleucine or L-valine [Squires *et al.*, 1981]. A DNA binding protein, termed *IlvIH*-binding (IHB) protein, can bind to two upstream regions of the *IlvIH* promoter, and the binding activity is decreased by L-leucine *in vitro*. The binding of the IHB protein to this upstream DNA is thought to be involved in the regulation of this operon by leucine [Ricca *et al.*, 1989]. A Tn*10* insertion within the *lhb* (*lrp*) gene, the gene that encodes IHB (Lrp) protein, decreases transcription of the *IlvIH* operon more than 30-fold [Platko *et al.*, 1990]. These results suggest that the IHB (Lrp) protein activates transcription of *IlvIH* and that L-leucine represses transcription by affecting the IHB (Lrp) function [Platko *et al.*, 1990].

Lrp also affects transcription from the four genes involved in leucine biosynthesis, *leuABCD*, located in a single operon. The expression of these genes depends upon L-leucine availability, decreasing in the presence of L-leucine and increasing when L-leucine is limiting [Burns *et al.*, 1966]. This regulation has been ascribed solely to transcription attenuation-which was thought to be the only mechanism of regulation modulated by L-leucine availability [Umbarger, 1987]. The leader sequence of the *leu* operon contains four adjacent Leu codons, which have been shown to be very important in transcription attenuation control of *leu* operon expression [Carter *et al.*, 1985]. In this study we show that the *leu* operon may also be subject to the regulation of Lrp protein.

Lrp may also be involved in a third branched-chain amino acid function, transport of branched chain amino acids into the cell. There are at least three kinetically distinct active transport systems for branched-chain amino acids transport [Quay *et al.*, 1977; Oxender *et al.*, 1977; Yamato and Anraku, 1980]. A low-affinity membrane bound system (LIV-I) is specific for all of the three branched-chain amino acids. Two shockable binding proteins are involved in two high affinity transport systems. One binding protein can mediate the transport of L-leucine, L-isoleucine, L-valine, L-threonine, L-alanine, or L-serine (LIV-binding protein, *livJ* gene product). The other binds specifically only to L- or D-leucine (Leucine-binding protein, *livK* gene product). A mutation in *livR* (*lrp*) results in an increase in both osmotic shock-

sensitive high affinity transport systems and in the level of both binding proteins [Oxender *et al.*, 1977]. It has been shown that transcription from *IlvJ* and *IlvK* occurred at high, constitutive levels in an *lrp*::Tn*10* mutant [Haney *et al.*, 1992]. These results suggest that *IlvJ* and *IlvK* are subject to the negative control of Lrp, and that L-leucine is required for this repression [Haney et *al.*, 1992].

IV-2. Other Genes Affected by L-Leucine

The expression of many genes that are unrelated to L-leucine metabolism are also affected by the presence of L-leucine in the growth medium. These gene products are involved in biosynthesis, transport systems and biodegradation.

IV-2-1. Genes Involved in Biosynthesis

Among genes whose products are involved in biosynthesis unrelated to L-leucine, three are known to be affected by L-leucine. They partipitate in amino acid biosynthesis (*serA*), amino acyl-tRNA synthesis (*lysU*), and methyl group donor synthesis (*metK*).

The *serA* gene of *E. coli* codes for phosphoglycerate dehydrogenase, the first enzyme in the biosynthesis pathway of L-serine and glycine. The level of this enzyme is decreased in cells when L-leucine and certain amino acids are present in the growth medium [McKitrick and Pizer, 1980]. We found that either the presence of L-leucine or a mutation in *lrp* gene can decrease transcription from the *serA* promoter as measured by the expression of a *lacZ* reporter gene.

Two-dimensional gel electrophoresis studies have shown that there are multiple forms of lysyl-tRNA synthetase (LysRS) in *E. coli* [Hirshfield *et al.*, 1981]. These multiforms of LysRS are encoded by two differently regulated genes. Forms I and III are encoded by a constitutive gene, *lysS*, which has been mapped at 62.1 min [Emmerich and Hirshfield, 1987]. An inducible gene located at 93.5 min, *lysU*, codes for LysRS forms II and IV [VanBogelen *et al.*, 1983]. The *lysU* gene shows almost no expression under normal growth conditions and can be induced in a number of conditions, including the presence of L-leucine, glycine-L-leucine dipeptides, L-alanine, and heat shock [Hirshfield *et al.*, 1981; Matthews and Neidhardt, 1989].

It was reported that the expression of *lysU* is increased in a *metK* mutant [Hirshfield *et al.*, 1977; Matthews and Neidhardt, 1989]. In this study we show that the *metK* mutant used in those experiments is in fact a *metK lrp* double mutant, and suggest that the constitutive expression of *lysU* may result from the *lrp* mutation rather than the *metK* mutation.

The *metK* gene encodes S-adenosylmethionine (SAM) synthetase which catalyses the reaction of L-methionine with ATP to produce SAM [Greene *et al.*, 1973; Hafner *et al.*, 1977]. SAM has been shown to play numerous important roles in cell metabolism. It serves as the major methyl group donor in biological systems, is involved in polyamine biosynthesis, and functions as a corepressor in the regulation of the methionine biosynthesis regulon [Tabor *et al.*, 1958; Greene *et al.*, 1973; Hobson, 1974; Shoeman *et al.*, 1985]. The presence of L-leucine in the growth medium increases the level of SAM synthetase [Greene *et al.*, 1973].

IV-2-2. Genes Involved in Biodegradation

As mentioned before, there are two differently regulated genes coding for L-serine deaminases (L-SDs). L-SD I, the gene product of *sdaA*, is produced in both minimal medium and rich medium, whereas L-SD II, the *sdaB* gene product, is only produced in rich medium [Su and Newman, 1991]. In the presence of L-leucine in minimal medium, the expression of *sdaA* is increased [Su *et al.*, 1989; Su and Newman, 1991]. In this study, we show that the effect of L-leucine is mediated by the Lrp protein. In the regulation of *sdaA* expression, Lrp may act as a repressor and L-leucine as an inducer. In the case of *sdaB*, Lrp may indirectly act to activate the expression in an L-leucine-dependent manner.

L-threonine dehydrogenase (TDH) is the first enzyme of the two-step pathway converting L-threonine to glycine and acetyl coenzyme A [Newman *et al.*, 1976]. Some organisms such as *Pseudomonas multivorans* and certain *Fusarium* can use threonine as carbon and energy source via this pathway [Lessie and Whiteley, 1969; Willets 1972]. In *E. coli*, which is unable to use L-threonine as sole carbon source, TDH is involved in the formation of glycine [Newman *et al.*, 1976; Ravnikar and Somerville, 1987], but the actual function of this enzyme may be degradation rather than biosynthesis [Newman *et al.*, 1976, Potter *et al.*, 1977]. The level of TDH is induced by the presence of L-leucine [Newman *et al.*, 1976]. TDH is encoded by *tdh* which has been mapped at 81.2 min [Ravnikar and Somerville, 1987].

The work in this thesis shows that the activity of TDH is increased in an *lrp* mutant, suggesting that Lrp represses the expression of *tdh* gene, and that L-leucine acts as an inducer. Recent studies showed that the Lrp target site is located in a 25-bp region between -69 and -44 of the *tdh* promoter [Rex *et al.*, 1991].

IV-2-3. Genes Involved in Transport

The binding protein-dependent oligopeptide transport system of *E. coli* and *S. typhimurium* can transport di- and oligopeptides with up to five amino acid residues [Payne and Gilvary, 1968]]. Four genes (*oppABCD*) specifying an oligopeptide permease similar to the system in *S. typhimurium* have been mapped at 27 min on the *E. coli* chromosome [Lenny and Margolin, 1980].

Transcription from these genes is affected both by L-leucine and by Lrp. Studies with *opp-lac* operon fusions showed that the addition of either L-leucine or L-alanine increased transcription of the *opp* operon [Andrews and Short, 1986a; 1986b]. Transcription of this operon was also shown to be negatively regulated by the gene product of *oppI*, which was shown to be identical to *lrp* [Austin *et al.*, 1989, Platko *et al.*, 1990].

Another transport system of *E. coli* may also be regulated by Lrp, L-serine can be transported by four different transport systems: the serine-threonine transport system, one of the LIV systems, the glycine-alanine system, and the L-

serine-specific transport system [Hama *et al.*, 1988]. The L-serine-specific system is a H^+-coupled cotransport and is an inducible system. Like L-serine deaminase I (the product of *sdaA*), the L-serine-specific system is induced by L-leucine, but not by L-serine [Hama *et al.*, 1988]. This suggests that the structural gene for the L-serine-specific transport system and the *sdaA* gene may be regulated in the same way [Hama *et al.*, 1988].

What is the metabolic function of L-leucine and Lrp? These observations and our studies suggest that the *lrp* gene product is a regulator of a global response to L-leucine-the leucine regulon. Our further studies show that *gcv* operon are regulated by Lrp and are not affected by L-leucine. This system was renamed the leucine/*lrp* regulon.

MATERIALS AND METHODS

PART 1. STRAINS, BACTERIOPHAGES, AND PLASMIDS

The strains, bacteriophages, and plasmids used in this study are described in Table 1.

PART 2. CULTURES, MEDIA, AND GROWTH TESTS

2-1. Minimal Medium:

The minimal medium used, neutralized to pH 7, has been described elsewhere [Newman *et al.*, 1985a]. Cultures were grown with L-isoleucine and L-valine, 50 µg/ml each, added to all media used to grow strain CU1008 and all its derivatives to compensate for the deletion in *ilvA* carried by these strains.

2-2. SGL Medium:

Medium with a combination of L-serine, glycine and L-leucine as the only carbon source other than L-isoleucine and L-valine is called SGL medium. L-Serine, glycine and L-leucine were usually provided at 2,000, 300, and 300 µg/ml, respectively (unless otherwise noted).

2-3. NSIV Medium

Medium with 0.2% of L-serine to replace D-glucose as sole carbon source.

TABLE 1. Strains, Bacteriophage and Plasmids

Strain, phage or plasmid	Description	Reference or source
E. coli K-12		
CU1008	*E. coli* K-12 *ilvA*	L.S. Williams
KEC9	CU1008 *ssd*	Newman *et al.*, 1982b
DRN-1	*serA*::Mu d1	Ramotar and Newman, 1986
MEW1	Δ*lacZ* derivative of strain CU1008	Newman *et al.*, 1985b
MEW22	MEW1::λp*lac*Mu9 SGL$^-$ Kan^r isolated by transduction from MEW21 to MEW1 and selecting for Kan^r	Su *et al.*, 1989
MEW26	MEW1 *lrp*::Tn*10*	This study
MEW28	MEW1 *sdaA*::Cm^r	Su, 1991
MEW30	MEW1 *metK62*	This study
MEW31	MEW1 *metK62 rbl*::Tn*10*	This study
MEW32	MEW31 *sdaA*::Cm^r from MEW28	This study
MEW33	MEW1 *rbl-62*	This study
MEW34	MEW1 *metK62 rbl-62*	This study
MEW35	MEW22 *ssd*	This study
MEW36	MEW22 *rbl*::Tn*10*	This study
MEW37	DRN-1 *ssd*	This study
MEW38	DRN-1 *rbl*::Tn*10*	This study
MEW39	DRN-1 *rbl-62*	This study
MEW40	MEW38 *sdaA*::Cm^r	This study
MEW41	DRN-1 *ihb*	This study

Cont.

MEW42	CV975 *rbl*::Tn*10*	This study
MEW43	MEW1 *lhb*::Tn*10*	This study
MEW44	MEW22 *lhb*::Tn*10*	This study
MEW45	MEW1 *lrp*::*lacZ*	This study
MEW85	MEW1 *serA* (select by UV treatment)	This study
MEW86	MEW85 *lrp*::Tn*10*	This study
MEW87	MEW45 *metJ*::Cm^r	This study
CP1-68	MEW1 carrying λp*lac*Mu9 insertions in Lrp-regulated genes	This study
A401	HfrC *polA1*	Russel and Holmgrem, 1988
XL1	*recA*$^-$ (*recA1*, *lac*$^-$, *endA1*, *gyrA96*, *thi*, *hsdR17*, *supE44*, *relA1*, (F' *proAB*, *lacI*^{Q}, *lacZ* M15, Tn*10*)	Stratagene Co.
RG	Wild-type *E. coli* K-12	Matthews and Neidhardt, 1988
RG62	RG *metK* spontaneous ethionine-resistant	Matthews and Neidhardt, 1988
CV975	*IlvIH*::*lacZ*	J. M. Calvo
CV1008	CV975 *lhb*::Tn*10*	J. M. Calvo
TK4100	MC4100 *metJ*::Cm^r	R. C. Greene
CAG5051	HfrH *nadA57*::Tn*10*	Singer *et al.*, 1989
CAG5052	KL227 *btuB3191*::Tn*10*	Singer *et al.*, 1989
CAG5053	KL208 *zbc-280*::Tn*10*	Singer *et al.*, 1989
CAG5054	KL96 *trpB83*::Tn*10*	Singer *et al.*, 1989

Cont.

CAG5055	KL16 *zed-3069*::Tn*10*	Singer *et al.*,1989
CAG8209	KL228 *zgh-3075*::Tn*10*	Singer *et al.*,1989
CAG8160	KL14 *thl-39*::Tn*10*	Singer *et al.*,1989
CAG12206	HfrH *nadA3052*::Tn*10*kan	Singer *et al.*,1989
CAG12204	KL227 *btuB-3192*::Tn*10*kan	Singer *et al.*,1989
CAG12203	KL208 *zbc-3105*::Tn*10*kan	Singer *et al.*,1989
CAG12202	KL96 *trpB-3193*::Tn*10*kan	Singer *et al.*,1989
CAG12200	KL16 *zed-3120*::Tn*10*kan	Singer *et al.*,1989
CAG12205	KL228 *zgh-3159*::Tn*10*kan	Singer *et al.*,1989
CAG12201	KL14 *thl-3178*::Tn*10*kan	Singer *et al.*,1989
CAG12079	MG1655 *fuc-3072*::Tn*10*	Singer *et al.*,1989
CAG12135	MG1655 *recD-1901*::Tn*10*	Singer *et al.*,1989
CAG18604	MG1655 *zgb-3156*::Tn*10*kan	Singer *et al.*,1989
CAG12112	MG1655 *zbl-3109*::Tn*10*	Singer *et al.*,1989
CAG18478	MG1655 *zbl-1203*::Tn*10*	Singer *et al.*,1989
CAG18528	MG1655 *zbl-3110*::Tn*10*kan	Singer *et al.*,1989
CAG12095	MG1655 *zac-3051*::Tn*10*	Singer *et al.*,1989
CAG12168	MG1655 *zgd-210*::Tn*10*	Singer *et al.*,1989
CAG18450	MG1655 *zhf-50*::Tn*10*	Singer *et al.*,1989
CAG18638	MG1655 *zhg-3086*::Tn*10*	Singer *et al.*,1989
CAG18488	MG1655 *zjd-2231*::Tn*10*	Singer *et al.*,1989
CAG18427	MG1655 *zje-2241*::Tn*10*	Singer *et al.*,1989
WM1776	*asnB*, *asnC*::Km, *lacY*, *thl*	W. Messer

Cont.

Phages		
M13K07		Vieira and Messing 1987
λp*lac*Mu9	λ p*lac*MuKm^r	Bremer *et al.*, 1985
λpMu507	λ cI ts857 *Sam*7 Mu*A+B+* (Helper phage)	Bremer *et al.*, 1985
λTn10	λcts *Sam*S53	Wood, 1981
λ1F10	λKohara phage214	Kohara *et al.*, 1987
Plasmids		
pBR322		Bolivar *et al.*, 1977
Bluescript	KS+ and KS−,	Stratagene Co.
pMC1871	pBR322 carrying *lacZ*	R.K. Storms
pACYC184		Chang and Cohen, 1978
pSD100		Moskaluk and Bastia, 1988
pLSK35-3		Kolling and Lother, 1985
pLSK92-2		Kolling and Lother, 1985
pMES28	plasmid pSD100 containing the *sdaA* gene in-frame fusion	Su, 1991
pMES22	pBR322 carrying *sdaA* gene	Su, 1991
pGS146	pACYC184 carrying *gcv*	Stauffer *et al.*, 1987
pGSU1	Bluescript$^+$ carrying 2 kp *sal*I-to-*Sma*I fragment from pGS146	This study
pWB6	Bluescript$^+$ carrying *sdaB*	Z.Q. Shao
pFN120	pBR322 carrying *lysU* gene	Clark and Neidhardt, 1990

Cont.

pFU2	Bluescript$^+$ carrying 3.3 Kb *Eco*RI fragment from pFN120	This study
pFU10	Bluescript$^+$ carrying 108 bp *Sau*3A-to-*Dra*I fragment from pFU2	This study
pFU20	Bluescript$^+$ carrying 256 bp *Sau*3A fragment from pFU2	This study
pGT17	pBR325 carrying *serA* gene	G. A. Grant
pGU2	Bluescript$^+$ carrying 1.3 Kb *Hin*dIII-to-*Bam*HI fragment from pGT17	This study
pGU3	Bluescript$^+$ carrying 522 bp *Bst*EII-*Pvu*II fragment from pGU2	This study
pGU4	Bluescript$^+$ carrying 214 bp *Bst*EII-to-*Mse*I fragment from pGU2	This study
pGU5	Bluescript$^+$ carrying 238 bp *Ssp*I-to-*Hin*PI fragment from pGU2	This study
pGU6	Bluescript$^+$ carrying 277 bp *Taq*I fragment from pGU2	This study
pGU7	Bluescript$^+$ carrying 144 bp *Mse*I-to-*Ssp*I fragment from pGU2	This study
pSU1	Bluescript$^+$ carrying 684 bp *Sal*I-to-*Ava*II fragment from pMES3	This study
pSU2	Bluescript$^+$ carrying 344 bp *Eco*RV-to-*Hin*dIII fragment from pSU1	This study
pSU3	Bluescript$^+$ carrying 173 bp *Ssp*I-to-*Hin*dIII fragment from pSU2	This study

Cont.

pSU4	Bluescript$^+$ carrying 171 bp *Ssp*I-to-*Eco*RV fragment from pSU2	This study
pLR1	pACYC184 carrying 6.5 Kb *Irp*$^+$ *Hin*dIII fragment from λ1F10	This study
pLR2	pBR322 carrying 6.5 Kb *Irp*$^+$ *Hin*dIII fragment from pLR1	This study
pLR3	Bluescript$^+$ carrying 1.2 Kb *Irp*$^+$ *Hin*dIII-to-*Hpa*II fragment from pLR1	This study
pLR4	Bluescript$^-$ carrying 1.2 Kb *Irp*$^+$ *Hin*dII-to-*Hpa*II fragment from pLR1	This study
pLRN1	pACYC184 carrying 1.2 Kb *Irp*$^+$ *Hin*dII-to-*Hpa*II fragment from pLR1	This study
pLRN2	pACYC184 carrying a *Irp*::*lacZ* fusion	This study
pLRE1	pLR4 carrying a *Eco*RI site in *Irp* stop codon created by in vitro mutagenesis	This study
pLRE2	Bluescript$^+$ carrying *Irp' Eco*RI fragment from pLRE1, stop codon near *Kpn*I site	This study
pLRE3	Opposite orientation of *Eco*RI fragment in pLRE3	This study
pLRE4	Bluescript$^+$ carrying an *Irp-lacZ* 2-part fusion	This study
pLRE5	Bluescript$^+$ carrying an *Irp*-collagen-*lacZ* 3-part fusion	This study
pLRD1	Bluescript$^+$ carrying downstream region of *Irp* gene	This study

Cont.

pLRD2	Bluescript$^-$ carrying downstream region of *lrp* gene	This study
pLRU1	Bluescript$^+$ carrying 0.5 Kb *Eco*RI-to-*Bgl*II fragment from pLR3	This study
pLRU2	Bluescript$^-$ carrying 0.5 Kb *Eco*RI-to-*Bgl*II fragment from pLR3	This study
pTS7		S. Adhya
pBRZ1	pBR322 carrying promoterless *lacZ* from pTS7	This study
pSUZ2	pBRZ1 carrying 373 bp *Hin*dIII-to-*Bam*HI fragment from pSU2	This study
pSUZ3	pBRZ1 carrying 202 bp *Hin*dIII-to-*Bam*HI fragment from pSU3	This study
pSUZ4	pBRZ1 carrying 201 bp *Hin*dIII-to-*Bam*HI fragment from pSU4	This study

2-4. Determination of Doubling Times

The doubling times of cultures were calculated from growth curves determined by measuring turbidity with sidearm flasks fitted for a Klett colorimeter and using a blue (#42) filter. To do this, overnight cultures were subcultured, allowed to grow to exponential phase and diluted into the medium in which growth was tested. Turbidity was determined every 30 or 60 minutes.

2-4. Determination of Nutritional Requirements

To determine whether amino acids could be used as nitrogen source, precultures were grown with a reduced amount of ammonium sulfate (500 µg/ml) and then diluted 4,000 times into test medium so as to minimize the possibility of carry over of ammonium sulfate. Growth was followed as indicated above.

To determine whether a compound could be used as carbon source, it was used to replace D-glucose at a final concentration of 0.2%.

2-5. Other Additions to the Medium

Antibiotics were used at the following concentrations, in µg/ml: Ampicillin (Amp) 100, tetracycline (Tet) 15, kanamycin (Kan) 80, chloramphenicol (Chl) 25 and streptomycin (Str) 100.

PART 3. ENZYME ASSAYS

3-1. L-Serine Deaminase

L-SD was assayed as previously described in toluene-treated whole cells [Newman *et al.*, 1985b]. One unit of L-SD

in the whole cell assay is defined as the amount of enzyme which catalyzes the formation of 1 µmol of pyruvate in 35 min.

Assays of L-SD in LB-grown cells present the problem that results could be confused by deamination of L-serine due to another known enzyme, the biodegradative L-threonine deaminase [Goss and Datta, 1984]. To avoid this, LB cultures were grown with increased (0.5%) glucose so as to repress formation of that enzyme.

3-2. ß-Galactosidase Assay

Cells were grown to the log-phase in the test medium. ß-Galactosidase activity was assayed in whole cells according to the method described by Miller and expressed in his units [Miller, 1972].

3-3. Protein Assay

Protein concentration was determined by the method of Lowry [Lowry *et al.*, 1951] with bovine serum albumin as the standard.

3-4. Cystathionine-ß-Lyase Assay

Cystathionine-ß-lyase was assayed in whole cells by the method of Hunter *et al.* (Hunter *et al.*, 1975), using toluene instead of lysozyme, as suggested to us by R.C. Greene.

3-5. Proline Oxidase Assay

Proline oxidase was assayed by a slight modification of the method of Dendinger and Brill (Dendinger and Brill, 1970),

using whole cells grown in glycerol minimal medium with 1 mg/ml L-proline.

3-6. Tryptophanase Assay

Tryptophanase was assayed by the method of Botsford and DeMoss (Botsford and DeMoss, 1971), using whole cells grown in glycerol minimal medium in which the nitrogen source was replaced with 500 µg/ml L-tryptophan.

3-7. Threonine Dehydrogenase Assay

L-Threonine dehydrogenase assays were made in toluene-permeabilized cells by the modified method of McGilvray and Morris [McGilvray and Morris, 1971] as previously described by Potter *et al.* [Potter *et al.*, 1977]. Mid-log phase cells grown in 250 ml flasks were chilled in ice, harvested by centrifugation, washed once with cold Tris buffer (10 mM, pH 8.5), and resuspended in Tris buffer at a turbidity of 500 Klett units (540 filter). The assay mixture contained: 0.3 ml of cells, 0.1 ml of 0.3 M L-threonine, 0.1 ml of 25 mM NAD, and 0.02 ml toluene. Assay blanks with water in place of L-threonine were used. The reaction (37^0C 20min) was started with the addition of L-threonine and stoped with 0.05 ml of cold 30% trichloroacetic acid. Protein were removed by centrifugation, and samples were assayed for aminoacetone by the rapid determination method of Urata and Granick [Urata and Granick, 1963] with the modified Ehrlich reagent of Mauzerall and Granick [Mauzerall and Granick, 1956] using delta-aminolevulinic acid as a standard.

3-8. Determination of Yield

Yield of protein per unit carbon was determined by the method described by Newman *et al.* [Newman *et al.*, 1976], on cultures which had reached stationary phase as judged by constancy of turbidity measurements.

3-9. Analog Sensitivity Tests

Analog sensitivity was tested by placing various concentrations of γ-glutamyl methyl ester (GGME) on filter disks [Kraus *et al.*, 1979], and also by streaking strains to be tested on glucose minimal medium plates containing 2 mM GGME.

PART 4. STRAIN CONSTRUCTIONS

4-1. Isolation of λ Tn*10* Insertion in the *lrp* Gene

Strain MEW1 was infected with λ::Tn*10* as described by J. Wood (Wood, 1981) and tetracycline resistant cells able to use L-serine as sole carbon source were selected on NSIV plate containing tetracycline (15 µg/ml). The *lrp* mutation from one isolate was transduced to strain MEW1, selecting for tetracycline resistance and verifying use of L-serine as sole carbon source. This strain was called MEW26.

4-2. Isolation of Mutant Strains Carrying Protein Fusions in Genes Whose Expression is Controlled by L-Leucine and the Lrp Protein

Principle: λp*lac*Mu9(kan^r) [Bremer *et al.*, 1985] was inserted into the strain MEW1 genome and kanamycin-resistant

cells whose expression was affected by L-leucine were screened and kept for further studies.

Detail: Strain MEW1 was infected with λp*lac*Mu9 and helper phage λpMu507, as described by Weinstock and his co-workers [Bremer *et al.*, 1984, Bremer *et al.*, 1985]. Cells were incubated for 1 hour in LB medium, centrifuged, washed and plated on lactose minimal medium with kanamycin 80 µg/ml, with or without L-leucine 300 µg/ml. The resultant colonies were streaked on LB kanamycin, and on minimal kanamycin medium with or without L-leucine. ß-galactosidase assays were made on strains in which the growth rate on lactose minimal medium seemed to be altered by the presence of 300 µg/ml L-leucine. The *lrp*::Tn*10* mutation was then transduced into relevant strains, and ß-galactosidase again assayed. Mutations in genes affected by Lrp and L-leucine were transduced by kanamycin resistance into strain MEW1, and the *lrp* double mutants recreated by the same transduction as before. The strains with target mutations transduced into strain MEW1, known here as CP1-CP66, and the corresponding *lrp* derivatives, were used in this study.

4-3. Isolation of an L-Serine-Requiring Mutant

Strain MEW1 was treated by ultraviolet irradiation, subcultured in minimal medium for 1 hour, and ampicillin was added to 50 µg/ml. After about 4 hours further incubation, cells were washed and spread on LB plates. Survivors of the

ampicillin selection were screened for those which might carry a *serA* mutation by testing for a strain which required either L-serine or glycine. That the mutation was in *serA* was confirmed by demonstrating 20% cotransduction with *metK*, and by cotransduction with a Tn*10* insertion in Singer strain CAG 18604.

4-4. Isolation of Mutations Carrying a Protein Fusion in the Glycine Cleavage (GCV) Genes

Principle: From a random λp*lac*Mu insertion pool in a *serA* strain, the mutants which were able to grow on minimal lactose kanamycin plate with serine but unable to use glycine as serine source were screened and kept for further study.

Detail: Strain MEW85 *serA* was infected with λp*lac*Mu9 and λpMu507 as above. After being washed with saline, cells were incubated in minimal glucose kanamycin medium with glycine for 60 min. Ampicillin was then added to final concentration of 50 µg/ml and the culture incubated for a further 3-4 hours. The cells were then washed, resuspended in saline and spread on minimal lactose kanamycin plates with L-serine. Colonies which grew with glucose and L-serine but not with glucose and glycine were screened and retained as possible GCV mutants [Newman *et al.*, 1974]. This attribution was confirmed by mapping with the Singer mapping kit [Singer *et al.*, 1989], and by a further transduction into MEW85 *serA* to verify the phenotype. For physiological studies, two independently isolated GCV mutations were transduced into strain MEW1, creating strains CP67 and CP68.

4-5. Isolation of a Strain Carrying a Protein Fusion in the *lrp* Gene

Strain MEW85 *serA* was infected with λp*lac*Mu9 and treated as above, and then spread on an NSIV kanamycin plate to select for *lrp* mutants. Strains able to grow on glucose minimal medium when supplemented with L-serine but not glycine were retained as possible *lrp* mutants. The strains that carried insertions in *lrp* were confirmed by transducing from an *lrp*::Tn*10* (Tet^r) mutant into the putative *lrp*::λp*lac*Mu (Kan^r), and showing that 100% of the 50 tetracycline-resistant transductants tested were sensitive to kanamycin.

4-6. Constructions of Other Strains

Other strains were constructed by transductions as follows.

lrp from strain MEW26 was transferred by selecting for tetracycline resistance and verifying that the transductants grew with L-serine as carbon source, except for derivatives of MEW22 and MEW28, in which this verification could not be made.

To transfer *rbl-62*, the Tn*10* from Singer strain 18478 was first transduced into RG62, selecting for tetracycline resistance, and the tetracycline-resistant transductants which grew with L-serine as carbon source were used as donors of *rbl-62*. *rbl-62* was cotransduced with Tn*10* and verified by growing on NSIV plate.

metK62 from strain RG62 was transduced into strain DRN1

serA or MEW85 *serA*, selecting for serine-independent growth. These transductions were done in the presence of L-leucine, which facilitates growth of *metK* mutants. Strains were verified by GGME resistance. Transductions of *metK62* were also done by transducing a Tn*10*kan from Singer strain 18604 into RG62, transducing kanamycin resistance and screening for transductants which were resistant to GGME.

ssd was transferred from strain KEC9 in two steps. First the recipient was made *metB* by selecting tetracycline resistance from Singer strain 5052 *btuB*::Tn*10* *metB*. A methionine-requiring recipient was transduced to methionine independence using an *ssd* donor, and other characteristics of the strain were verified as required.

lhb::Tn*10* was transferred from strain CV1008 by selecting tetracycline resistance.

sdaA::Cm^r was transferred from strain MEW28 by selecting chloramphenicol resistance and verifying that the strains could not grow on SGL plates.

PART 5. HYBRIDIZATION

Hybridization was performed by the method of Southern [Maniatis *et al.*, 1982] with minor modifications, and using gels dried according to the method of R.K.Storms. The DNA fragments, isolated from agarose gels, were randomly oligo-labelled with α-^{32}P-dATP by the procedure of the supplier (Boehringer-Mannheim).

PART 6. SITE-DIRECTED MUTAGENESIS

Oligonucleotide-directed mutagenesis was performed by the method of Kunkel [Kunkel,1987] with the Muta-Gene Phagemid in vitro mutagenesis kit from Bio-Rad Laboratories.

6-1. Creating an *Eco*RI Site in the Stop Codon of the *lrp* Gene

In order to fuse the ß-galactosidase to the C-terminal of the Lrp protein, an *Eco*RI site was created in pLR4 by using the oligonucleotide 5'-CACCTGTTCCGAATTCGCGCGTCTTAATAAC-3' (mismatches are underlined). To do this, 200 ng of oligonucleotide was phosphorylated, and annealed to U-strand DNA prepared from strain CJ236 carrying plasmid pLR4. The synthesis of the complementary strand was carried out following the protocol in the instruction manual from Bio-Rad. The reaction mixture was transformed into strain XL1. Plasmids isolated from 24 transformants were digested with *Eco*RI and analyzed on agarose gels. One of these plasmids showed a new *Eco*RI site and was named pLRE1.

6-2. Creating an *Eco*RI Site in the Upstream Region of the *lrp* Gene

In order to construct a hybrid plasmid in which the expression of *lrp* gene could be under the control of other promoters, an *Eco*RI site was created at 15 bp upstream of the start codon. This new *Eco*RI site was created by using the oligonucleotide 5'-GTCTCTCTGAATTCCTTC-3' (mismatch is underlined) using the method described above.

PART 7. PLASMID CONSTRUCTIONS

7-1. Subcloning the *Irp* Gene from Kohara Phage

The *serA Irp* double mutant cannot use glycine as a source of L-serine. I took advantage of this to subclone the *Irp* gene from Kohara phage. Kohara phage λ1F10 was digested with *Hin*dIII, and the resulting *Hin*dIII fragment ligated into the *Hin*dIII site of plasmid pACYC184. The ligation mixture was transformed to a *serA Irp* strain, and transformants which could grow on glucose minimal medium with glycine were selected. One plasmid isolated from one of these transformants, pLR1, carried an ~6.5 Kb *Hin*dIII fragment from λ1F10. This 6.5 Kb *Hin*dIII fragment was inserted into the *Hin*dIII site of pBR322 to yield plasmid pLR2. An ~1.25Kb *Hin*dIII-to-*Hpa*I fragment from pLR1 was subcloned into the *Hin*dIII and *Eco*RV sites of plasmids pACYC184, Bluescript^{+} and Bluescript^{-} to produce pLRN1, pLR3, and pLR4, respectively (Fig. M1). All these plasmids could complement the *Irp* mutation, as judged by ß-galactosidase activity transcribed from the *sdaA* and other promoters.

7-2. Construction of Plasmids Carrying *Irp* Upstream and Downstream Region for Sequencing

To subclone the downstream region of the *Irp* gene, plasmids pLR3 and pLR4 were digested with *Mlu*I and *Eco*RI, made blunt ended with klenow enzyme and dNTPs, and ligated with T4 ligase. Self-ligation of the large *Mlu*I-to-*Eco*RI fragments

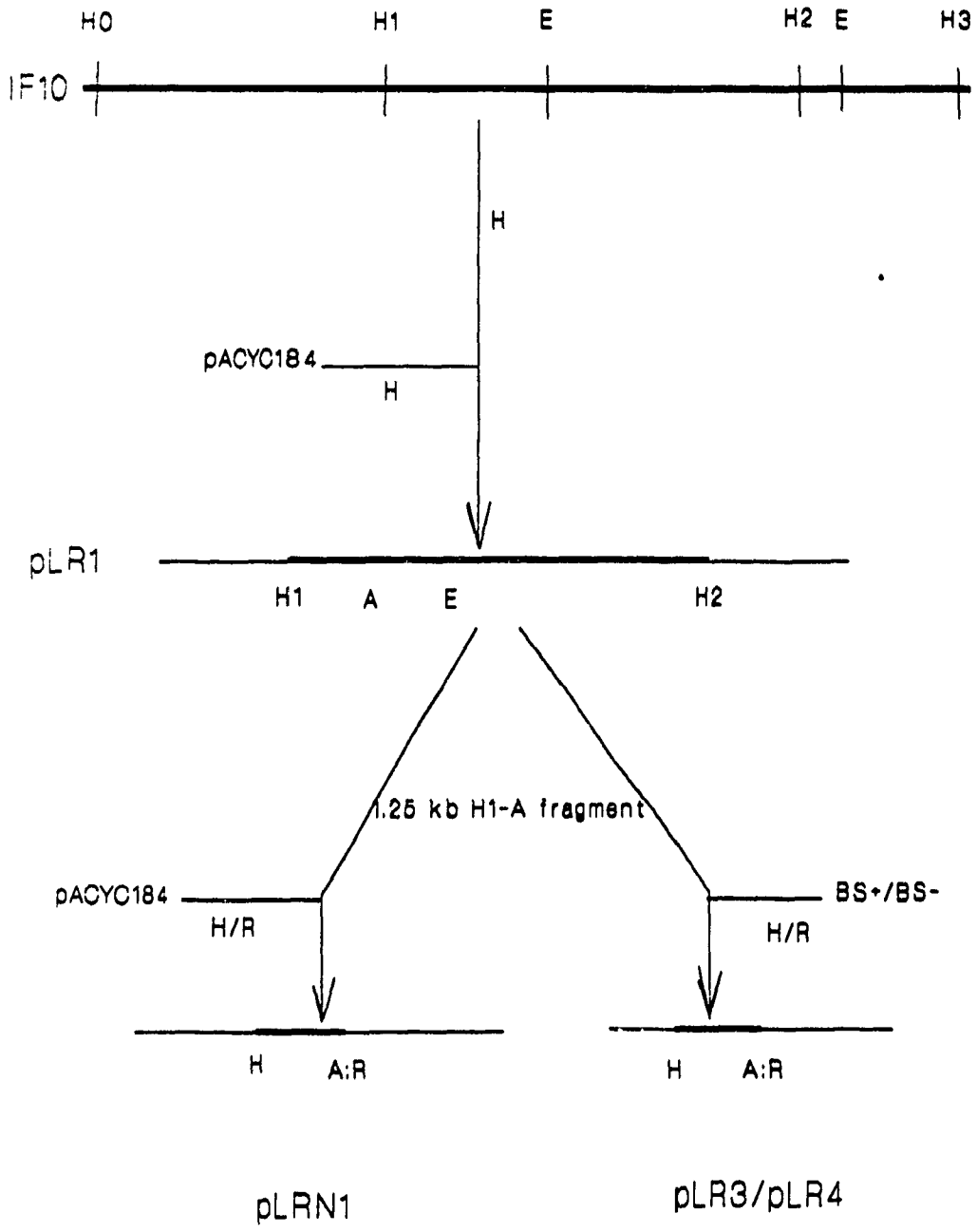


Figure M1. cloning of Irp gene from Kohara phage lamda 1F10

Figure M1. Cloning of the *Irp* gene from Kohara phage λ1F10. The Kohara phage λ1F10 was digested with *Hin*dIII, ligated with pACYC184 digested with the same restriction enzyme, and transformed into a *serA Irp* strain. Plasmid pLR1, which carrying a ~6.5 kb *Hin*dIII fragment from λ1F10, was isolated from one of transformants growing on glucose minimal medium supplied with glycine. A ~1.25 kb *Hin*dIII-to-*Hpa*I fragment from plasmid pLR1 was isolated from an agarose gel, inserted into the *Eco*RV, *Hin*dIII sites of plasmids pACYC184, Bluescript^{+}, and Bluescript^{-} to produce plasmids pLRN1, pLR3, and pLR4, respectively. Abbreviations: H, *Hin*dIII; E, *Eco*RI; A, *Hpa*I; R, *Eco*RV. The thick line corresponds to *E. coli* DNA and the thin line corresponds to plasmid DNA.

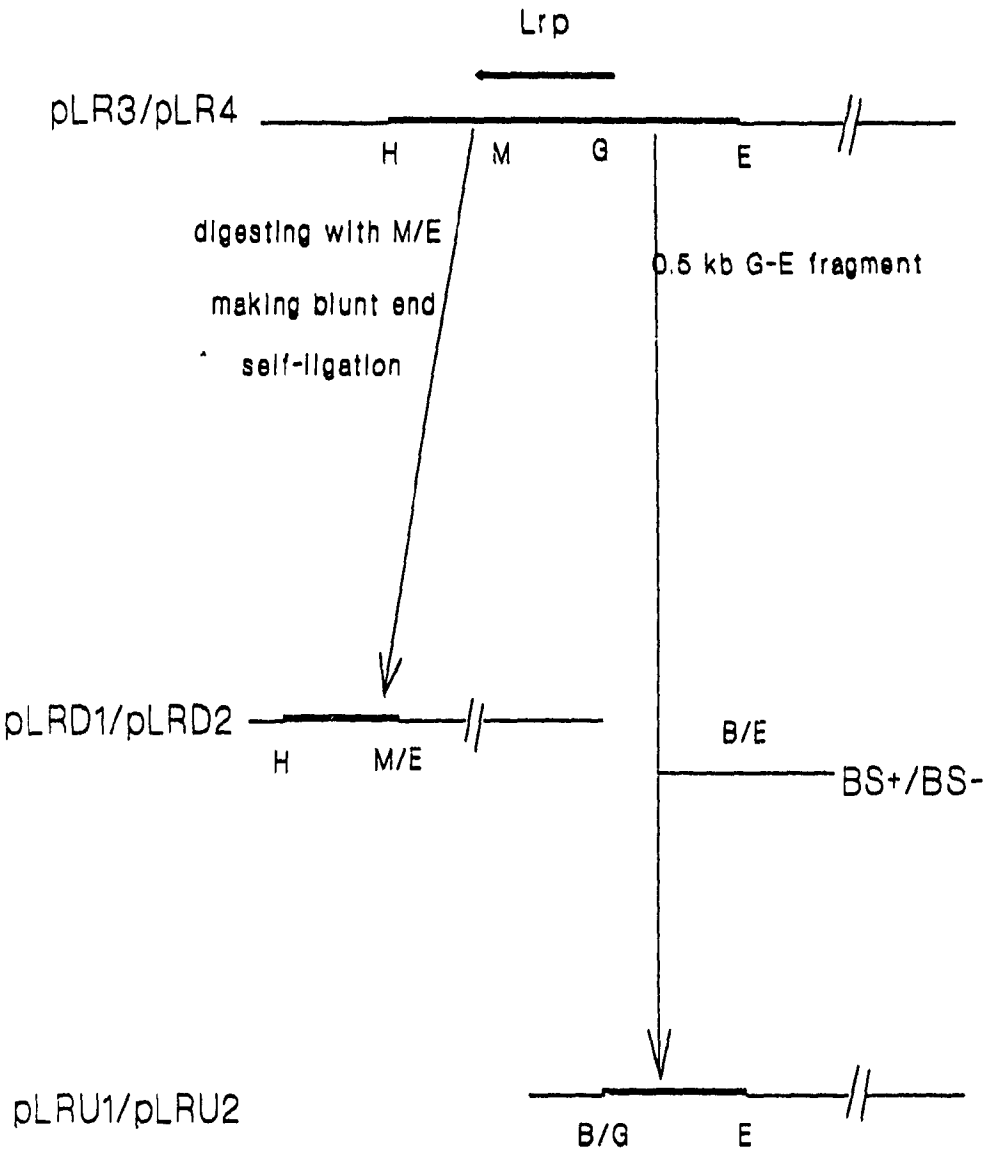


Figure M2. Construction of plasmids carrying the lrp upstream or downstream region

Figure M2. Construction of plasmids carrying the *lrp* upstream or downstream region. To construct plasmids carrying the *lrp* upstream region, a ~500 bp *Bgl*II-to-*Eco*RI fragment from plasmid pLR3 was isolated from an agarose gel and inserted into the *Bam*HI, *Eco*RI sites of Bluescript$^+$ and Bluescript$^-$, forming plasmids pLRU1 and pLRU2, respectively. To construct plasmids carrying the *lrp* downstream region, plasmids pLR3 and pLR4 were digested with the *Eco*RI, *Mlu*I, made blunt end with Klenow enzyme and 4 dNTPs and subjected to self-ligation. The resulting plasmids, pLRD1 and pLRD2, carrying the *lrp* downstream region in Bluescript$^+$ and Bluescript$^-$, respectively. Abbreviations: H, *Hin*dIII; E, *Eco*RI; M, *Mlu*I; G, *Bgl*II. The thick line corresponds to *E. coli* DNA and the thin line corresponds to plasmid DNA. The size and direction of translation of *lrp* are marked by an arrow.

from pLR3 and pLR4 generated plasmids pLRD1 and pLRD2, respectively (Fig. M2). The 0.5 Kb *Eco*RI-to-*Bgl*II fragment from pLRN3 was isolated, inserted into the *Eco*RI and *Bam*HI sites of Bluescript$^+$ and Bluescript$^-$, producing pLRU1 and pLRU2, respectively (Fig. M2).

7-3. Construction of Plasmids Containing an *lrp*::*lacZ* Fusion

Plasmid pLRN2 carried an in-frame *lacZ* fusion to *lrp* gene at the unique *Bgl*II site. To construct this plasmid, a 3.1 Kb *Bam*HI fragment carrying part of the *lacZ* gene from pMC1871 was isolated, and inserted into the *Bgl*II site of pLRN1 (Fig. M3).

7-4. Construction of Plasmids Carrying *lrp-lacZ* and *lrp*-Collagen-*lacZ* Fusions.

A 1 Kb *Eco*RI fragment from pLRE1 was isolated and inserted into the *Eco*RI site of Bluescript$^+$. In one of the resulting plasmids, pLRE2, the newly created *Eco*RI site was near the *Eco*RV site (Fig. M4). Another plasmid, pLRE3, carried the same fragment, inserted in the opposite direction (Fig. M5). These two plasmids were used for the construction of 2-part and 3-part fusion proteins.

The *lrp-lacZ* fusion plasmid was constructed by isolating a 3.1 Kb *Sma*I-to-*Sal*I fragment from PMC1871, ligating with pLRE2 which was digested with *Eco*RV and *Sal*I (Fig. M4). The resulting plasmid was named pLRE4.

To construct an *lrp*-collagen-*lacZ* fusion plasmid, a 6 Kb *Sma*I-to-*Sal*I fragment from pMES28 was isolated from agarose

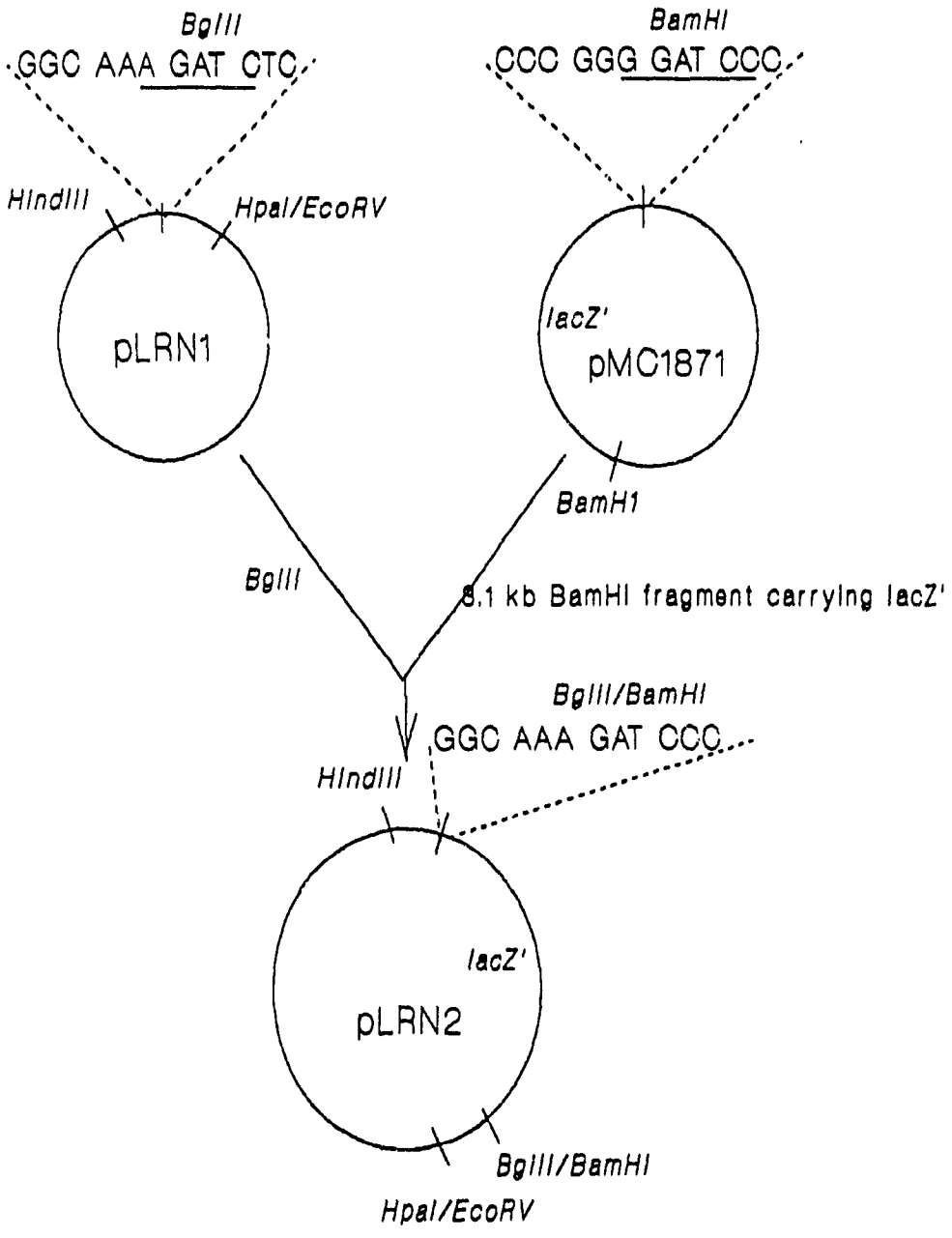


Figure M3. Construction of plasmid pLRN2 (Irp::lacZ).

Figure M3. construction of plasmid pLRN2 (*lrp*::*lacZ*). Plasmid pMC1871 was digested with *Bam*HI and a ~3.1 kb fragment carrying *lacZ'* (lacking both its promoter and the coding sequence for the several N-terminal amino acids) was isolated from an agarose gel. This fragment was inserted into the *Bgl*II site of pLRN1 and transformed into MEW1. Plasmid pLRN2 was isolated from one of the resulting colonies which showed ß-galactosidase activity. Its size (8.6 kb) and restriction digest pattern were checked.

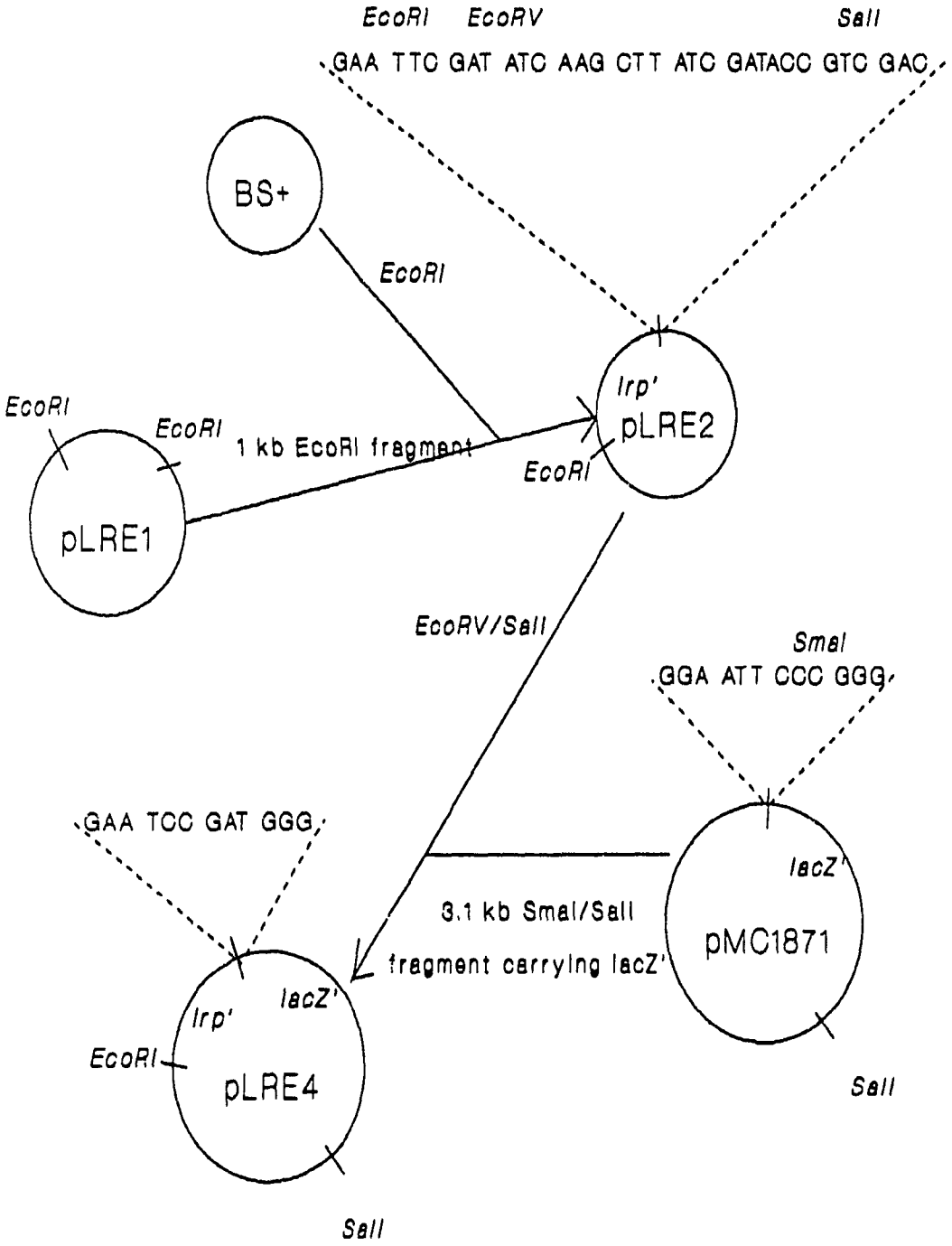


Figure M4. Construction of plasmid pLRE4.

Figure M4. Construction of plasmid pLRE4 carrying the intact Lrp coding region fused in frame to the *lacZ* gene. To construct a plasmid coding for an Lrp-ß-galactosidase protein fusion, a 1 kb *Eco*RI fragment from plasmid pLRE1 (stop codon of *lrp* was mutated; see section 6-1) was isolated from an agarose gel and inserted into the *Eco*RI site of Bluescript^{+}. In one of the resulting plasmids, pLRE2 (4 kb), the insertion was in the direction which placed its newly created *Eco*RI site near the *Eco*RV site. A ~3.1 kb *Sma*I-to-*Sal*I fragment carrying *LacZ'* was isolated from plasmid pMC1871 and inserted into the *Eco*RV and *Sal*I sites of pLRE2, and the ligation mix was used to transform MEW1. Plasmid pLRE4 was isolated from one of resulting colonies which shown ß-galactosidase activity. The size (7.1 kb) and restriction pattern of pLRE4 were checked. The reading-frame of both genes is indicated.

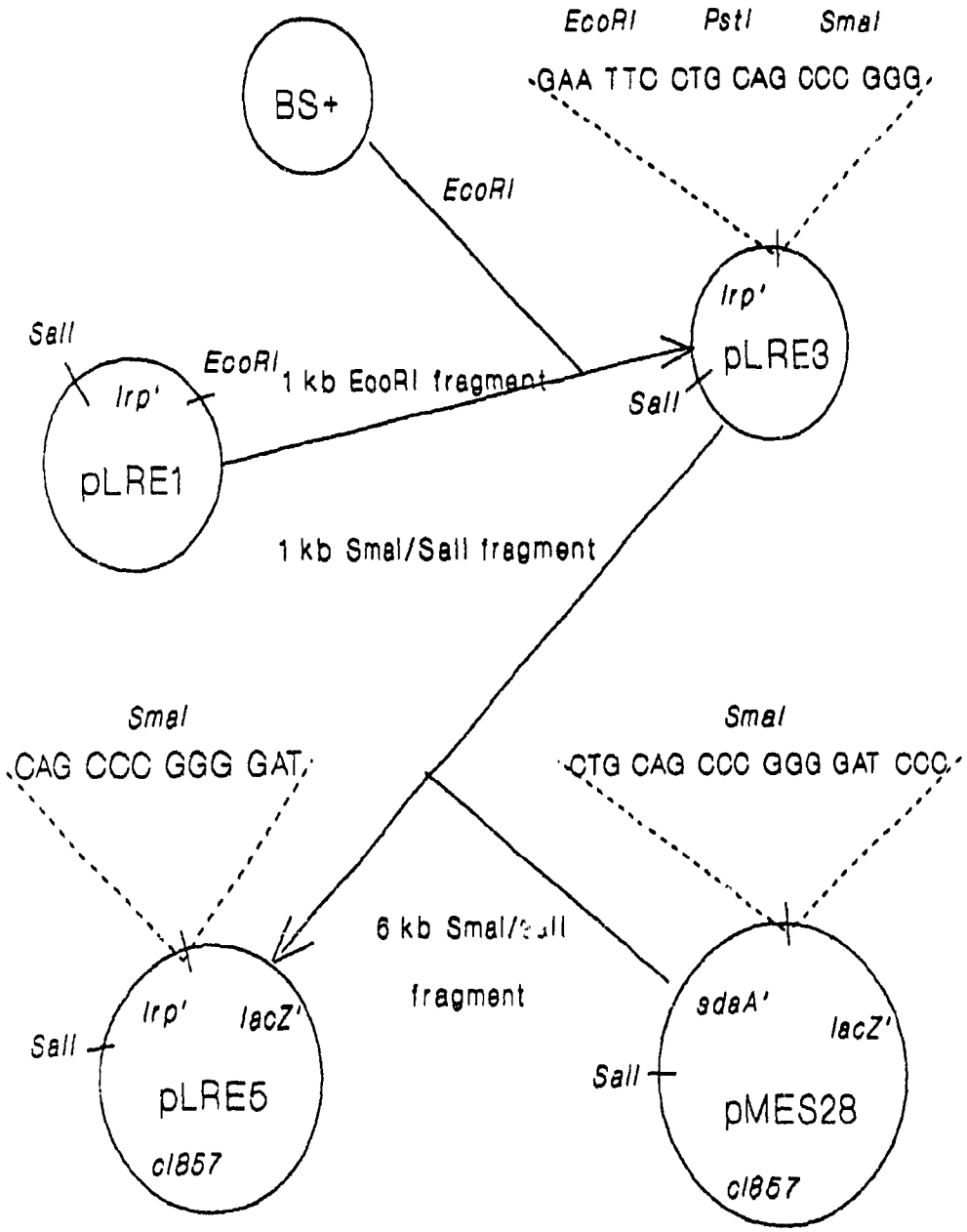


Figure M5. Construction of plasmid pLRE5.

Figure M5. Construction of plasmid pLRE5 carrying the intact *Irp* coding region fused in frame to the *lacZ* gene by means of an intervening collagenase sequence. A 1 kb *Eco*RI fragment from plasmid pLRE1 was isolated from an agarose gel and inserted into the *Eco*RI site of Bluescript^{+}. In one of the resulting plasmids, pLRE3 (4 kb), the newly created *Eco*RI site was near the *Pst*I site, that is, insertion in the opposite direction from pLRE2 (Fig. M4). A 1 kb *Sal*I-to-*Sma*I fragment from plasmid pLRE3 and a 6 kb *Sal*I-to-*Sma*I fragment from plasmid pMES28 (Su, 1991) were isolated from agarose gels and ligated with T4 DNA ligase. The ligation mixture was transformed into strain MEW1 and Amp resistant colonies were selected at 28^{0}C. A 7 kb plasmid isolated from one of these colonies showing the expected restriction site was named pLRE5. The reading-frame of these genes is indicated.

gels, and ligated with a ~1 kb *Sma*I-to-*Sal*I fragment isolated from pLRE3 to produce plasmid pLRE5 (Fig. M5).

Only preliminary experiments were done with these constructs, so results are not included in the text.

7-5. Construction of Plasmids Carrying *lysU* Upstream Region

Plasmid pFU2 (Table 1) containing the DNA fragments from the *lysU* upstream region was produced in the construction of plasmids pFU10 and pFU20, the sources of DNA fragment for binding assays and footprinting. A 3.3 Kb *Eco*RI fragment from *lysU*$^+$ plasmid pFN120 [Clark and Neidhardt, 1990] was subcloned into the *Eco*RI site of Bluescript$^+$, yielding plasmid pFU2 (Fig. M6). Plasmid pFU2 was then digested with *Dra*I, producing a 497 bp fragment carrying the *lysU* upstream sequence. This fragment was digested with *Sau*3A, and the resulting 108 bp fragment was ligated into the *Bam*HI and *Eco*RV sites of Bluescript$^+$, generating the plasmid pFU10 (Fig. M6). The plasmid pFU20 was constructed by inserting a 256 bp *Sau*3A fragment from pFU2 into *Bam*HI site of Bluescript$^+$ (Fig. M6).

7-6. Construction of Plasmids Containing the *sdaA* Upstream Region

Four plasmids (Table 1) carrying *sdaA* upstream regions were produced as the sources of DNA fragments for gel retardation assays. A 681 bp *Sal*I-to-*Ava*II fragment from pMES3 [Su *et al.*, 1989] was made blunt with Klenow enzyme and dNTPs and ligated into the *Eco*RV site of Bluescript$^+$ to yield plasmid pSU1 in which the *Ava*II end was near the *Hin*dIII site (Fig. M7). A 349 bp *Hin*dIII-to-*Eco*RV fragment was then

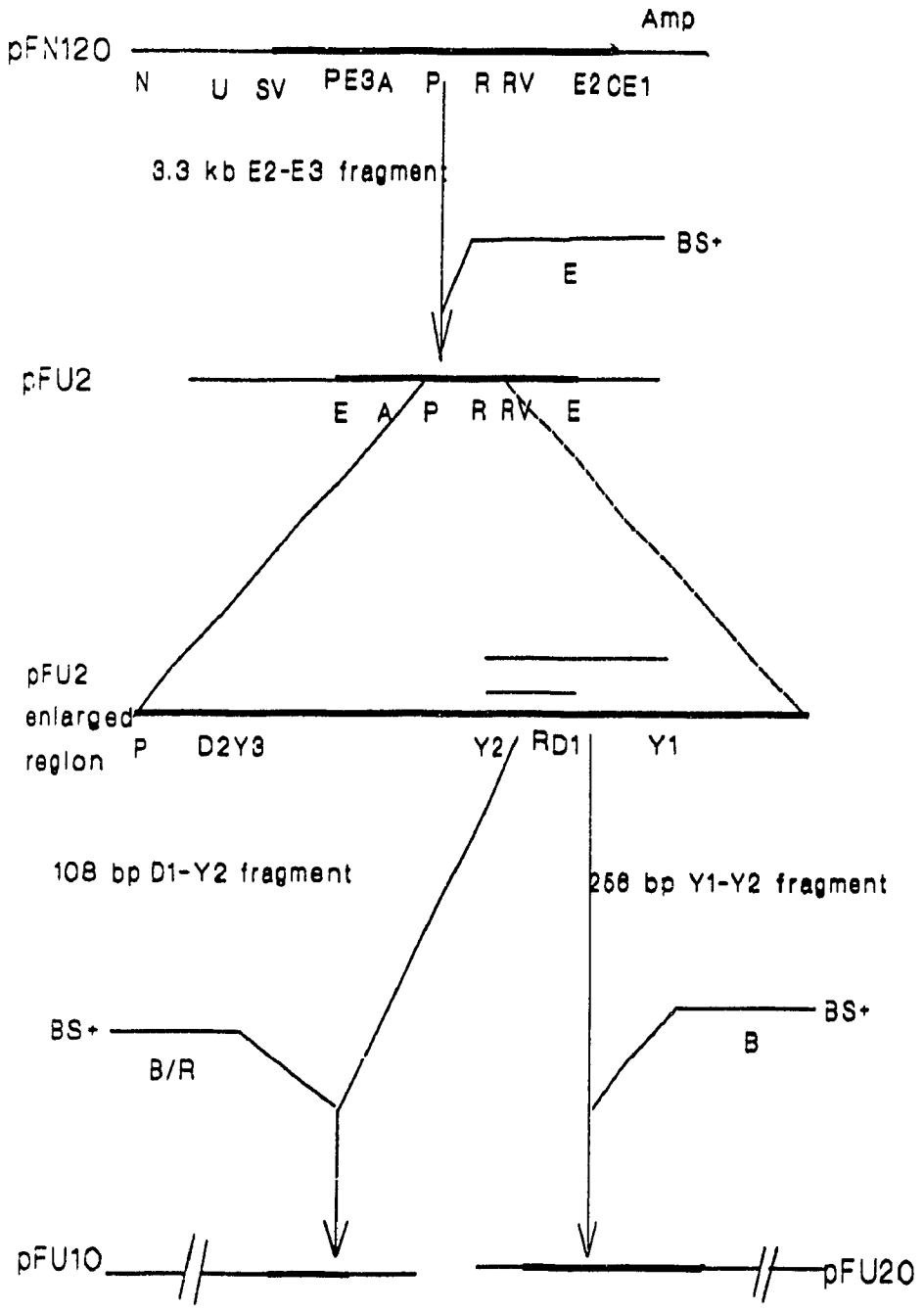


Figure M6. Construction of plasmids carrying the lysU upstream region.

Figure M6. Construction of plasmids carrying the *lysU* upstream region. A 3.3 kb *Eco*RI fragment carrying *lysU* upstream region from plasmid pFN120 [Clark and Neidhardt, 1990] was isolated from an agarose gel and inserted into the *Eco*RI site of Bluescript$^+$ to produce plasmid pFU2. Two subclones were made from pFU2. (1). A 108 bp *Dra*I-to-*Sau*3A1 fragment carrying *lysU* promoter region was isolated from plasmid pFU2 and inserted into the *Bam*HI and *Eco*RV sites of Bluescript$^+$, yielding plasmid pFU10. (2). A 256 bp *Dra*I fragment from plasmid pFU2 was isolated from an agarose gel and inserted into *Bam*HI of Bluescript$^+$ to produce plasmid pFU20. Abbreviations: H, *Hin*dIII; E, *Eco*RI; P, *Pst*I; V, *Pvu*II; B, *Bam*HI; C, *Cla*I; S, *Sal*I; N, *Nde*I; U, *Nhu*I; A, *Hpa*I; R, *Eco*RV; D, *Dra*I; Y, *Sau*3A; and BS$^+$, Bluescript$^+$. The thick line corresponds to *E. coli* DNA and the thin line corresponds to plasmid DNA. The dotted line represents the area of pFU2 from which subclones were made.

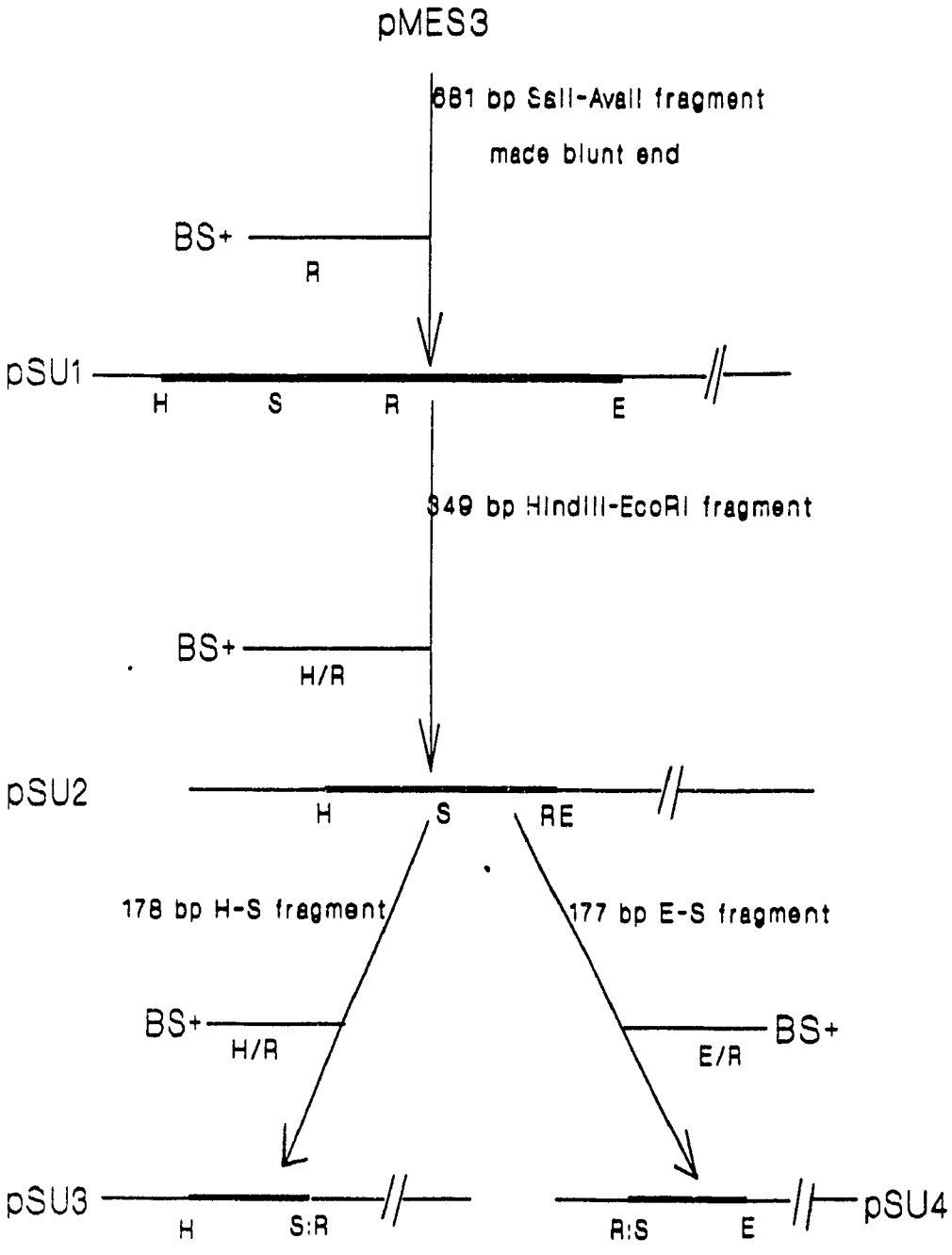


Figure M7. Construction of plasmids carrying the sdaA upstream region.

Figure M7. Construction of plasmids carrying the *sdaA* upstream region. The ends of a 681 bp *Sal*I-to-*Ava*II fragment from plasmid pMES3 (Su *et al.*, 1989) was made blunt with Klenow enzyme, isolated from a agarose gel and inserted into the *Eco*RV site of plasmid Bluescript$^+$ to yield plasmid pSU1. From plasmid pSU1, a 349 bp *Hin*dIII-to-*Eco*RV fragment was isolated and inserted into the *Hin*dIII and *Eco*RV sites of Bluescript$^+$, producing plasmid pSU2. Two subclones were made from plasmid pSU2. (1). A 178 bp *Hin*dIII-to-*Ssp*I fragment was isolated from plasmid pSU2 and inserted into the *Hin*dIII and *Eco*RV sites of Bluescript$^+$, producing plasmid pSU3. (2). Plasmid pSU4 was constructed by isolating a 177 bp *Eco*RI-to-*Ssp*I fragment from plasmid pSU2 and inserting into the *Eco*RI and *Eco*RV sites of Bluescript$^+$. Abbreviations: H, *Hin*dIII; E, *Eco*RI; S, *Ssp*I; R, *Eco*RV; and BS$^+$, Bluescript$^+$. The thick line corresponds to *E. coli* DNA and the thin line corresponds to plasmid DNA.

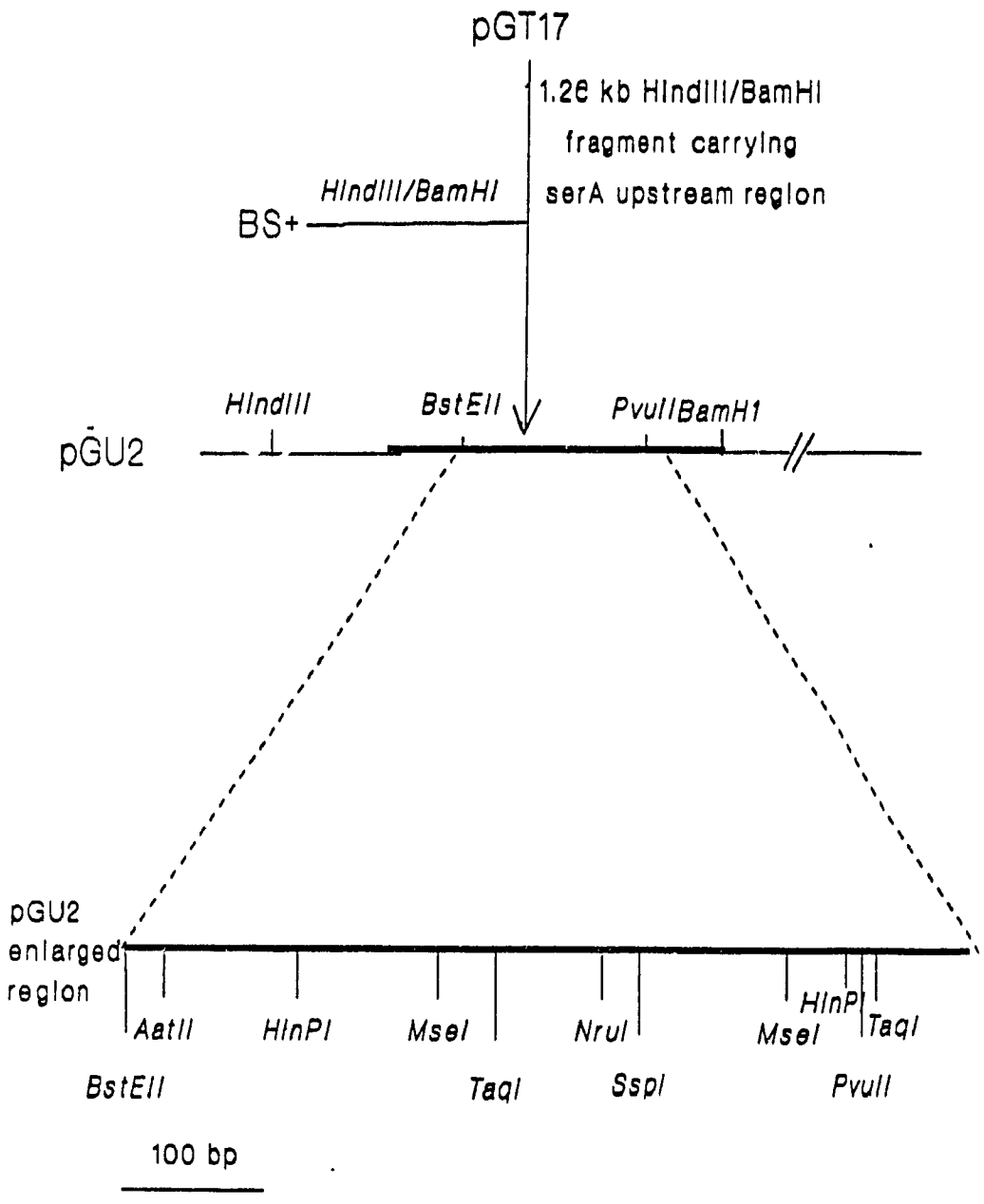


Figure M8. Construction of plasmid pGU2 and the description of the restriction map of the serA upstream region

Figure M8. Construction of plasmid pGU2 and the description of the restriction map of the *serA* upstream region. A 1.26 kb *Hin*dIII-to-*Bam*HI fragment carrying the *serA* upstream region from plasmid pGT17 (Tobey and Grant, 1986) was isolated and inserted into the *Hin*dIII and *Bam*HI sites of Bluescript$^+$ to produce plasmid pGU2. The restriction map of *serA* upstream region carried in this plasmid was shown in bottom. BS$^+$: Bluescript$^+$. A section of the *serA* region of plasmid pGU2 is enlarged and its restriction map is indicated. Several fragments discussed in section 7-7 were isolated from this plasmid.

Isolated from pSU1 and subcloned into *Hin*dIII and *Eco*RV sites of Bluescript$^+$, yielding plasmid pSU2 (Fig. M7). Plasmid pSU3 was constructed by ligating a 178 bp *Hin*dIII-to-*Ssp*I fragment from plasmid pSU2 into *Hin*dIII and *Eco*RV sites of Bluescript$^+$ (Fig. M7). A 177 bp *Eco*RI-to-*Ssp*I fragment from pSU2 was isolated , and inserted into *Eco*RI and *Eco*RV sites of Bluescript$^+$ to produce plasmid pSU4 (Fig. M7).

7-7. Construction of Plasmids Carrying the *serA* Upstream Region

A series of plasmids (Table 1) were constructed for producing fragments for binding assays. The plasmid pGT17 [Tobey and Grant, 1986] was digested with *Hin*dIII and *Bam*HI, a 1.26 Kb fragment was isolated and inserted into the *Hin*dIII and *Bam*HI sites of Bluescript$^+$ to yield plasmid pGU2 (Fig. M8). All plasmids carrying the *serA* upstream region were derived from plasmid pGU2. Its restriction map is shown in Fig. M8. A 522 bp *Bst*EII-to-*Pvu*II fragment was isolated from pGU2, the ends were made blunt with Klenow enzyme and NTPs, and ligated into *Sma*I site of Bluescript$^+$, producing the plasmid pGU3. A 364 bp *Bst*EII-to-*Ssp*I fragment from pGU2 was digested with *Mse*I, made the ends blunt, and inserted into *Eco*RV site of Bluescript$^+$ to produce plasmid pGU4, which contains 222 bp *Bst*EII-to-*Mse*I fragment, and pGU7, which carrys a 144 bp *Mse*I-to-*Ssp*I fragment. A 333bp *Aat*II-to-*Ssp*I fragment from pGU2 was digested with *Hin*PI. The large fragment (238 bp) was isolated and ligated into *Eco*RV and *Cla*I sites of Bluescript$^+$. The resulting plasmid, pGU5, contained the upstream sequence of *serA* gene from -23 to -260 related to

transcription start site. A 728 bp *Bam*HI-to-*Bst*EII fragment from pGU2 was digested with *Taq*I and subcloned into *Cla*I site of Bluescript$^+$. The resulting plasmid, pGU6, carried a 277 bp *Taq*I fragment from pGU2.

7-8. Construction of Plasmid pGSU1

A 2 kb *Sal*I-to-*Sma*I fragment containing the *gcv* upstream region was isolated from an agarose gel and inserted into the *Sal*I and *Eco*RV sites of Bluscript$^+$ polylinker to yield plasmid PGSU1.

7-9. Construction of Plasmids Carrying *sdaA-lacZ* Operon Fusion

A 4 kb *Eco*RI-to-*Pst*I fragment carrying promoterless *lacZ* from plasmid pTS7 was isolated from an agarose gel and inserted into the *Eco*RI and *Pst*I sites of pBR322. The resulting plasmid, named pBRZ1, was used for the construction of fusion plasmids. The *Hin*dIII-to-*Bam*HI fragments from plasmids pSU2 (373 bp), pSU3 (202 bp), and pSU4 (201 bp) were isolated from agarose gels and inserted into the *Hin*dIII and *Bam*HI sites of pBRZ1, yielding plasmids pSUZ2, pSUZ3, and pSUZ4, respectively.

PART 8. DNA SEQUENCING

8-1. DNA Sequencing by the Dideoxy-Mediated Chain Termination Method

DNA sequences of upstream and downstream regions of *Irp* gene were determined by the dideoxy-chain termination method of Sanger *et al.* [Sanger *et al.*, 1977], following the protocol

Indicated in the Sequenase kit from United States Biochemical Corporation, Cleveland, Ohio. The sequencing markers for primer extension were also determined by this method.

Single-stranded DNAs were prepared from strain XL1 harbouring Bluescript plasmids carrying the required fragment by superinfection with the helper phage K07. This was done by methods described in Su's thesis [Su, 1991].

8-2. DNA Sequencing by the Maxam-Gilbert Method

The A+G reaction of Maxam-Gilbert sequencing [Maxam and Gilbert, 1980] was used to prepare the sequencing marker for the DNase I footprint. To do this, 20,000 cpm 3'-end-labeled DNA was mixed with 4 mg sonicated herring sperm DNA in a volume of 24 µl. 4 µl of 1M piperidine formate was added and incubated at 37^0C for 15 min. After adding 240 µl of cold hydrazine stop solution, the DNA was precipitated with ethanol and cleaved in 100 µl of 1M piperidine at 90^0C for 30 min. The piperidine was removed by evaporating in a rotary vacuum lyophilizer. The DNA was redissolved in 10 µl of sequencing gel-loading buffer.

8-3. Sequencing Gel

The DNA sequencing gel system used was developed by F. Lang [Lang and Burger, 1990].

PART 9. GEL RETARDATION ASSAY

The DNA fragments used for gel retardation and footprinting studies were isolated from agarose gels using the MERmaidTM kit (Bio-101 Inc) by the procedure indicated in the

kit, and 3'-end labelled with α-^{32}P-dATP or dCTP using Klenow enzyme by the method described by Maniatis [Maniatis *et al.*, 1982].

The Lrp protein used in these studies was the generous gift of Joseph Calvo and Debbie Willins. The protein was purified to greater than 98% as described by Willins *et al.* [Willins *et al.*, 1991].

Binding of Lrp protein to the upstream region of *lysU*, *serA*, *sdaA*, *lrp* genes and other DNA fragments was determined by the gel shift assay described by Ricca *et al.* [Ricca *et al.*, 1989] with slight modifications. From 1-5 ng (5000 cpm) of 3'-end-labelled DNA fragment was incubated at room temperature for 10 min with 2 µg sonicated herring sperm DNA and 0-1000 ng Lrp protein in 20 µl binding buffer. Samples were resolved by electrophoresis through a 5% polyacrylamide gel pre-electrophoresed at 10 v/cm and electrophoresis of the samples was performed in the same conditions. The gel was cast and run in 1x TBE buffer [Maniatis, 1982]. The gels were dried at 60^{0}C in gel drier and subjected to radioautography.

PART 10. DNase I FOOTPRINTING

For DNase I footprinting, the plasmids carrying the source DNA fragments were digested with two different restriction enzymes, of which one generated 5' protruding ends which could be labelled with Klenow enzyme, and the other generated 3' protruding ends which could not be labelled with Klenow enzyme. After treatment with Klenow enzyme and α-P^{32}-dNTP,

these fragments were only labelled in one end and thus were suitable for DNase I footprinting. Alternatively, the plasmids were digested with one of the restriction enzymes which generated 3' protruding ends, labelled with klenow enzyme, cut with another restriction enzyme, and the appropriate fragment isolated from an agarose gel.

To determine the footprint, 5-20 ng (20,000 cpm) of end-labelled DNA fragment were incubated with 3 µg bovine serum albumin and Lrp protein in 20 µl of binding buffer as above. After 10 min at room temperature, 0.4 units of DNase I were added and incubated at 30^{0}C for 2 min. the reaction was terminated by adding 20 µl of stop solution (0.6M pH5.2 sodium acetate, 50 mM pH8.0 EDTA and 1 µg/µl sonicated herring sperm DNA). After ethanol precipitation, the pellet was resuspended in 6 µl of formamide loading buffer. Half of this sample was loaded onto a 7% or 8% polyacrylamide-8 M urea sequencing gel for electrophoretic resolution.

PART 11. RNA-PRIMER EXTENSION

11-1. 5'-End-Labelled Oligonucleotides

The oligonucleotide primers were purified from 20% polyacrylamide gel by the "Crush and Soak" method described by Maniatis [Maniatis *et al.*, 1982]. 100 ng of oligonucleotide primer 20-24 nucleotides long was 5' end-labelled with γ-^{32}P-ATP using f4 polynucleotide kinase. The probe was precipitated three times at -20^{0}C in 2 M ammonium acetate and 10 volumes of cold ethanol. After precipitation, the probe was resuspended in 100 µl of 0.3 M sodium acetate.

11-2. Isolation of Total Cellular RNA

The total cellular RNA was isolated using acid guanidinium thiocyanate-phenol-chloroform extraction modified by the method from Chomczynski and Sacchi [1987]. The strains carrying the plasmids pMES22 or pGU2 were grown in minimal medium. 40 ml of mid-log phase cells were chilled in ice, harvested by centrifugation, and then resuspended in 5 ml of denaturing solution which contained 4 M guanidinium thiocyanate, 25 mM sodium citrate (pH 7), 0.5% sarcosyl, and 0.1 M 2-mercaptoethanol. Sequentially, 0.5 ml of 2 M sodium acetate (pH 4), 5 ml of water saturated phenol, and 1 ml of chloroform-isoamyl alcohol mixture (24:1) were added. The mixture was vortexed for 10 seconds and cooled on ice for 15 min. After centrifugation at 10,000 rpm for 20 min at 4^0C, the aqueous phase which contained RNA was transferred to a fresh tube, mixed with 1 ml of cold isopropanol, and RNA precipitated at -20^0C for more than 1 hour. After centrifugation (10,000 rpm for 20 min), the RNA pellet was dissolved in 0.6 ml of denaturing solution, transferred to a 1.5-ml Eppendorf tube, and precipitated with 0.6 ml of isopropanol at -20^0C for more than 1 hour. After centrifugation, the RNA pellet was washed with 75% ethanol and dissolved in 100 µl Of 0.5% SDS.

11-3. RNA Primer Extension

Primer extension was performed on the RNA following the method of Gellebter modified from Current Protocols [Ausubel *et al.*, 1987] as suggested to us by Karin Everett. To do

this, 1-5 ng of labelled primer was mixed with 10 µg of RNA in a volume of 12 µl which contained 250 mM KCl and 10 mM Tris, pH8.3. Primer was annealed to RNA by heating in a heat block at 80°C for 2 min, and slow cooling at 30°C. After 4 hours, the annealed mixture was placed at 42°C. 12 ml of reverse transcriptase buffer was added, which contained 24 mM Tris (pH 8.3), 16 mM $MgCl_2$, 8 mM dithiothreitol, 0.7 mg/ml actinomycin D, 0.4 mM each dATP, dCTP, and dTTP, 0.8 mM dGTP, and 20 units of AMV reverse transcriptase. The reaction mixture was incubated at 42°C or 47°C for 90 min. 1 µl each of 0.5 M EDTA, and 1mg/ml RNase was added and the mixture was incubated for 10 min at 37°C. The mixture was treated by adding 100 µl of 2.5 M ammonium acetate and extracting with 100 µl phenol/chloroform/isoamyl alcohol (24:24:1). The aqueous phase was transferred to a fresh tube and the cDNA extension product was precipitated with 300 µl ethanol at -20°C. After centrifugation, the pellet was resuspended in 6 µl formamide loading buffer. Half of this sample was loaded onto an 8% polyacrylamide-8 M urea sequencing gel. A sequencing ladder labelled with ^{35}S and primed with the same oligonucleotide was run in adjacent lanes.

PART 12. OTHER GENETIC METHODS

12-1. Plasmid Isolation

The plasmid "miniprep" isolations were made according to the methods of Maniatis *et al.* [Maniatis *et al.*, 1982]. The large amount of plasmid required for gel retardation and DNase

I footprints was isolated by a method from F. Lang's lab(personal communication) with slight modification. Cells from a 100 ml culture of plasmid-containing cells were harvested by centrifugation and resuspended in 20 ml of Triton-mix buffer (0.4% Triton-100, 100 mM Tris pH8.0, 100 mM EDTA) in a 250 ml flask, then 1 ml of 10mg/ml lysozyme was added. The flask with the cell suspension was put in a boiling water-bath for 1-2 minutes. After that, the samples were poured into centrifuge tubes and centrifuged at 10000 rpm for 15 minutes. The supernatant was transferred to centrifuge tubes and precipitated with 2-3 volumes of 0.5 M ammonium acetate in ethanol. The DNA pellets were then resuspended in 0.3 M sodium acetate to precipitate proteins, and treated with 3.3 M LiCl to precipitate RNA, and with RNase, proteinase K digestion, and phenol and chloroform extraction. Finally, plasmid DNA was precipitated with ethanol and resuspended in TE buffer.

12-2. Chromosomal DNA and λ DNA Isolation

Chromosomal DNA isolation was carried out as described by Silhavy *et al.* [Silhavy *et al.*, 1984] and λ DNA was isolated according to the methods of Maniatis *et al.* [Maniatis *et al.*, 1982].

12-3. Transformation

Transformations were performed according to Maniatis *et al.* [Maniatis *et al.*, 1982].

12-4. Transduction and Conjugation for Mapping

Initial mapping of *lrp*::Tn*10* and some of the λp*lac*Mu insertion mutations was carried out by conjugation with Hfr strains from Singer *et al.* [Singer *et al.*, 1989]. Spontaneously-occurring streptomycin-resistant or chloramphenicol-resistant transduced from MEW28 derivatives of these strains were used as recipients. The protocol used was from Miller [Miller, 1972].

More accurate mapping was done by P1-mediated transduction also according to the method described by Miller [Miller, 1972]) using Singer's mapping kit, a collection of strains carrying Tn10 insertions at convenient locations (12-22 min. for *lrp*, 62-64 min. for *serA* and GCV, 60-65 min for CP8, CP41 CP52, and CP57, 70- min for CP36, and 91-95 min for CP61) [Singer *et al.*, 1989].

RESULTS

This thesis is devoted to the definition and characterization of the leucine/Lrp regulon of *E. coli*. The experiments are reported in two sections. In Part A, I will describe the experiments that defined the regulon, and characterized it as a metabolic and regulatory entity. In Part B, I will present molecular experiments concerning the interaction of Lrp with the regulatory sequences of some genes it controls.

PART A. *IN VIVO* STUDIES OF LEUCINE/LRP REGULON

It has been a surprise of *E. coli* metabolism that synthesis of many diverse gene products is regulated according to the availability of exogenous L-leucine [Newman *et al.*, 1976]. In this part of the thesis, I will describe the isolation and characterization of an *E. coli* mutant which, as the result of a single mutation, has lost the leucine-mediated control on the expression of several of these leucine-regulated genes. The genes so affected are now considered to be members of the leucine regulon.

We originally used the name *rbl* (regulation by leucine) for the gene in which this pleiotropic mutation occurred. This has now been renamed *lrp* (leucine-responsive regulatory protein) according to a consensus of investigators in this field. I show here that the *rbl*//*lrp* gene product is the regulator of global response to leucine.

1. Isolation and Characterization of an *rbl*::Tn*10* Mutant

I began this study by trying to isolate an insertion mutation in a regulatory gene affecting L-SD synthesis. Under the assumption that a regulatory mutation which increased L-SD activity would allow the strain to grow on L-serine, I selected tetracycline-resistant strains for those which could grow with L-serine as carbon source, as described in the following section.

1-1. Isolation of Mutants Using L-Serine as Sole Carbon Source by Tn*10* Insertion

Wild-type *E. coli* is unable to use L-serine as sole carbon source unless L-leucine and/or glycine are also present in the medium. Two classes of mutations which allow the cell to grow on L-serine alone have been isolated and/or described in this lab. One class of mutants acquired defects in methionine metabolism along with the ability to grow with L-serine as carbon source. Surprisingly, they did not show increased L-SD activity [Brown *et al.*, 1990]. In this class are *metJ* mutants which grew on L-serine by derepression of methionine biosynthetic enzymes, and *gos* mutants which could grow on L-serine by an unknown metabolic event causing a conditional requirement for L-methionine. A second class of mutants, *ssd*, could use L-serine as sole carbon source by overproduction of the enzyme L-SD.

Since the expression of L-SD is induced by L-leucine and glycine, and increased in some mutants [Newman *et al.*, 1982b],

one would expect the gene coding for L-SD to be subject to negative control. If this were true, the destruction of the gene encoding the *sdaA* repressor would result in the overproduction of L-SD and a mutant with an ineffective repressor would be able to grow on L-serine as sole carbon source. To determine whether it was possible to isolate an L-serine utilizer by a null mutation, I selected mutants growing on L-serine from a pool of MEW1 cells carrying random insertions of transposon Tn*10* elements on the chromosome [see materials and methods]. The Tn*10* insertion from one such isolate was transduced back into strain MEW1 via generalized transduction mediated by P1 phage, selecting for tetracycline resistance, and verified that the transductants could grow on L-serine. This procedure showed that the insertion mutation itself was sufficient to permit growth on L-serine. This mutation was named *rbl*-1::Tn*10*, for regulation by leucine. One *rbl*::Tn*10* transductant, strain MEW26, was kept for further study.

I-2. Genetic Locus of the *rbl* Gene

To determine the chromosomal location of this mutation, a spontaneous streptomycin-resistant derivative of strain MEW26 was isolated and mated with the various strains from the Hfr kit of Singer *et al*. [Singer *et al*., 1989]. Recombinants resistant to both kanamycin (carried by the Hfr strains) and streptomycin were selected and scored for loss of the *rbl* phenotype (i.e. loss of tetracycline resistance and of the

ability to grow on the L-serine). 95% of the recombinants with Hfr strain CAG12203 (Tn*10*kan at 12 min) and 67% of those with Hf$^-$ strain CAG12206 (Tn*10*kan at 17 min) lost the *rbl* mutation, This result suggested that the *rbl* gene was located on the *E. coli* chromosome between 12-21 min.

The *rbl* gene was located more precisely by P1-mediated transduction. In P1 transduction with the donor strain CAG18528, *rbl* was found to be 15 to 20% linked to Tn*10*kan, inserted at 20 min. (72 of 412 transductants scored). The mutation was also shown to be 40 to 50% cotransducible (67 of 147 transductants scored) with *serC* at 20.1 min. Thus, the *rbl* gene is located near 20 min. This map location is obviously different from that of previously described mutations permitting growth on L-serine: *ssd* at 88 min [Newman *et al.*, 1982b], *metJ* at 89 min and *gos* at 90 min [Brown *et al.*, 1990]. I concluded therefore that this mutation identified a new gene governing L-SD expression.

I-3. Physiological Characterization of Strain MEW26

The *rbl*::Tn*10* mutation of strain MEW26 is a very pleiotropic mutation. It confers the ability to grow with L-serine as the sole carbon source, and it also results in many other changes in cell metabolism. The characteristics of the mutant strain include a decreased growth rate, an inability to grow in minimal medium at elevated temperature, inefficient use of exogenous L-serine source and loss of the ability to use glycine as nitrogen and L-serine source.

1-3-1. Decreased Growth Rate of Strain MEW26 in Minimal-Glucose Medium

As a preliminary assessment of the physiological state of strain MEW26, I compared its apparent doubling time (a.d.t.) with that of its parent strain MEW1 (Table 2). The *rbl* mutant MEW26 grew more slowly than its parent strain MEW1 in glucose minimal medium,(a.d.t. 77 min compared with 56 min for strain MEW1). The addition of either L-leucine or L-serine increased the growth rate considerably. The presence of 100 µg/ml of L-leucine in glucose minimal medium had essentially no effect on the growth rate of the wild-type strain but increased that of the *rbl* mutant (a.d.t. 67 min.) This result suggested that L-leucine biosynthesis might be limited in the mutant, which was later shown to be true. The addition of L-serine to glucose-minimal medium also increased the growth rate of *rbl* mutant (a.d.t. 69 min.). The effects of L-leucine and L-serine were somewhat additive: a combination of L-leucine and L-serine decreased the mutant a.d.t. to 64 min but had no effect on that of wild-type strain.

1-3-2. Strain MEW26 Required L-Serine for Growth at Elevated Temperature

While strain MEW26 grew more slowly than its parent at 37^0C, it could not grow at all in glucose minimal medium at 42^0C, conditions in which the wild type strain grew well (Table 3). The mutant grew well at the higher temperature if also supplied with L-serine (100 µg/ml), suggesting that it

TABLE 2. Doubling times of the *rbl*::Tn*10* mutant and its parent strain MEW1 during exponential growth at 37°C[a].

Addition to glucose minimal medium[b]	Doubling time (min) of strain[c]	
	MEW1 (*rbl*$^+$)	MEW26 (*rbl*$^-$)
none	56	77.5
L-leucine	56	67
L-serine	55	69
L-leucine + L-serine	56	64

a: Values are averages of at least three different experiments.

b: Both L-leucine and L-serine were added at 100 µg/ml final concentration.

c: Strains were grown in glucose-minimal medium overnight, subcultured in the media to be tested for 3-4 hours to allow adaptation to these conditions and assure that the cells were in the exponential growth phase, and then subcultured, each in the medium used for subculture in sidearm flasks equipped for determining turbidity with a Klett colorimeter. Turbidity was measured at 30 minute intervals using a blue (#42)filter. The doubling time was calculated from semi-log plots of turbidity as a function of incubation time. Data presented are the average of three different determinations.

was deficient in L-serine at both temperatures, but more seriously deficient at the high temperature. If this were true, it might be that a *serA* gene cloned in a high-copy number plasmid could increase L-serine biosynthesis and restore growth of the mutant at 42°C. I transformed plasmid pGU17 (pBR325 carrying *serA*$^+$ gene, kindly provided by G. A. Grant) into strain MEW26 *rbl*::Tn*10* and in fact this strain with pserA grew well at 42°C. This is consistent with the idea that the *rbl* mutation decreases synthesis of phosphoglycerate dehydrogenase, the serA gene product, and the first enzyme in the L-serine biosynthetic pathway.

A deficiency in L-serine might be due to increased degradation rather than (or in addition to) decreased synthesis. Temperature-sensitive growth might have resulted from increased L-SD activity, the more so since L-SD activity is known to be temperature-inducible, and was increased when cells were grown at elevated temperature. However strains which expressed much higher L-SD activity - MEW1 carrying plasmid pMES22 *sdaA* and KEC9 *ssd*- grew well at 42°C (Table 3), and also at 37°C. This demonstrates that high L-SD activity did not suffice to starve the cell for L-serine.

I-3-3. Inefficient Use of Exogenous L-Serine

Nonetheless, high levels of L-SD would be expected to lead to less efficient use of exogenous serine, since the conversion of serine to pyruvate is not reversible. This was described earlier for an *ssd serA* strain which had a much higher requirement for L-serine than the parent serA strain.

TABLE 3. Growth and L-SD activity[a] of the *rbl* mutant and related strains.

Strain	Plasmid	Relevant genotype	37°C growth[c]	37°C L-SD	42°C growth	42°C L-SD	42°C + L-serine[b] growth	42°C + L-serine[b] L-SD
MEW1	None	Wild-type	+	10	+	29	+	32
MEW1	pMES22	MEW1 psdaA	+	150	+	505	+	507
MEW26	None	*rbl*	+	64	-	NA[d]	+	96
MEW26	pMES22	*rbl* psdaA	+	590	-	NA	+	480
MEW26	pGT17	*rbl* pserA	+	64	+	87	+	ND[e]
KEC9	None	*ssd*	+	214	+	ND	+	ND

a: Cells were grown in glucose-minimal medium, in one case with L-serine 500 µg/ml added, at the temperatures noted above, subcultured, and assayed in exponential phase. L-SD activity was determined in whole cells and is expressed as milliunits of activity catalyzed by 0.1 ml of a 100 K.U. suspension of cells in 35 min, 1 unit of enzyme is the amount of enzyme which synthesized 1 µmol of pyruvate in the assay. Values are averages of three different experiments. Antibiotics was added to the cultures of strains carrying plasmids.

b: L-Serine was added at final concentration of 100 µg/ml.

c: Growth was measured as follows: overnight cultures grown at 37 C in glucose minimal medium were diluted 1000 times into test medium and incubated for 36 hours at the temperature noted. + represents good growth within 18 hours, and - represents the absence of turbidity after 36 hours.

d: NA, Not available; cells do not grow in this condition.

e: ND, Not determined.

Since the *rbl* mutation also resulted in overproduction of L-SD, one would expect to see the same inefficient use of L-serine. To investigate this, I constructed strains MEW37 (*serA*::Mu d1 *ssd*) and MEW38 (*serA*::Mu d1 *rbl*::Tn*10*) and compared the yield per unit serine of these two strains and their parent strain DRN1 (*serA*::Mu d1). In fact the high L-SD was accompanied in both cases by a decreased yield per serine (9.5 and 6.6 ug protein/ug serine in the mutants compared : 14.4 in DRN1; Table 4). It was surprising that the yield of the *ssd* mutant was higher than that of the *rbl* mutant, since the *ssd* mutant has much higher L-SD activity. No explanation can be given at this time, other than the suggestion that some other aspect of the *rbl* mutant phenotype may also be involved in the inefficient use of L-serine.

I-3-4. The rbl Mutant Cannot Use Glycine as L-Serine Precursor or as Nitrogen Source

A *serA* mutant can grow in glucose-minimal medium with either exogenous serine or exogenous glycine. It converts some of the glycine to C1-THF via the glycine cleavage enzymes (*gcv* product), and condenses this with other glycine molecules to form L-serine via the *glyA* gene product.

The *rbl serA* double mutant (MEW38) differs from its parent in that it can use exogenous serine but not exogenous glycine to compensate for the *serA* mutation. This deficiency is not likely to result from high L-SD activity, because an *ssd serA* double mutant (MEW37) grew well with glycine as L-serine source.

TABLE 4. The yield of cell material obtained by a *serA* mutant and its derivatives from L-serine[a].

Strain	Relevant genotype	L-serine (µg/ml)	Turbidity (K.U.)	Protein (µg/ml)	Protein/µg L-serine (µg)
DRN-1	*serA*	40	59	576	14.4
		80	109	1090	13.6
MEW37	*serA ssd*	40	31	380	9.5
		80	62	736	9.2
MEW38	*serA rbl*	40	27	264	6.6
		80	48	528	6.6

a: The strains were grown in glucose minimal medium with 200 µg/ml of L-serine overnight, centrifuged, washed twice with minimal medium without glucose and L-serine (NIV), diluted 100 fold in NIV. 0.1 ml each of the diluted suspension was used to inoculate 20 ml cultures with the serine concentrations noted. Cells were grown to constant turbidity, water added to 20 ml to compensate for evaporation, and the turbidity measured using Klett colorimeter with green (#54) filter. 9 ml samples of the final cultures were precipitated with trichloracetic acid, and digested with NaOH, and the amount of protein assayed, as described in the Materials and Methods. The values given are the averages of three determinations.

This physiological deficiency is much like that caused by a *gcv* mutation. The *gcv* mutant would be unable to cleave glycine to C1-THF, and would therefore be unable to use glycine as either L-serine or nitrogen source. In fact, strain MEW26 *rbl*::Tn*10* was also unable to grow in glucose minimal medium with glycine as nitrogen source while the parent strain MEW1 grew well. The *metJ serA* double mutant showed the same physiological deficiency. This suggests the possibility that both *rbl* and *metJ* gene products may directly or indirectly be involved in the regulation of *gcv* expression.

II. The Expression of *sdaA* and *tdh* is Regulated by the *rbl* Gene Product

In this section I show that the activity of two amino acid-degrading enzymes, L-SD and TDH, is regulated by the *rbl* gene product. However, the *rbl* gene is not a general regulator of amino acid-degrading enzymes since other amino acid-degrading enzymes are not affected in *rbl* mutant.

II-1. Increased L-SD Activity in the *rbl* Mutant

The previously described *ssd* mutant KEC9, selected for its ability to use L-serine as carbon source, was shown to have an increased L-SD enzyme activity [Newman *et al.*, 1982b]. When grown in minimal glucose medium, the *rbl*::Tn*10* mutation strain MEW26 also produced more L-SD than the parent strain MEW1, but less than the *ssd* mutant KEC9 (Table 5). This presumably accounts for the ability of the mutant to use L-serine as carbon source, since an increase in L-SD activity has been

shown to be sufficient to permit the growth of *E. coli* strain on L-serine [Su *et al.*, 1989]. In this respect, the *rbl* mutant is different from the *gos-3*, *gos-5*, and *metJ* mutants, which do not overproduce L-SD though they do grow with L-serine [Brown *et al.*, 1990].

Formation of active L-SD requires posttranslational activation [Newman *et al.*,1985]. A regulatory system might act by affecting either gene expression or posttranslational steps. However, the increased activity observed in the *rbl* and *ssd* mutation strains reflects an increased rate of expression of the *sdaA* gene, the structural gene which codes for L-SD [Su, 1992]. This effect of *rbl* and *ssd* on *sdaA* expression was shown by means of an *sdaA::lacZ* protein fusion, constructed by insertion of λp*lac*Mu9 in frame into the *sdaA* coding sequence, putting the structural gene for ß-galactosidase under the control of the *sdaA* promoter [Su *et al.*, 1989]. ß-galactosidase activity was increased sevenfold in the *sdaA::lacZ rbl::Tn10* strain MEW36 and ninefold in the *sdaA3::lacZ ssd* strain MEW35 (Table 5). It is thus clear that the *rbl* and *ssd* mutations cause a significant increase in *sdaA* expression.

As previously described, L-SD activity in the wild-type strain was induced during growth in the presence of L-leucine. The four-fold induction factor is about half that due to the *rbl* mutation. These effects were not additive; in the *rbl* mutation strain MEW26, the presence of L-leucine stimulated the L-SD activity no more than 1.1-fold (Table 6). L-Leucine exerted similar effects on the expression of *lacZ* from *sdaA*

TABLE 5. Synthesis of ß-galactosidase from the *sdaA* promoter in the *rbl* mutant and related *E. coli* K-12 strains[a]

Strain	Relevant genotype	L-SD activity[b]	ß-galactosidase from the *sdaA* promoter[c]
MEW1	Wild type	10	NA[d]
MEW26	*rbl*	96	NA
KEC9	*ssd*	214[e]	NA
MEW22	*sdaA::lacZ*	ND[f]	50
MEW36	MEW22 *rbl*	ND	350
MEW35	MEW22 *ssd*	ND	450

a: Assays were done on mid-exponential-phase cultures grown in glucose minimal medium at 37°C. Values are averages of three different experiments.

b: Expression as in Table 3.

c: Reported in Miller units.

d: NA, Not available.

e: Value taken from Table 3.

f: ND, Not determined.

TABLE 6. The effect of amino acids on synthesis from the *sdaA* promoter[a]

Addition to glucose minimal medium[b] (relevant genotype)	L-SD activity of ß-galactosidase activity MEW1 Wild-type	MEW26 *rbl*	MEW22 *sdaA::lacZ*	MEW36 MEW22 *rbl*
None	10	68	73	423
Glycine	20	81	201	558
Glycine + L-leucine	84	95	507	645
L-leucine	42	70	268	474
L-Alanine	24	72	191	428

a: Values are averages of two or three different experiments and expressed as in Table 5.

b: The amino acids were added at the following concentrations: 300 µg/ml for glycine and L-leucine, 500 µg/ml for L-alanine.

promoter: 4-fold induction in the *rbl*$^{+}$ strain, compared with no more than 1.2-fold further induction in the *rbl* mutant (Table 6). Table 6 also shows that the presence of L-alanine or glycine also could induce the expression of *sdaA* gene, and stimulated both L-SD activity and ß-galactosidase from *sdaA* promoter to 2- to 2.4-fold. In the *rbl* mutant, the addition of glycine still induced the expression from *sdaA* promoter, whereas the presence of L-alanine had only a slight effect (Table 6). This may indicate that L-alanine acts through the *rbl* gene product but glycine does not.

II-2. Increased TDH Activity in the *rbl* and *ssd* Mutants

If the *rbl* mutation affects regulation by L-leucine, other enzymes regulated by this amino acid might also show altered activity in the *rbl* mutant. L-threonine dehydrogenase (TDH), which catalyses the first step in a pathway of glycine biosynthesis from threonine, has been shown to be induced by the addition of L-leucine to the growth medium [Newman *et al.*, 1976]. Since the *rbl* mutation seemed to mimic the effect of exogenous L-leucine on L-SD activity, I examined its effect on TDH activity. The TDH activity in the *rbl* mutant MEW26 was 7 to 8-fold higher than that in the parent strain MEW1 when grown in glucose minimal medium (Table 7). The addition of L-leucine to the growth medium stimulated the wild-type level threefold but had essentially no effect on the TDH level in the *rbl* mutant. It is clear then that leucine regulates TDH via the *rbl* gene product (Table 7).

Since the *ssd* mutant is highly pleiotropic [Newman *et al.*,

1981] and affects L-SD synthesis, I examined the possibility that this mutation also affects TDH activity. As shown in Table 7, TDH activity was increased fourfold in the *ssd* mutant. Although the effects were quantitatively different, both L-SD and TDH activities were regulated by *rbl* and *ssd* gene products. The L-SD activity was higher in the *ssd* mutant than in the *rbl* mutant; whereas the fourfold increase of TDH activity in the *ssd* mutant was about half of the increase resulting from the *rbl* mutation.

To determine what other similarity there might be in the regulation of these two amino acid- degrading enzymes, TDH activity was assayed in cells grown in the presence of other amino acids which were known to increase L-SD activity. The results show that the addition of glycine to the growth medium had no effect on TDH activity in wild-type strain but decreased the level of TDH in the *rbl* mutant. On the other hand, the presence of L-alanine in the medium resulted in the TDH level increasing 2.5-fold in the wild-type strain but decreasing slightly in the *rbl* mutant. Since the expression of another *rbl* regulated gene, *lysU*, was also stimulated by both L-leucine and L-alanine, it is possible that regulation by L-alanine may be mediated by the *rbl* gene product, but this question has not been settled in this work.

Newman et *al*. showed that strain JEV73R, a mutant strain which was able to derive its glycine from endogenously synthesized L-threonine, has constitutively increased levels of TDH in minimal medium [Newman *et al*., 1976]. This increase in TDH level might be caused by a mutation in the gene coding

TABLE 7. TDH activity of different strains grown in glucose minimal medium with and without addition of amino acids[a]

Strain	Relevant genotype	TDH activity in minimal medium with[b] None	L-leucine	Glycine	L-alanine
MEW1	Wild type	5.1	17.5	5	17
MEW26	*rbl*	40.5	44	27	30
KEC9	*ssd*	24.5	ND[c]	ND	ND

a: Assays were done on mid-exponential-phase cultures grown in glucose minimal medium at 37°C. Values are averages of at least three different experiments. TDH activity was expressed as nanomoles of aminoacetone formed by 0.3 ml of a 500-Klett-unit suspension of cells in 20 min as described in Materials and Methods.

b: The amino acids were added at same concentration as in Table 6.

c: ND, Not determined.

for TDH. However it could also have been due to an *rbl* mutation. To see whether strain JEV73R carries an *rbl* mutation, the *rbl* gene was cotransduced with the linked *serC*$^+$ gene from JEV73R to an MEW1 *serC* strain. Of 187 L-serine-independent transductants, 68 were able to use L-serine as sole carbon source. All three L-serine-utilizers assayed had high levels of L-SD (data not shown). This observation suggested that JEV73R harboured an *rbl* mutation, which resulted in the overproduction of TDH.

II-3. Tryptophanase and Proline Oxidase Are not Affected by the *rbl* Mutation

Since both L-SD and TDH are amino acid-degrading enzymes, it seemed possible that the *rbl* mutation might also stimulate the synthesis of other enzymes involved in amino acid catabolism. We assayed L-proline oxidase and L-tryptophanase activities in wild-type strain MEW1 and in MEW26 *rbl*::Tn*10*. The *rbl* mutation had no effect on the level of either enzyme (data not shown), excluding the hypothesis that the *rbl* gene product is a general regulator of all amino acid-degrading enzymes.

It has previously been reported that L-proline oxidase activity is not affected by the addition of L-leucine to the growth medium [Tam *et al.*, 1978]. I also found that L-tryptophanase activity was not affected by the presence of exogenous L-leucine (data not shown). These observations suggest that the *rbl* gene product is the specific mediator of L-leucine effects, but the metabolic thread connecting L-leucine effects is as yet obscure.

III. The *rbl* Mutation Decreased Transcription from the *serA* and *IlvIH* Promoters

If the regulation of leucine-sensitive genes by L-leucine is mediated by *rbl* gene product, the expression of other leucine-regulated genes should be altered in the *rbl* mutant. Both *serA* and *IlvIH*, whose gene products are involved in amino acid biosynthesis, are regulated by L-leucine and the *rbl* gene product.

III-1. Decreased *serA* Transcription in the *rbl* Mutant

The *serA* gene codes for the enzyme phosphoglycerate dehydrogenase, which is responsible for the first step specific to L-serine biosynthesis. The phosphoglycerate dehydrogenase level has also been shown to be regulated by L-leucine. However, unlike L-SD and TDH, it is expressed at lower levels when L-leucine is added to the growth medium [McKitrick and Pizer, 1980]. To examine the effect of the *rbl* mutation on *serA* transcription, the *rbl*::Tn*10* allele was transduced into a *serA-lacZ* operon fusion strain, DRN1. β-galactosidase activity was measured in cells grown in the presence and absence of L-leucine (Table 8). The data demonstrated a fivefold decrease of *lacZ* expression from the *serA* promoter in the *rbl* mutant. Exogenous L-leucine repressed *serA-lacZ* expression twofold in the wild-type strain but had essentially no effect in the *rbl* mutant. These results indicated that the transcription of *serA* is under the positive control of Lrp.

The *ssd* mutation had no effect on transcription from the *serA* promoter (Table 8). This observation suggested that the

TABLE 8. Regulation of *serA-lacZ* expression

strain	Relevant genotype	ß-galactosidase[a]	
		- L-leucine	+ L-leucine[b]
DRN-1	*serA::lacZ*	2350	1360
MEW38	*serA::lacZ rbl*	450	500
MEW37	*serA::lacZ ssd*	2250	ND[c]
MEW40	*serA::lacZ rbl* *sdaA::*Cm^r	425	ND

a: Assays were carried out on log-phase cultures grown in glucose minimal medium with 100 µg/ml L-serine at 37^{0}C. The data in this Table is the average of two or three different experiments and reported in Miller units.

b: L-Leucine was added at the concentration of 100 µg/ml.

c: ND, Not determined.

ssd and *rbl* gene products belong to different, although possibly overlapping, regulatory systems.

Since the *rbl*::Tn*10* mutant was selected by its ability to grow on L-serine and increased L-serine degradation, it seemed possible that the decrease of transcription from *serA* promoter might be an indirect result of high L-SD activity. However I constructed a triple mutant in which the *sdaA* gene was inactivated and no L-SD could be made, strain MEW40, *serA*::*lacZ rbl*::Tn*10 sdaA*::Cm^r. In this strain, transcription of *lacZ* from the *serA* promoter was also fivefold decreased (Table 8). This data demonstrates that the decrease of *serA* transcription is not a secondary effect resulting from L-SD overproduction.

III-2. Decreased *IlvIH* Transcription in the *rbl* Mutant

The *IlvIH* operon encodes the two subunits of acetolactate synthase III, an enzyme involved in the biosynthesis of branched-chain amino acids [Squires et *al.*, 1983]. Expression of IlvIH is also affected by L-leucine. A Northern blot analysis demonstrated that the addition of L-leucine to growth medium resulted in an 8- to 10-fold decrease in the synthesis of *IlvIH* mRNA [Squires et *al.*, 1981].

To study the effect of the *rbl* mutation on *IlvIH* transcription, the *rbl*::Tn*10* mutation was transduced into strain CV975, an *IlvIH*::*lacZ* fusion strain kindly provided by J.M. Calvo. The *rbl* mutation resulted in a 27-fold decrease in the level of transcription from the *IlvIH* promoter (Table 9).

TABLE 9. Regulation by L-leucine in the *rbl* mutant[a]

Activity assayed	Strain	Relevant genotype	Activity in minimal medium no addition (A)	with L-leucine (B)	effect of L-leucine B/A X 100
L-SD[b]	MEW1	*rbl*$^+$	11	46	418
	MEW26	MEW1 *rbl*	67	73	109
TDH[c]	MEW1	*rbl*$^+$	5	18	305
	MEW26	MEW1 *rbl*	36	43	112
ß-galactosidase					
from *sdaA*[d]	MEW22	*sdaA::lacZ*	56	225	402
	MEW36	MEW22 *rbl*	425	530	125
serA[d]	DRN-1	*serA::lacZ*	2350	1360	58
	MEW38	DRN-1 *rbl*	450	500	113
ilvIH[e]	CV975	*ilvIH::lacZ*	322	38	12
	MEW42	CV975 *rbl*	9	9	100

a: Values are averages of at least two different experiments.

b: Expressed as in Table 3.

c: Expressed as in Table 7.

d: Expressed as in Table 5.

e: Cells were grown in glucose minimal medium with 50 µg/ml L-proline and 1µg/ml thiamine. Expressed as in Table 5.

As summarized in Table 9, four leucine-regulated genes (*serA*, *sdaA*, *tdh* and *IlvIH*) whose products are involved in different metabolic pathways are all regulated by the *rbl* gene product. In the *rbl* mutant, the expression from these leucine-regulated promoters was no longer affected by the presence of L-leucine. That is, in the absence of the rbl gene product, L-leucine does not affect expression of these genes. Indeed, the mutant behaves as though L-leucine were constantly present in the growth medium. Based on these results, we suggested that the *rbl* gene product might be the regulator of a global response to L-leucine.

IV. Effects of *rbl* Mutations in a *metK* Strain

This section shows that the *rbl* mutation suppressed the slow growth of a *metK* mutant, and that the most commonly used *metK* strain carried an *rbl* mutation.

IV-1. A True *metK* Mutant Grows Slowly in Glucose Minimal Medium

Strain RG62, a *metK* mutant deficient in S-adenosylmethionine (SAM) synthetase activity, was isolated by Greene *et al.* by virtue of its ethionine-resistance [Greene *et al.*, 1973]. This strain shown less than 3% residual SAM synthetase activity and high levels of methionine biosynthetic enzymes and was resistant to γ-glutamyl methyl ester (GGME). It was reported recently that strain RG62 had an unusual ability to degrade exogenously supplied L-serine [Matthews and Neidhardt, 1989], and they suggested that the strain might have high L-SD activity. R. Matthews kindly provided us with

TABLE 10. L-SD, ß-galactosidase activities and GGME resistance of *rbl* and *metK* mutants and related strains of *E. coli* K-12

Strain	Relevant genotype	L-SD activity[a]	GGME resistance[b]	ß-galactosidase from promoter of[c]	
				sdaA	*serA*
RG	wild-type	30	-		
RG62	*metK62 rbl-62*	71	+		
MEW1	Wild-type	11[d]	-		
MEW26	*rbl*::Tn*10*	67[d]	-		
MEW33	*rbl-62*	66	-		
MEW30	*metK62*	11	+		
MEW31	*metK62 rbl*::Tn*10*	65	+		
MEW22	*sdaA*::*lacZ*			56[d]	
MEW36	*sdaA*::*lacZ rbl*			425	
MEW46	*sdaA*::*lacZ rbl-62*			385	
DRN-1	*serA*::*lacZ*				2350
MEW38	*serA*::*lacZ rbl*				450
MEW39	*serA*::*lacZ rbl-62*				386

a: Expressed as in Table 3. Values are averages of three different assays.

b: As judged by growth on glucose minimal medium containing 1 mM GGME. +, Resistance; -, sensitivity.

c: Expressed as in Table 5.

d: values taken from Table 9.

her isolate of strain RG62. It did indeed have increased L-SD levels: 71 mU/100 Klett units versus 30 for the *metK*$^+$ parent strain RG (Table 10). Furthermore, strain RG62 was able to use L-serine as sole carbon source, whereas the parent strain RG could not.

Since another *metK* mutant in this lab did not show overproduction of L-SD and could not grow with L-serine as carbon source (data not shown), it seemed possible that these properties of RG62 might be due not to the *metK* mutation but rather to an unsuspected *rbl*- or *ssd*-like mutation present in the strain. To test this possibility, the *metK* mutation from RG62 was transduced into our strain background, by cotransduction with *serA*$^+$ to
recipient strain DRN-1. Of the transductants, 10 to 20% were resistant to GGME, consistent with the established genetic map distance between *serA* and *metK* [Bachmann,1983].

One such *metK62* transductant, MEW30, was characterized in more detail. It showed some of the characteristics of the original *metK* strain: GGME resistance as stated above, and high ß-cystathionase activity (the product of the *metC* gene), as expected from a *metK* strain, which is derepressed for the methionine biosynthetic genes (data not shown). However, unlike RG62, it showed low L-SD activity (Table 10) and was unable to grow with L-serine as sole carbon source. Furthermore, MEW30 grew exceedingly slowly in glucose minimal medium. The fastest growth rate measured in MEW30 (a.d.t. 120 min) was only half that of the parent strain, but the growth rate was variable and usually even slower (Table 11). On the

other hand, RG62 had a doubling time of 80 min, only slightly longer than that of the wild-type strain. The growth rate of MEW30 was considerably increased by the presence of exogenous L-leucine in the growth medium (a.d.t. 89 min) (Table 11); This might be related to the observation that exogenous L-leucine i eased the residual SAM synthetase activity three- to four-fold in strain RG62 [Greene *et al.*, 1973].

IV-2. Demonstration that Strain RG62 Carries an *rbl* Mutation

As indicated above, when we transferred the *metK* gene out of strain RG62, the ability to use L-serine as carbon source did not accompany it. It seems that strain RG62 must carry a second mutation responsible for L-serine utilization. I located this mutation by conjugation and transduction, and then moved the mutation to our laboratory reference strain MEW1 and confirmed that the resultant strain showed the *rbl* phenotype.

Conjugation studies with the Singer Hfr kit located the second mutation responsible for L-SD overproduction in RG62 between 12 min and 21 min; It was highly linked to Tn*10* elements at 12 min (44/50 exconjugants with strain CAG5053) and at 17 min (24/49 exconjugants with strain CAG5051), consistent with it being in the *rbl* gene.

The mutation was more accurately mapped by P1-mediated transduction, and located at the same map position as *rbl*::Tn*10*, 50 to 70% cotransducible with a Tn*10* element at 20 min (strain CAG18478). I then used the Tn*10* element inserted next to the mutation in the preceding experiment to transfer

TABLE 11. The growth rate of the *metK* mutant and related strains in glucose minimal medium with and without L-leucine

Strain	Relevant genotype	a.d.t. in glucose minimal medium with	
		no addition	L-leucine[b]
RG62	*metK-62 rbl-62*	80	ND[c]
MEW1	Wild-type	56	56
MEW26	*rbl*::Tn*10*	77.5	67
MEW30	MEW1 *metK-62*	142	89
MEW31	*metK-62 rbl*::Tn*10*	81	80
MEW32	MEW31 *sdaA*::Cm^r	82	ND[c]

a: A.d.t. was measured as described in Table 2. Values are averages of more than two different experiments.

b: L-Leucine was added at a concentration of 100 µg/ml.

c: ND, Not determined.

the mutation to strain MEW1, screening tetracycline resistant derivatives of MEW1 for strains which could grow on serine. These transductants had high L-SD activity but were not resistant to GGME (Table 10), and they had low cystathionase activity (data not shown).

The preceding experiments confirm that strain RG62 carried two mutations, one mapping in the *rbl* gene, and one in *metK*. Because of the physiological interest of this finding, I verified it further by transducing the mutation at 20 minutes into the *sdaA*::*lacZ* fusion strain MEW22 and *serA*::*lacZ* fusion strain DRN-1. The RG62-derived mutation caused a six-fold increase in *sdaA* expression and a six-fold decrease in *serA* transcription, similar to results for the *sdaA*::*lacZ* *rbl*::Tn*10* strain MEW36 and *serA*::*lacZ* *rbl*::Tn*10* strain MEW38 (Table 10). I conclude that RG62 harbours a second, heretofore unsuspected, mutation, an *rbl* allele, and suggest for it the allele designation *rbl-62*.

The fact that exogenous L-leucine increased the growth rate of the *metK-62* transductant, strain MEW30, suggested that the relatively fast growth rate of RG62 might be due to its *rbl-62* mutation. To test this possibility, the *rbl*::Tn*10* mutation from strain MEW26 was transduced into a newly constructed, slow-growing *metK-62* strain, selecting for tetracycline resistance on LB plates and screening for growth with L-serine. In glucose minimal medium, the *metk* *rbl* double mutant grew at the same rate as RG62 (a.d.t. 80-min Table 11), and much faster than the *metK-62* parent strain. This growth rate was about the same as that of the *metK-62* strain growing

in the presence of exogenous L-leucine (Table 11). The same experiment was performed with the *rbl-62* mutation from RG62, and with a mutation previously known in J. M. Calvo's lab as *lhb*::Tn*10* mutation from their strain CV1008, and recognized by them as an allele of *rbl*. These gave the same acceleration of growth rate in the *metK* background (data not shown).

I wondered which of the pleiotropic effects of the *rbl* mutation might be involved in increasing the *metK* growth rate. The compensation for slow growth of the *metK* mutant did not depend on a functional L-SD enzyme. As shown in Table 11, a *metK rbl*::Tn*10 sdaA*::Cm^r triple mutant grew as rapidly as the *metK rbl*::Tn*10* parent strain (and much faster than the *metK* single mutant), indicating that the *rbl* mutation could compensate for the slow growth of the *metK* in the total absence of L-SD activity.

Together with the observation that the *serA*::*lacZ rbl*::Tn*10 sdaA*::Cm^r also decreased *serA* transcription, it could be conclude that these phenotypes are not secondary effects resulting from L-SD overproduction in *rbl* strains.

IV-3. During Growth in Glucose-Minimal Medium, *metK* Mutants Accumulated Derivatives Able to Grow on L-Serine

Strain RG62, like most *metK* strains described, was isolated by resistance to a methionine analog, in this case ethionine [Greene *et al.*, 1973]. It grew a little more slowly than the parent strain but was not severely handicapped, especially as used in later studies [Matthews and Neidhardt, 1988; 1989]. Since the same *metK-62* allele transduced into

another strain still conferred slow growth, it was possible that the secondary *rbl-62* mutation had arisen early during culture of the original *metK* isolate and that the double mutant had been selected by virtue of its faster growth rate such that most published studies used the double mutant inadvertently.

To determine whether this type of event could be reproduced, the *metK-62* mutation was transduced into a newly isolated single colony of strain DRN-1 *serA*, a serine-independent GGME-resistant transductant (*serA*$^{+}$ GGME^{r}) was isolated, and grown in glucose minimal medium, and the number of cells able to grow with L-serine as sole carbon source were periodically tested. Whereas the first transductant could not use serine as carbon source, by the fourth subculture in glucose-minimal medium (about 30 generations), 19% of the cells could grow on L-serine (Table 12). No such accumulation of L-serine utilizers was observed during growth of the parent strain MEW1. It could be concluded that an *rbl*-like mutation(s), permitting growth on L-serine, accumulated during cultivation of a *metK* mutation strain in glucose minimal medium- i.e. that a mutation which increased the growth rate in glucose minimal medium also conferred ability to grow with serine.

Although many rapidly growing *metK* derivatives were able to grow with L-serine as carbon source, others were not. This suggest that there is more than one way to restore rapid growth to *metK* strains.

TABLE 12. Selection of cells growing on L-serine during growth of a *metK* mutant[a] in glucose-minimal medium

Step	Culture	No. of cells forming colonies on minimal medium with given carbon source		% L-Serine utilizers, B/A X 100
		Glucose (A)	L-Serine (B)	
a	Inoculum	1.4×10^{8}	5×10^{2}	4×10^{-4}
b	Culture 1	6.4×10^{7}	1.5×10^{4}	2.5×10^{-2}
c	Culture 2	8×10^{8}	2.4×10^{6}	0.3
d	Culture 3	1.6×10^{9}	1.6×10^{7}	1.0
e	Culture 4	7.5×10^{8}	1.4×10^{8}	19

a: A recently transduced isolate of strain MEW30 was plated on glucose and L-serine minimal medium plates (step a), inoculated into liquid glucose minimal medium at 37°C, grown overnight, and plated on the same two medium (step b). It was then subcultured into fresh glucose minimal medium (culture 2) at 0.1 ml/20 ml of culture, and that culture was incubated overnight, plated (step c), and subcultured. The experiment was continued in the same way for cultures 3 and 4. Each subculture allowed approximately eight generations of growth. Columns A and B show the number of cells in 1 ml of each culture at the time of subculture, as judged by the colony counts on glucose and L-serine minimal media, respectively. Glucose and L-serine was provided at 0.2%. The results given are those of one experiment; two others gave similar results.

V. *rbl* Is Identical to *lhb* and Might Be Identical to *oppl*

There are several reports of regulatory genes located near 20 min of *E. coli* chromosome whose function is mediated by leucine. These genes, including *rbl*, *lhb*, *oppl*, and *llvR*, might all be alleles of the same gene. In this section I show that the *rbl* gene is in fact identical to *lhb*.

V-1. *rbl* and *lhb* Have Similar Phenotypic Effects

J.M. Calvo and his collaborators have described a protein that binds to the regulatory region of the *E. coli llvIH* operon [Ricca *et al.*, 1989] and mediates its repression by L-leucine. They have also isolated a mutant, *lhb*::Tn*10*, in which this protein is absent [Platko *et al.*, 1990]. The *lhb* mutant was located near 20 min and caused decreased expression of the *llvIH* operon. J. M. Calvo suggested to us that our instances of L-leucine regulation might be modulated by the same protein. In this case, the *rbl* gene might be identical to *lhb*, the tructural gene for the *llvIH* binding protein.

To test this, I transduced *lhb*::Tn*10* mutation from CV1008 (provided by J. M. Calvo) into our strains, and examined the effect of the *lhb* and *rbl* mutations on TDH and L-SD activity and on the expression of *lacZ* from *sdaA*, *serA* and *llvIH* promoters. The results show that the two mutations had quantitatively similar effects in all cases: increasing L-SD, TDH activity, and *sdaA* expression and decreasing the expression from *serA* and *llvIH* promoters (Table 13). Furthermore, the strains carrying *lhb* mutations (MEW43 and CV1008) were able to grow with L-serine as sole carbon source, whereas parent strains (MEW1 and CV975) could not.

TABLE 13. Comparison of strains carrying *rbi* and *lhb* mutations[a]

Strain	Relevant genotype	Activity of β-galactosidase from promoter of[d] L-SD[b]	TDH[c]	*sdaA*[e]	*serA*[f]	*IlvIH*[g]
MEW1	Wild-type	10	5.1			
MEW26	*rbi*	96	40			
MEW43	*lhb*	93	42			
MEW22	*sdaA::lacZ*			50		
MEW36	MEW22 *rbi*			350		
MEW44	MEW22 *lhb*			375		
DRN-1	*serA::lacZ*	14			2350	
MEW38	DRN-1 *rbi*	83			375	
MEW41	DRN-1 *lhb*	84			375	
CV975	*IlvIH::lacZ*	13				275
MEW42	CV975 *rbi*	51				10
CV1008	CV975 *lhb*	8				10

a: Assayed in exponential-cells grown in glucose minimal medium (with additions as needed). Values are averages of at least two different experiments.

b: Expressed as in Table 3.

c: Expressed as in Table 7.

d: Expressed as in Table 5.

e: Expressed as units from an *sdaA::lacZ* fusion.

f: Expression as units from a *serA::lacZ* fusion.

g: Expression as units from an *IlvIH::lacZ* fusion.

Another regulatory gene, *oppI*, is also known to be located near 20 min on the *E. coli* chromosome. The *oppI* gene product was shown to be a negative regulator of *oppABCDF* operon expression [Austin *et al.*, 1989]. It has been reported that the addition of L-leucine to the growth medium resulted in a dramatic increase in the expression of the *opp* operon-encoded transport system [Andrews *et al.*, 1986]. The *oppI* mutation- or L-leucine-mediated elevation of peptide transport was accompanied by increased sensitivity to toxic tripeptides. To test the relationship between *rbI* and *oppI*, the sensitivity of our wild-type and *rbI* mutant strains to tripeptides triornithine (Orn_3) was determined. Fifty µg Orn_3 on a 7-mm disk resulted in a 26-mm zone of growth inhibition of strain MEW1 *IlvA*$^+$ grown on glucose-minimal medium; the addition of L-leucine to the medium increased the inhibition zone to 31-mm. In the same conditions, MEW26 *Ilv*$^+$ (*rbI*) showed a larger inhibition zone without leucine (33 mm). This was not further increased by the presence of exogenous L-leucine. This observation is consistent with the idea that the *rbI* gene is identical to *oppI*.

V-2. Subcloning of the *rbI* Gene from the Kohara Phage Collection.

To study it in more detail, I chose to subclone the *rbI* gene from the Kohara phage which carried DNA in the region defined by transduction as the locus of *rbI*. Thus DNA was isolated from Kohara phage λ1F10, digested with *Hin*dIII, and

inserted onto plasmid PACYC184. A 6.5 Kb *Hin*dIII fragment containing the *rbl* gene was cloned by using a *serA rbl* double mutant, which cannot use glycine as a source of L-serine, and selecting for transformants which could grow on glucose minimal medium with glycine. From this plasmid, which was termed pLR1, a 1.3 Kb fragment was subcloned first onto Bluescript KS^+ and then onto pACYC184, yielding plasmid pLRN1, which was used for complementation tests.

The putative *rbl*-carrying clones were all isolated by their ability to use glycine as a source of L-serine. The *rbl* mutant cannot cleave glycine to C-1 THF source. If a plasmid in fact carried a wild-type rbl^+ gene, it might also be expected to restore the expression of *sdaA* and *serA* of an *rbl* mutant to wild-type levels. To test this, plasmid pLRN1 was transformed into strains MEW26, MEW36, and MEW38. As shown in Table 14, plasmid pLRN1 complemented the *rbl* mutation with respect to both *sdaA* and *serA* expression.

V-3. Hybridization of a 2.5 Kb Fragment from pLR1 to Chromosomal DNA Isolated from Strains with an Insertion in *rbl*::Tn*10* or *lhb*::Tn*10*

It is clear that plasmid pLRN1 can complement the regulatory deficiencies caused by the *rbl* mutation. However this does not demonstrate that the *rbl* gene was in fact carried on plasmid pLRN1, but shows only that the plasmid carries some gene which complements an *rbl* deficiency.

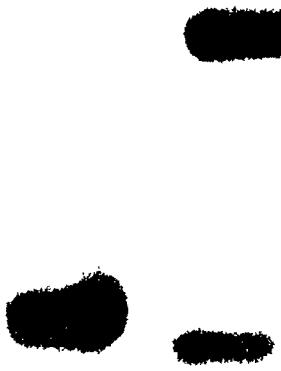
TABLE 14. L-SD and ß-galactosidase activities in strains with plasmids carrying the putative *rbl* gene[a]

Strain	Plasmid	L-SD activity[b]	ß-galactosidase from promoter of	
			sdaA[c]	*serA*[c]
MEW1	None	11		
MEW26	None	76		
MEW26	pLRN1	9		
MEW22	None		65	
MEW36	None		420	
MEW36	pLRN1		66	
DRN-1	None	14		2387
MEW38	None	95		430
MEW38	pLRN1	15		2312

a: Cells were grown in glucose minimal medium (with additions as needed) at 37°C, and assayed in exponential-phase. Antibiotics was added to plasmid-carrying cells during both overnight and subculture as listed in material and methods. Values are averages of two or three different measurements.

b: Expressed as in Table 3.

c: Expressed as in Table 5.



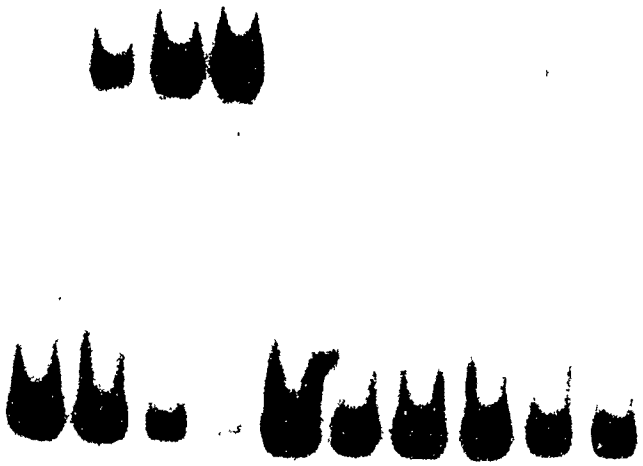
1 2 3

Figure 1. Southern blot analysis of chromosomal DNA from mutants MEW26 and CV1008. Chromosomal DNA was isolated from strains MEW1, MEW26 and CV1008, digested with *Eco*RI and *Hin*dIII, and fractionated by electrophoresis through agarose gel. The gel was dried and hybridized with ^{32}P-labelled *Eco*RI-*Hin*dIII fragment derived from plasmid pLR1. Lanes 1 through 3 represent the hybridization of chromosomal DNA from strain MEW1, MEW26, and CV1008, respectively.

To determine whether pLR1 carried the *rbl*/*Ihb* gene, a 2.5 Kb *Eco*RI-*Hin*dIII fragment from pLR1 was used to probe *Eco*RI-*Hin*dIII digests of chromosomal DNA isolated from wild-type strain MEW1, *rbl*::Tn*10* mutation strain MEW26, and *Ihb*::Tn*10* mutant CV1008 (Fig. 1). The Southern blot analysis showed that the digest of DNA from strain MEW1 contained only one band hybridizing to pLR1 i.e. that the gene carried on pLR1 was intact in strain MEW1. This band (2.5 Kb) was replaced in digests of DNA from both mutants by two other bands. It is clear then that the gene carried on pLR1 is split by the insertions into two parts, i.e. that pLR1 carries the wild type version of the gene mutated in strains MEW26 and CV1008, that is, the *rbl* gene located in this region.

V-4. protein Encoded by pLRN1 Has DNA Binding Activity

The *Ihb* gene product has been shown to be a DNA binding protein [Ricca *et al.*, 1989]. If *rbl* is identical to *Ihb*, then its product should be the same DNA binding activity. Then extracts of cells grown with plasmid pLRN1 should contain the binding protein. To test this, the DNA fragment carrying *sdaA* promoter region was used for binding assay, and the binding activity in crude extracts prepared from strains MEW1, MEW26, and MEW26 carrying plasmid pLRN1 was compared with purified IHB protein (Fig. 2). The extracts of MEW1 and MEW26 carrying pLRN1 showed a pattern of band shifts similar to that of purified IHB protein, though the plasmid-containing strain MEW26 had much higher activity than that of MEW1. In all cases addition of L-leucine to the binding buffer prevented retardation. Consistent with the idea that strain MEW26 has a



1 2 3 4 5 6 7 8 9 10

Figure 2. Gel retardation assays performed with *sdaA* DNA. A 349 bp end labelled *Hin*dIII-*Eco*RV fragment from pSU1 was incubated with cell extracts of MEW1, MEW26, or MEW26 pLRN1, or with purified Lrp protein, and fractionated by electrophoresis. Lanes 1 through 5 represents the addition of 0, 17, 33, 66, and 66 ng of purified Lrp protein; 1 µl cell extract from MEW26, 1 µl extract from MEW1, 0.25, 0.5, 0.5 µl extract from MEW26 pLRN1 were added to lanes 6 through 10, respectively. Total protein concentrations (µg/µl) in each extract were as follows: MEW1, 8.5; MEW26, 8.1; MEW26 pLRN1, 16.2. Lanes 5 and 10 also incubated with 38 mM L-leucine.

nonfunctional *rbl* gene, the extract of strain MEW26 did not show detectable binding activity. The observations described in sections V-3 and V-4 allow the conclusion that the *rbl* gene is identical to the *lhb* gene to be made.

Since the *rbl/lhb/oppl* gene product is involved in the regulation of a number of operons with different functions, and is now studied in several labs, it seemed better to agree on a new name which would apply to all of these. The researchers working on this locus agreed that the locus mapping near *aroA* and *serC* that affects expression of *sdaA*, *serA*, *tdh*, and *IlvIH* is assigned the designation *lrp* (leucine-responsive regulatory protein). From the next section, I will therefore call this gene *lrp* instead of *rbl*.

VI. Isolation and Preliminary Characterization of λp*lac*Mu Inserts in Genes Controlled by the *lrp* Gene Product

As described above, the regulation by L-leucine of several genes from different metabolic pathways is mediated by Lrp protein. There is no reason to suppose that this includes all the Lrp regulated genes. In the following experiments, I assess the possibility that other L-leucine-regulated genes might also be regulated by *lrp* gene product, characterize the patterns of regulation of the various genes, and attempt to assess the size of the leucine regulon.

For this purpose, I took advantage of the fact that λp*lac*Mu9 inserts randomly in the chromosome, creating protein

fusions of *lacZ* to various genes. This involved first isolating *lacZ* translational fusions to genes whose expression are affected by the presence of exogenous L-leucine; and then making the corresponding *lrp* derivatives by P1 transduction (see materials and methods) and measuring ß-galactosidase activities of cells grown in glucose-minimal medium with and without L-leucine 100 µg/ml. In all of λp*lac*Mu inserts isolated whose expression was affected by exogenous L-leucine, the ß-galactosidase levels were also regulated by Lrp protein.

The expression of *lacZ* from these promoters was regulated by L-leucine, and by the *lrp* gene product, but in several different ways (Table 15). In some cases (class 1), Lrp activated gene expression and L-leucine decreased activity, a pattern described above for *serA* and *IlvIH*. In others (class 2), Lrp repressed and L-leucine increased activity, as in *sdaA* and *lysU*. In class 3, both the *lrp* gene product and L-leucine were required for induction, and in class 4, both were needed to repress. Most of these genes have not yet been identified, nor is it known how many different genes are represented. However, all were assayed for L-SD and showed the same activity as the parent strain.

Among the 22 prototrophic isolates studied, I found 11 of class 1, 6 of class 2, 4 of class 3 and only 1 of class 4.

In class 1, the presence of Lrp increased the expression of ß-galactosidase (in the absence of L-leucine) from 4- to 35-fold. In most of the strains, the absence of Lrp had a much greater effect on gene expression than the addition of

TABLE 15. Activity of ß-galactosidase transcribed from Lrp-regulated promoters[a]

Strain	ß-Galactosidase activity in glucose minimal medium with		Effect of L-leucine (lrp^+/lrp^-) in	Effect of Lrp (lrp^+/lrp^-) in	
	None(A)	L-leucine(B)	B/A X 100	A	B
_____Class 1. Lrp activates, L-leucine relieves					
CP4	532	232	44	35.1	15.5
CP4 *lrp*	16	15	94		
CP5	362	21	5.8	22.6	1.4
CP5 *lrp*	16	15	94		
CP7	151	26	17	5	0.8
CP7 *lrp*	30	34	113		
CP8	498	242	49	19.2	9.7
CP8 *lrp*	26	25	96		
CP10	494	190	38	19.8	5.3
CP10 *lrp*	35	36	103		
CP14	457	197	43	18.3	7.3
CP14 *lrp*	25	27	108		
CP26	379	135	36	23.7	7.5
CP26 *lrp*	16	18	113		
CP27	426	144	34	32.8	12
CP27 *lrp*	13	12	92		

Cont.

CP28	265	144	54	14.7	9
CP28 *lrp*	18	16	89		
CP30	172	141	82	11.5	9.4
CO30 *lrp*	15	15	100		
CP31	911	377	41	26.8	10.5
CP31 *lrp*	34	36	106		
Class 2. Lrp represses, L-leucine relieves					
CP21	52	96	185	0.45	0.75
CP21 *lrp*	115	128	111		
CP23	28	117	418	0.08	0.21
CP23 *lrp*	538	556	103		
CP25	35	114	326	0.52	1.56
CP25 *lrp*	67	73	109		
CP59	19	63	332	0.17	0.52
CP59 *lrp*	114	121	106		
CP60	21	86	409	0.17	0.62
CP60 *lrp*	127	139	109		
CP61	30	80	270	0.02	0.07
CP61 *lrp*	1239	1184	96		

Cont.

Class 3. Lrp and L-leucine activate together					
CP1	92	229	248	1.26	2.4
CP1 *lrp*	73	71	97		
CP41	170	355	209	0.99	2.29
CP41 *lrp*	172	155	90		
CP52	35	114	326	0.97	3.26
CP52 *lrp*	36	35	97		
CP57	38	114	300	0.9	2.92
CP57 *lrp*	42	39	93		
Class 4. Lrp and L-leucine repress together					
CP36	496	7	1.4	0.96	0.02
CP36 *lrp*	515	443	86		

a: Cells were grown in glucose minimal medium either lacking or containing 100 µg/ml L-leucine at 37°C, and subcultured in the same medium. The ß-galactosidase was assayed in cells from exponential-phase cultures and expressed as in Table 5. Values are averages of two or three different experiments.

L-leucine to the growth medium of the wild-type, as described for *serA* and *ilvIH* above. This observation suggests that the *lrp* gene product is needed for the expression of these genes, and that L-leucine modulates expression of the activated genes, probably by virtue of its interaction with Lrp, but that in most cases it does not totally reverse it, as it would if it removed Lrp from the DNA.

The repressive effect of Lrp in class 2 is also greater than the effect of L-leucine in most isolates, as is the case in *sdaA*. The Lrp repressed expression from 2- to 40-fold, while the presence of L-leucine only increased the expression to from 1.8- to 4-fold. The *in vivo* data suggested that two conformations of Lrp, free or bound with L-leucine, could each be active. It is possible that there are some genes for which one Lrp conformation acts as activator and the other acts as repressor. In such a case, one would expect that the expression level in an *lrp* mutant would be intermediate between those in the lrp^+ strain in the presence and absence of L-leucine. In strain CP25, the presence of exogenous L-leucine indeed had a greater effect than the *lrp* mutation. This might suggest that in the absence of L-leucine, Lrp acts as a repressor of this gene. However, the Lrp-leucine complex could stimulate the expression of the strain CP25 inserted gene. In both class 3 and 4, neither Lrp nor L-leucine alone could affect the gene expression. Furthermore, L-leucine had no effect on gene expression in *lrp* mutants of all 4 classes.

VII. Identification of Some Lrp Regulated-Genes

In the last section I described that the inserts whose expression was affected by Lrp showed four modes of regulation. These mutations were isolated by a selection based on regulatory function, as was done in the isolation of *din* genes. However this in itself gives no information as to the identity of the genes affected.

In order to understand the physiological function of Lrp, I tried to identify some of the Lrp-regulated genes. These experiments are described in the following paragraphs, which also include the demonstration that the expression of L-leucine-biosynthetic enzyme(s) also falls under the control of Lrp.

VII-1. Insertions in Some Lrp-Regulated Genes Leads to L-Leucine Auxotrophy

The system used for isolating insertion mutants in Lrp-regulated genes would not allow the isolation of auxotrophs with the exception of those requiring L-leucine. I did isolate several leucine-requiring strains. This is consistent with the idea that L-leucine biosynthesis seemed to be limited in the *lrp* mutant, as judged by the fact that the addition of L-leucine increased the growth rate of the *lrp* mutant.

Some of the L-leucine auxotrophs were studied further. As shown in Table 16, the expression of ß-galactosidase from the 5 insert-carrying strains studied was regulated by the *lrp* gene product (Table 16). In all of the cases, the *lrp* mutation

reduced the expression of ß-galactosidase in cells grown with sufficient L-leucine for growth (50 µg/ml). Lrp stimulated expression between 2.6 and 12-fold in the 5 strains.

All of the strains were mapped using the Singer kit, and showed linkage to a Tn*10* element at 2 min (strain CAG12095): 82% linkage for strain CP66, and 97-100% for the other 4 strains. It seems likely that all 5 strains carry insertions in the L-leucine biosynthetic operon at 1.8 min.

Because strain MEW1 *ilvA* requires L-isoleucine for growth, each of these mutations were transduced into an *ilvA*$^{+}$ derivative of MEW1 to allow testing of the ability of the auxotrophs to grow in the absence of L-isoleucine. In the presence of only L-leucine in glucose- minimal medium, four strains (CP53, 54, 55, and 64) grew well, but one (CP66) grew slowly. Based on the low level of stimulation seen in strain CP66, its lower linkage to the insert at 2 min., and its difference in growth requirements, it is possible that strain CP66 constitutes a separate phenotypic class of leucine auxotrophs.

If these mutants were inserted in L-leucine biosynthetic genes, as the results suggest, one would expect them to be derepressed in L-leucine-starved cells. To test this, cells were grown with either 5 or 50 µg/ml of L-leucine overnight, subcultured in the same medium, and assayed ß-galactosidase. In 4 similar mutants, L-leucine starvation resulted in a 4- to 23-fold increase in ß-galactosidase activity, and it did so both in the *lrp* mutants and in the *lrp+* strains (Table 16).

TABLE 16. Regulation of gene expression in L-leucine auxotrophs

Strain	ß-Galactosidase activity in glucose minimal medium with L-leucine of 5 µg/ml (A)	50 µg/ml (B)	A/B	Effect of starvation A	Effect of Lrp (*lrp*$^+$/*lrp*$^-$) on B
CP53	3806	688	5.5	4.2	8.1
CP53 *lrp*	900	85	10.6		
CP54	3179	610	5.2	2.8	10
CP54 *lrp*	1126	61	18.5		
CP55	4449	831	5.4	2.6	11.2
CP55 *lrp*	1707	74	23.1		
CP64	3325	685	4.9	2.7	12.2
CP64 *lrp*	1251	56	22.3		
CP66	392	774	0.5	7.3	2.6
CP66 *lrp*	54	302	0.18		

Results expressed as in Table 15. Values are the average of 2 or 3 determinations.

This confirms that these mutants do carry insertions in one or more of the L-leucine biosynthetic genes.

The expression of these L-leucine-biosynthetic genes is clearly activated by Lrp, as judged by these results. It is also controlled independently by L-leucine. The effects of L-leucine starvation and the *lrp* gene product were additive. This is the first case of L-leucine affecting expression in an *lrp* mutant, and probably reflects the transcriptional attenuation mechanism by which the *leu* operon is regulated [Wessler and Calvo, 1981].

The regulation of strain CP66 was very different. ß-Galactosidase activity of CP66 was actually repressed by L-leucine starvation, both in the parent strain and in the *lrp* mutant. The *lrp* gene product increased expression in the presence of excess L-leucine, and mitigated the repressive effect of L-leucine considerably during L-leucine starvation. This is consistent with the idea that strain CP66 differs from the other 4, and perhaps suggest that the insertion may even be in a different, though closely linked, gene.

VII-2. Location of Some Inserts Near the *sdaB* Gene

Three of our class 3 inserts, including strain CP41, CP52, and CP57, mapped very close to the inserts in strain CAG12079 (*fuc*::Tn*10*) and strain CAG12135 (*recD*::Tn*10*), showing linkage of 79, 94, and 85% to *fuc*::Tn*10*, and 50, 80 and 67% to *recD*::Tn*10*, respectively. This strongly suggests that the inserts are in or near the *sdaB* gene which is immediately adjacent to the fucose operon on the *E. coli* chromosome (Shao,

Z. Q., personal communication), and codes for a second L-serine deaminase expressed mainly in rich medium [Su and Newman, 1991].

To test whether these insertions were located within the *sdaB* gene, ß-galactosidase activity was assayed in cells grown in three media: in glucose minimal medium where the expression of *sdaB* is repressed; in glucose minimal medium with L-leucine where it is slightly induced; and in LB medium where *sdaB* is expressed at its highest level [Su and Newman, 1991]. In these strains, ß-galactosidase activity followed exactly that pattern, 150, 350, and 2000 in CP41, and 35, 125 and 350 in CP52.

sdaB function in the insertion strains was tested by constructing *sdaA*::Cm^r derivatives. If these insertions in fact inactivated *sdaB*, the double mutants (*sdaA*::cm^r *sdaB*::λp*lac*Mu) should have neither of the L-SD coding genes, and thus should be
almost completely deficient in L-serine deaminating ability when grown in LB medium. As shown in Table 17, the CP41 and CP52 insertions decreased the L-SD levels in LB medium 5- and 10-fold, respectively.

The preceding data shows clearly that L-SD activity from *sdaB* is low in both strains CP41 and CP52. Nonetheless, hybridization experiments showed that there is no insertion in *sdaB* in either strain. When a 2.6 Kb DNA fragment from pWB6, carrying the intact *sdaB* gene, was hybridized with DNA from CP41, CP52, and the parent strain MEW1, only one band was seen, at the same position (2.6 Kb in size) in each strain

TABLE 17. The ß-galactosidase and L-SD activities in CP41, CP52, and related strains[a]

Strain	L-SD activity in LB	ß-Galactosidase activity in MM[b]	MM + L-leucine[b]	LB
MEW1 *sdaA*::Cm^r	63	ND[c]	ND	ND
CP41 *sdaA*::Cm^r	6	160	374	1482
CP52 *sdaA*::Cm^r	13			
CP41		161	352	1925
CP41 *lrp*		157[d]	147[d]	1499
CP52		35	114	384
CP52 *lrp*		37	35	378

a: L-SD and ß-galactosidase were assayed and expressed as in Table 3 and Table 5, respectively. The values are averages of two or three different determinations.

b: MM, glucose minimal medium. L-Leucine was added at 100 µg/ml.

c: ND, Not determined.

d: Values taken from Table 15.

(data not shown). But when an 8 Kb fragment from the same plasmid was used as probe, two hybridized bands could be seen in both CP41 and CP52, while only one band (8 Kb in size) was seen in MEW1 (Fig. 3). This observation indicated that the insertions in strain CP41 and CP52 are located immediately adjacent to the *sdaB* gene, but not within it.

The *lrp* mutation nonetheless largely decreased the transcription of both *sdaB*, and the gene carrying the insert. It might be suggested that there is a regulatory gene situated close to *sdaB* whose product is necessary for *sdaB* gene expression. The full expression of that gene then would require both Lrp and L-leucine. It is also possible that the regulation of *sdaB* by Lrp and L-leucine is an indirect metabolic effect.

In any case the major influence on the expression of both CP41 and CP52 is not Lrp. Lrp increased the ß-galactosidase activity in CP41 and CP52 in glucose minimal medium with L-leucine, as shown in Table 17. However the *lrp* mutation had almost no effect on the high level expression in both strains in LB medium. It can be concluded that the relatively slight L-leucine effects are mediated by Lrp, but that other effectors result in major changes in regulation of transcription in rich medium.

VII-3. Possible identification of an insert in *IlvJIK*

In the only mutant of class 4 isolated so far, strain CP36, Lrp together with L-leucine repressed expression profoundly, to about 1% of the level seen in cells grown in

1 2 3

Figure 3. Hybridization of chromosomal DNA of mutant strains CP41 and CP52 to a DNA fragment carrying *sdaB*. Chromosomal DNA from strain MEW1, CP41 and CP52 were isolated, digested with *pst*I, and electrophoresed on a 0.7% agarose gel. The gel was dried, hybridized with a ^{32}P-labelled *Pst*I fragment derived from plasmid pWB6, and subjected to radioautography. Lanes 1 through 3 represent the hybridization of chromosomal DNA from strain MEW1, CP41, and CP52, respectively.

the absence of L-leucine (Table 16). This mutation was mapped using the Singer's Hfr and transduction mapping kit and found that the wild-type allele of the insertion in strain CP36 was 73.5% cotransducible with a Tn*10* insertion at 75.5 min in strain CAG18450 (144 of 196 transductants), and 29.6% with one at 76.5 min in strain CAG18638 (61 of 206). This is consistent with the insertion being located in *IlvJ/K*, known to be regulated by Lrp, and coding for the branched chain amino acid transport operon [Quay *et al.*, 1977].

VII-4. Further Metabolic Screening of the *lrp* Mutant and the Strains Carrying Lrp-Regulated Insertions

In order to try to identify the function of the remaining genes carrying insertions regulated by Lrp. strains MEW1, MEW26 *lrp*::Tn*10*, and all the insertion strains were screened for their ability to use various carbon and nitrogen source. The *lrp* mutant could not use glycine as nitrogen source, and neither could several of the insert strains. However, several insert mutants also had difficulty using a whole group of nitrogen sources, L-serine, L-arginine, L-alanine and L-proline, which MEW1 used easily (Table 18). This suggests that the strains are deficient in reactions transferring ammonia/amino groups, and that the *lrp* gene product regulates synthesis of one or more of the enzymes for handling nitrogen obtained at low level from these amino acids. These insertion strains are found in both classes 1 and 2, and thus represent at least two different genes, and perhaps more. One of these might be *gltB*, the gene coding for glutamate synthase, since

TABLE 18. Ability of various strains to use amino acids as nitrogen source[a]

Strain	Growth on -N glucose minimal medium with				
	glycine	L-serine	L-alanine	L-proline	L-Arginine
MEW1	+	+	+	+	+
MEW26	-	+	-	-	-
CP4	-	+	+	-	-
CP8	-	-	+	-	-
CP10	-	-	-	-	-
CP14	-	-	-	-	-
CP25	-	-	+	-	-
CP26	-	-	+	-	-
CP27	-	-	+	-	-
CP30	-	-	+	-	-
CP31	-	+	+	-	-
CP67(CP68	-	+	+	+	+

a: To determine the ability of a strain to use an amino acid as nitrogen source, the ammonium sulphate of our minimal medium was replaced with the amino acid(5 mM) to be tested. Strains to be tested were streaked on plates with the amino acid indicated, and incubated at 37^{o}C for 24 hr. +: growth; -: no growth. All of our insertion strains grew well with L-aspartic acid, L-asparagine, or L-glutamine as nitrogen source. All strains other than those listed in this table could use glycine, L-serine, L-alanine, L-arginine, and L-proline as nitrogen source. This experiment was carried out three times, with similar results.

mutation in *gltB*, were originally described as being unable to use L-arginine, L-proline, or glycine as nitrogen source [Pahel *et al.*, 1978]. The *lrp* mutant grew poorly in the medium using L-alanine, L-proline, or L-arginine as sole nitrogen source. This may suggest that a functional Lrp protein is required for the expression of enzymes involving in using organic nitrogen sources.

Similarly strain CP8, in which activation by Lrp is lessened by L-leucine (class 1), was unable to use several carbon sources, D-xylose, D-ribose and rhamnose, which both the *lrp* mutant and its parent strain were able to use. Using the Singer mapping kit, the conjugation and P1-mediated transduction results showed that the wild-type allele of the gene carrying the insertion in CP8 was 47% (275/585) cotransducible with an Tn*10* element insert at 63.5 min in strain CAG12168. When the insertion mutation in CP8 was transduced into a *serA* mutant, 4 of 17 transductants were L-serine-independent. It is clear that mutation in CP8 located near 63.5 min.

These observations suggest that Lrp target genes code for products involved in a much larger array of metabolic reactions than was previously thought. This is also suggested by the following paragraph.

We have seen that L-leucine interacts with Lrp and a number of genes. Another amino acid, L-alanine, also interacts with some of the Lrp target genes. L-alanine induces synthesis of L-SD#1, and increases expression of *lacZ* from CP4, CP36, CP60, and CP61. Of these, CP61 can be tentatively identified

as carrying an insert in *lysU*. The expression of CP61 is increased at 42^{0}C, a characteristic of *lysU*, and it mapped in the same location as *lysU*. P1 transduction showed that in 82 of 89 transductants, the *lacZ* insertion was replaced by a Tn*10* element at 94.5 min in strain CAG18427. Similarly, *lacZ* was replaced in 7 of 82 transductants by a insertion at 93.75 min in strain CAG18488.

In short, the gene carrying the insert in strain CP61 is similar to lysU in that both are located in the same position, and both are known to be induced by L-alanine and high temperature. However, the expression of *lysU* is increased in a *lrp* mutant only 4- to 5-fold, while the ß-galactosidase level of strain CP61 was increased more than 40-fold in the *lrp* mutant. Whether CP61 carries an insertion in lysU cannot be decided at present.

VIII. Regulation of the Glycine Cleavage Operon by Lrp, but not by L-Leucine

In the previous section it was shown that Lrp could interact with L-leucine and the promoters it regulated in at least 4 ways. Study of the glycine cleavage (*gcv*) genes indicates that Lrp can also have drastic effects on expression of genes which are not affected by L-leucine.

VIII-1. Lrp is Required for the Expression of *gcv* Operon

Since Lrp is already known to control the synthesis of several enzymes involved in L-serine and glycine metabolism (L-SD, phosphoglycerate dehydrogenase, and L-TDH), one might expect that other enzymes of related function might also be

affected by the *lrp* gene product. The following experiments show that the expression of *gcv* genes is in fact also under the control of Lrp. To do this, I isolated insertions in the *gcv* operon, confirmed that the strains did in fact carry the appropriate insertion, and then studied the effect of L-leucine and Lrp on *gcv* expression.

Strains which carried λp*lac*Mu insertions in the *gcv* operon were isolated using a selection described earlier [Newman *et al.*, 1974], namely screening a random pool of inserts in a *serA* strain for those which were able to use L-serine but not glycine to fulfil their autotrophic requirement. The mutations were then transduced into strain MEW1 selecting antibiotic resistance, resulting in the formation of strains CP67 and CP68, which carried only the *gcv* mutation, but not *serA* (see materials and methods).

That strain CP67 indeed carried a *gcv* mutation was confirmed by the map position of the insertion, 40% linkage to *serA* (63 min) by P1 transduction, consistent with the reported location of the *gcv* operon at 62.6 min. Strain CP68 was mapped in the same way with similar results. To verify this further, plasmid pGS146 carrying the *gcv* operon was obtained from G.V. Stauffer and transformed into the *serA gcv* strains from which CP67 and CP68 were derived. This plasmid allowed the strain which was assumed to carry *serA* and *gcv* mutations to grow with glucose and glycine- confirming that the insert was indeed in *gcv*.

To determine the effect of Lrp and L-leucine on *gcv* operon expression, ß-galactosidase activity was measured in strains

TABLE 19. Regulation of gene expression in *gcv* mutants

Strain	Relevant genotype	ß-galactosidase levels in medium with	
		no addition	L-leucine (100 µg/ml)
CP67	wild-type	1675	1458
CP67 *lrp*	*lrp*::Tn*10*	68	65
CP68	Wild-type	1466	1247
CP68 *lrp*	*lrp*::Tn*10*	74	67

Results are expressed as in Table 5. Each experiment is the average of three different determinations.

CP67, CP68, and their respective *lrp* derivatives in the presence and absence of L-leucine. As shown in Table 19, the *lrp* mutation decreased transcription of *lacZ* from the *gcv* promoter to a very low level, 4% of that seen in the parent strain, whether L-leucine was present or not. The addition of L-leucine to either the parent strain or the *lrp* mutants had no significant effect on *gcv* expression (85% of the level without L-leucine). The data of Table 19 demonstrate that Lrp is absolutely required for *gcv* expression. This is the first operon identified in which Lrp regulation is unaffected by L-leucine.

VIII-2. Physiological Deficiency in Glycine Cleavage in the *lrp* Mutant

If the requirement of Lrp for the transcription of the *gcv* operon is indeed close to absolute, as the preceding experiments suggest, an *lrp* mutant (e.g. MEW26 *lrp*::Tn*10*) should have the same physiological deficiency as a *gcv* mutant. It would be unable to cleave glycine to C1-THF, and would therefore be unable to use any pathway which depends upon glycine cleavage. In particular, it could not derive either L-serine or nitrogen from glycine, since both of these require cleavage. In fact, the *lrp* mutant was deficient in both of these functions.

While the parent strain MEW1 could grow in glucose minimal medium with glycine as nitrogen source, strain MEW26 (*lrp*::Tn*10*) could not. Similarly, the *serA lrp* double mutant could grow with serine but was unable to use glycine as L-serine source. This strain could grow in glucose minimal

medium with L-serine but not with glycine, presumably due to its inability to derive C-1 units from glycine, resulting in a L-serine deficiency in glucose-glycine medium. On the other hand, its parent strain, a *serA* mutant, could grow in glucose minimal medium with either L-serine or glycine. It is clear then that the *lrp* mutation reduces *gcv* expression to a physiologically insignificant level.

A glycine-cleavage deficiency was further confirmed by showing similar growth problems in an *lrp glyA* double mutant. *E. coli* derives its C-1 units mainly from L-serine conversion to glycine via the *glyA* gene product, serine hydroxymethyltransferase (SHMT), and from glycine cleavage. An *lrp glyA* should have neither SHMT nor glycine cleavage enzymes, and so should have great difficulty in generating C-1 units.

An *lrp glyA* mutant was constructed by transducing *lrp*::Tn*10* into a MEW1 *glyA* strain. The *lrp glyA* double mutant grew well on LB plates which were supplemented with compounds whose syntheses require C-1 units: purines, L-methionine, L-histidine, and thymidine. However it grew slowly on LB plates without such supplement, or on glucose minimal medium with it, and could not grow on glucose minimal medium with glycine. It is clear therefore that Lrp is a major, physiological effector of synthesis of glycine cleavage enzymes.

VIII-3. Other Factors Regulating gcv Expression

During the assays of *gcv* expression, I noticed a great variation in the results from experiment to experiment. Further studies indicated that ß-galactosidase levels from

both *gcv::lacZ* fusions varied directly with the density of the culture. There was a gradual progression from 1150 units at low O.D.(0.155) to 2275 units in a much more dense culture (O.D. 1.089) in strain CP67 (Fig.4, curve A1, B1). This same progression was seen in the presence of L-leucine (A2) or glycine (B2). In the *lrp* mutant, there was so little activity that one could not determine whether such a pattern existed (A3, A4). It is clear then that *gcv* expression is influenced by some factor related to the growth phase, and that it may be difficult to compare the expression of *gcv* in different growth conditions without controlling for this.

Another factor regulating expression of the *gcv* genes is the availability of C-1 units from other sources, or the availability of compounds whose biosynthesis requires C-1 units [Newman and Magasanik, 1963; Newman *et al.*, 1974]. As reviewed by Stauffer [Stauffer, 1987], there is a good deal of confusion as to the effect of C-1 units on *gcv* expression. However, in these experiments, it is clear that addition of the C-1 end-products greatly decreased *gcv* expression. When strain CP67 was grown with the end-products of C-1 metabolism: adenine, guanine, L-methionine, L-histidine and thymidine, the level of ß-galactosidase fell to 35-75 units, varying little throughout the growth cycle (B5). Most of this decrease was due to the addition of purines, as indicated by the ß-galactosidase levels from 475-150 at different stages of the culture with purines (B3, B4).

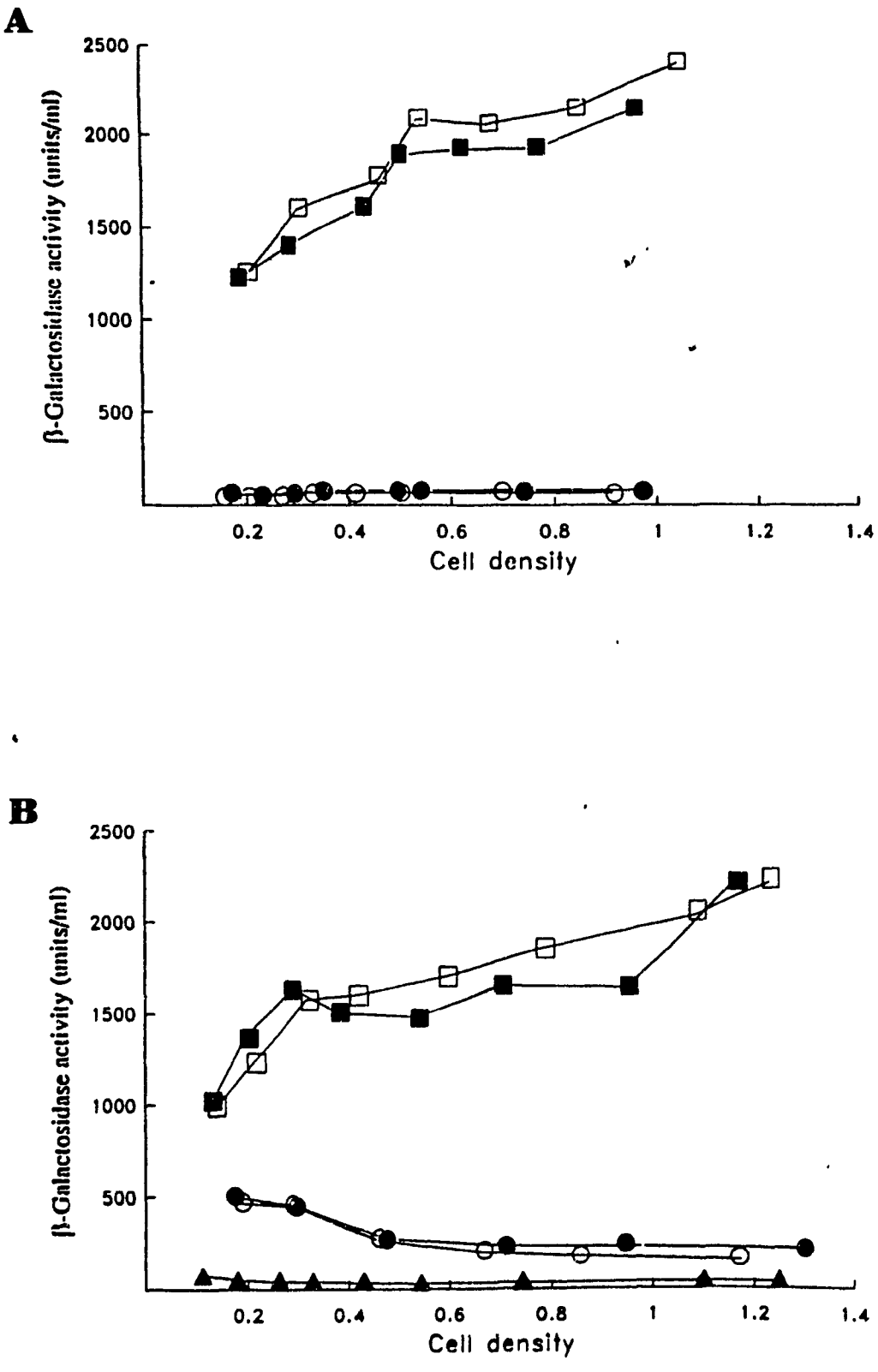


Figure 4. Regulation of *gcv* expression.

Figure 4. Regulation of *gcv* expression. Cells were grown in glucose minimal medium at 37^{0}C with the additions noted below, subcultured in the same medium, and ß-galactosidase was assayed at different points in the growth curve. Cell density is plotted in O.D. units as measured at 600 nm. Additions made were: Fig. 4A: strain CP67-- A1 (open squares), no addition; A2 (filled squares), L-leucine; Strain CP67 *lrp*--A3 (open circles), no addition; A4 (filled circles), L-leucine. Fig. 4B: Strain CP67-- B1 (open squares), no addition; B2 (filled squares), glycine; B3 (open circles), adenine; B4 (filled circles), guanine; B5 (filled triangles), adenine, guanine, L-histidine, L-methionine, and thymidine. L-Leucine, adenine, and guanine were provided at 100 µg/ml each, glycine at 400 µg/ml, and all others at 50 µg/ml.

IX. The Regulation of *lrp* Gene Expression

Since the *lrp* gene product is involved in regulating the expression of such a large and diverse array of genes, an interesting question is how the *lrp* gene itself is regulated. This section shows that the expression of *lrp* is autoregulated, Lrp decreasing its own synthesis - a neat homeostatic mechanism. However, Lrp synthesis is also affected by environmental factors, so that the cellular concentration of Lrp probably varies considerably despite this homeostatic mechanism. A variation in Lrp levels might then be expected to affect the expression of target operons.

IX-1. The Expression of *lrp* is Decreased by Lrp, and by Growth in Rich Medium but Increased in a *metJ* Mutant

To study the regulation of *lrp* gene expression, an *lrp-lacZ* translational fusion was constructed by *in vitro* methods in vector pACYC184 (pLRN2) and a mutant carrying an *lrp::lacZ* fusion in chromosome was isolated by λp*lac*Mu9 insertion (see Materials and Methods). To determine whether the expression of the *lrp* gene is autoregulated, plasmid pLRN2 was transformed into *lrp*$^+$ and *lrp* strains, and ß-galactosidase activities were measured from cells grown in glucose minimal medium with and without L-leucine (Table 20). The data indicated that a 3-fold repression of *lrp-lacZ* in the *lrp*$^+$ strain.

The preceding experiment shows that expression from a plasmid-carried lrp promoter is regulated according to the presence or absence of Lrp in the cell. I also measured

TABLE 20. Regulation of *lrp* expression[a]

Strain	Plasmid	Relevant genotype	ß-Galactosidase activity in minimal medium with none	minimal medium with L-leucine	LB
MEW45	pACYC184	*lrp::lacZ* pACYC184	1502	1507	123
MEW45	pLRN1	*lrp::lacZ* pACYC184 *lrp*$^+$	145	200	86
MEW1	pLRN2	p*lrp::lacZ*	4271	4625	858
MEW26	PLRN2	*lrp* p*lrp::lacZ*	12546	12234	838
MEW45	None	*lrp::lacZ*	1972	ND[b]	126
MEW87	None	*lrp::lacZ metJ*	3523	ND	342

a: The strains indicated were each grown at 37^{0}C in glucose minimal medium with and without L-leucine 100 µg/ml or in LB medium, with chloramphenicol 50 µg/ml for plasmid-carrying strains, subcultured in the same medium, and ß-galactosidase activity was measured and expressed as in Table 5. The values given are the average of 3 determinations.

b: ND, Not Determined.

expression from a chromosomal *lrp::lacZ* insertion. To do this, plasmids pACYC184 and pLRN1 (pACYC184 carrying lrp^+) were transformed into strain MEW45 which carried a λp*lac*Mu insertion in *lrp* gene, and ß-galactosidase activity was assayed as above (Table 20). This showed a 10-fold repression in the strain carrying the lrp^+ gene on a multi-copy plasmid.

It is clear from these experiments that *lrp* expression is autogenously regulated. On the other hand, the expression of *lrp* is essentially not affected by L-leucine. This is the first representative we have identified of the 6th class of Lrp/leucine interactions, those promoters are repressed by Lrp but unaffected by L-leucine.

A major decrease in *lrp* expression was seen in LB medium. Both the chromosomal and the plasmid-carried *lrp::lacZ* were repressed over 10-fold by growth in LB medium (Table 20). It seems then that there is a higher concentration of Lrp in cells growing in minimal medium than in those in rich medium. Even though the expression of *lrp* is autoregulated, some other regulation must override this, in order to result in this low level of Lrp in cells grown in LB medium.

Like an *lrp serA* double mutant, a *metJ serA* double mutant is also unable to derive L-serine from glycine. Since the expression of *gcv* was not repressed in this double mutant (data not shown), I wondered whether the *metJ* gene product is required for the expression of *lrp* gene. To test this, a *metJ*::Cm^r mutation was transduced into MEW45, an *lrp::lacZ* strain, and ß-galactosidase activity was assayed from cells grown in glucose minimal medium and LB medium. The data

demonstrate that the expression of *lrp-lacZ* in the *metJ* mutant was increased 2-fold in glucose minimal medium and 3-fold in LB medium (Table 20), rather than decreased as expected.

IX-2. Changes of Expression of *lacZ* from Target Insertions in LB-Grown Cells May Be Secondary to Changes in Lrp Concentration

Since the expression of *lrp* is repressed in LB medium, one would expect that the expression of Lrp-repressed genes might be increased in LB medium, while that of Lrp-activated genes might be decreased. This is true for most of the class 1 and class 2 mutation strains (Table 21). If the effect of LB on gene expression is due to the low level of Lrp, introduction of a plasmid carrying lrp^+ should at least partly restore Lrp regulation.

To examine this, the plasmids pACYC184 and pLRN1 were transformed into strains carrying target insertions, and ß-galactosidase activity assayed in cells grown with and without plasmids in LB medium. As shown in Table 21, plasmid pLRN1 increased ß-galactosidase levels of class 1 mutants but decreased that of class 2 strains.

It is clear then that the decreased Lrp concentration in LB is one of the major factors determining the expression of target genes in LB. Since Lrp increases transcription from *serA* and *IlvIH* promoters (class 1) and decreases the activity of biodegradative enzymes L-SD and TDH (class 2), it is possible that Lrp is a mechanism for controlling a switch from rich environments to the poor environments.

TABLE 21. Regulation of gene expression in LB medium by Lrp[a]

Strain	ß-galactosidase activity in strain carrying plasmid			
	none	pACYC184 (A)	pLRN1 (B)	B/A X 100
class 1	Activated by Lrp; decreased by L-leucine.			
CP4	29	16	113	706
CP8	40	19	131	689
CP10	24	15	92	613
CP14	25	10	88	880
CP26	18	11	124	1127
CP27	12	11	78	709
CP28	12	13	61	469
CP30	18	15	81	540
CP31	40	19	85	447
class 2	Repressed by Lrp; induced by L-leucine.			
CP23	328	136	19	14
CP59	ND[b]	52	10	19
CP60	ND	50	13	26
CP61	216	102	14	14

a: ß-Galactosidase activity was measured in exponential-phase cultures grown in LB medium with antibiotics appropriate to the particular plasmids at 37°C and expressed as in Table 5. Values are averages of two or three different experiments.

b: ND, Not Determined.

IX-3. The Similarities between *lrp* and *asnC*

Lrp is a basic DNA-binding protein composed of two subunits of molecular weight 18.8 kDa each [Willins *et al.*, 1991]. A search of Genbank indicated that the sequence of *lrp* gene did not resemble any of the known sequences of *E. coli* DNA-binding proteins except AsnC, the *asnC* gene product [Kolling and Lother, 1985]. The amino acid sequences of Lrp and AsnC shown 24% identity. Furthermore, Lrp and AsnC each act as both repressor and activator. The *asnC* gene product stimulates the expression of *asnA* (coding for asparagine synthetase A) but it inhibits its own synthesis, i.e. is autogenously regulated like lrp. Stimulation of transcription from the *asnA* gene by AsnC is abolished by asparagine, while the autoregulation of *asnC* is not affected by this amino acid [Kolling and Lother, 1985].

In view of these similarities between asnC and Lrp, I wished to determine whether AsnC has any physiological relation to Lrp. I therefore transformed plasmid pLSK35-3, (a kind gift of W. Messer), which carried the $asnC^{+}$ under the control of the *tac* promoter into a *serA lrp* double mutant. If the plasmid could substitute for the lrp gene, it might overcome the physiological deficiency of glycine cleavage enzymes and restore the ability to use glycine as L-serine source. Indeed the *serA lrp/pasnC* strain did grow in glucose minimal medium with glycine.

This experiment suggests that pasnC substitutes for *lrp* in the *serA lrp* double mutant. However the experiment is not unambiguous, since one would expect the same result if pasnC

TABLE 22. The effect of p*asnC* on the expression of Lrp-regulated genes[a]

Strain	Plasmid[b]	ß-galactosidase activity in minimal medium with	
		none	IPTG (1 mM)
MEW36	none	446	ND[c]
MEW36	pLSK35-3 (p*asnC*)	433	278
CV1008	none	12	ND
CV1008	pLSK35-3 (p*asnC*)	15	32
CP67 *lrp*	pLSK92-2	52	43
CP67 *lrp*	pLSK35-3 (p*asnC*)	84	703
CP68 *lrp*	pLSK92-2	61	ND
CP68 *lrp*	pLSK35-3 (p*asnC*)	87	639
CP36 *lrp*	none	585	594
CP36 *lrp*	pLSK35-3 (p*asnC*)	501	104
CP52 *lrp*	none	35	ND
CP52 *lrp*	pLSK35-3 (p*asnC*)	42	63

a: Cells were grown in glucose minimal medium with or without IPTG overnight and subcultured in the same medium, 100 µg/ml ampicillin was added to the cultures carrying plasmids. ß-galactosidase activity was measured and expressed as in Table 5. Values are averages of two or three different experiments.

b: Plasmid pLSK35-3 contains the *asnC* gene under the control of the *tac* promoter of plasmid pJF118u [Kolling and Lother, 1985]. Plasmid pLSK92-2 as a control.

c: ND, Not Determined.

did not complement *lrp* but instead acted directly to stimulate *gcv* expression. This expectation is strengthened by the fact that this double mutant carrying a gcv^+ gene in high copy plasmid also can use glycine as L-serine source.

To distinguish between complementation of *lrp* and stimulation of *gcv* expression by pasnC, I determined the effect of pasnC on several Lrp-regulated genes (Table 22), Several of these were affected by pasnC. These results strongly suggest that AsnC is similar enough to Lrp to substitute for it. Whether this indicates a physiologically important role is not clear.

PART B. *IN VITRO* BINDING STUDIES OF LRP PROTEIN TO SOME OF PROMOTERS OF THE LEUCINE/LRP REGULON

Lrp, under the name of lhb, has been shown to bind to two regions upstream of the *IlvIH* promoter [Ricca *et al.*, 1989]. This binding stimulates transcription from the *IlvIH* operon [Willins *et al.*, 1991]. It is possible that Lrp regulates transcription initiation via protein-DNA interactions. If this were true, one would expect Lrp to bind to the upstream regions of other genes of the leucine/Lrp regulon. To test this, I investigated the ability of Lrp to bind to the upstream regions of *lysU*, *sdaA*, and *serA*.

1. *In vitro* Study of Interaction between Lrp Protein and the *lysU* Upstream Region

Lysyl-tRNA synthetase isozyme II, the *lysU* gene product, was shown to be induced in a *metK* strain, RG62 [Hirshfield *et al.*, 1981]. This induction was thought to be due to the *metK* mutation. However, the overproduction of LysU protein was not corrected when strain RG62 was transformed with a plasmid carrying a wild-type *metK* gene [Matthews and Neidhardt, 1988]. As shown in this thesis, strain RG62 carries both *lrp* and *metK* mutations. Both assays of enzyme activity [I.N. Hirshfield] and 2-D gel electrophoresis analysis [R. Matthews] indicated that the constitutive expression of the *lysU* gene seen in the *metK lrp* strain is due to the *lrp* mutation. However, measurement of the levels of protein expression does not

provide any indication of whether the regulation of *lysU* expression by Lrp is direct or indirect, a question which is addressed in the following sections.

I-1. Lrp Binds to the *lysU* Promoter Region

It has been shown that plasmid pFN121 can complement the *lysU* mutation [Clark and Neidhardt, 1990] and also responds to heat shock [R. Clark, personal communication]. It must therefore carry the regulatory and coding regions of *lysU*. This plasmid carries E. coli DNA starting from a 5'-end *Eco*RV site (128 bp upstream from translation start site, corresponding to nucleotide 188) that is just in front of the putative promoter.

To test the possibility of interaction between Lrp protein and the *lysU* regulatory region, a corresponding 497 bp *Dra*I fragment was isolated from pFN120 [kindly provided by R. Clark] including nucleotides 158-654 from the *lysU* sequence [Clark and Neidhardt, 1990]. The three fragments (382, 108, and 7 bp, respectively) produced from this *Dra*I fragment after digestion with *Sau*3A were 3'-end labelled with Klenow enzyme as described in Materials and Methods.

These fragments were incubated with purified Lrp protein (kindly provided by J. M. Calvo) followed by polyacrylamide gel electrophoresis. The 108 bp fragment, containing the putative *lysU* promoter and regulatory region (bases 50 to 158 upstream from the translation start site), migrated more

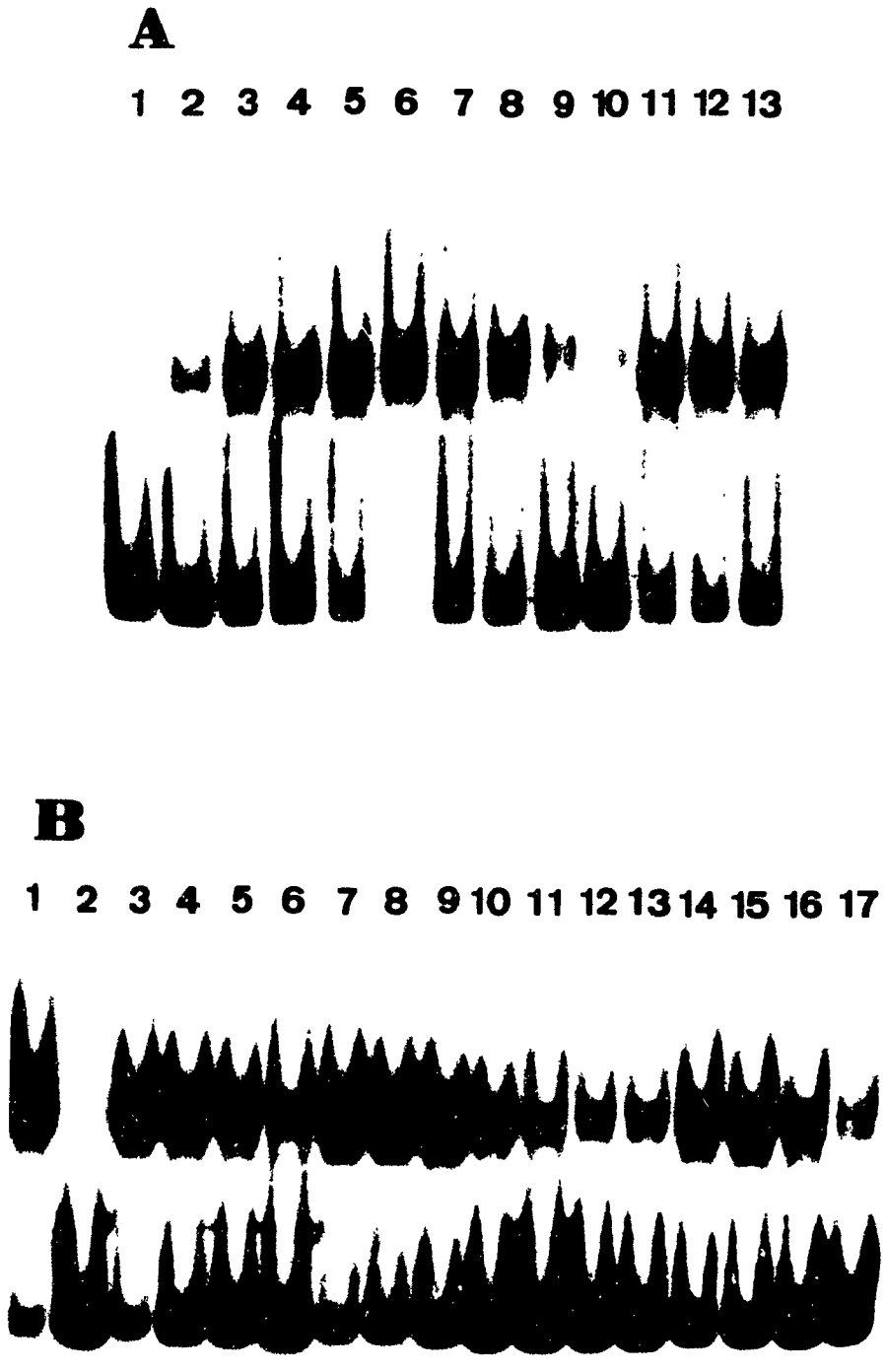


Figure 5. Gel retardation assays with Lrp protein and the *lysU* promoter region.

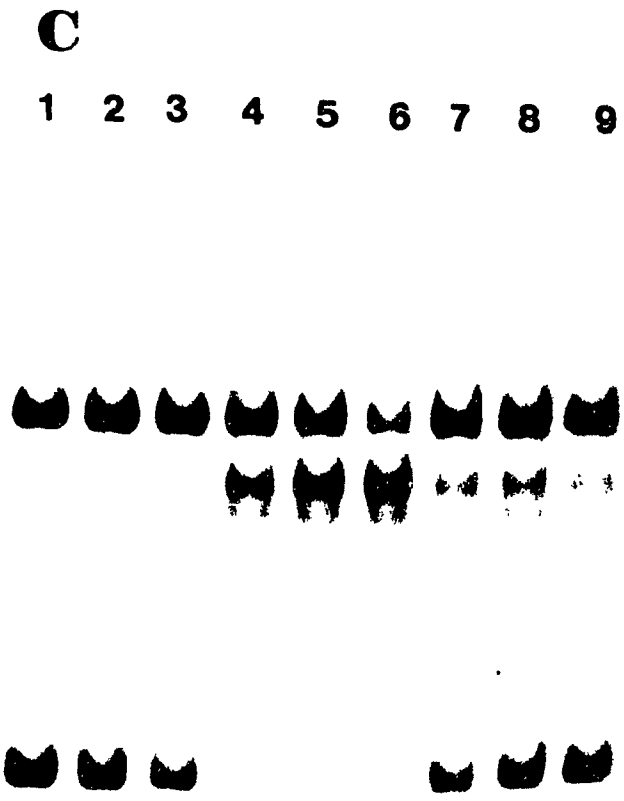


Figure 5. Gel retardation assays with Lrp protein and the *lysU* promoter region.

Figure 5. Gel retardation assays with Lrp protein and the *lysU* promoter region. Gel retardation analysis was performed as described in Materials and Methods. (A). Gel retardation assays performed with the 150 bp *Hin*dIII-to-*Sac*I fragment from pFU10 which carries 108 bp of *lysU* upstream region. Lanes 1-6 represent incubations with 0, 25, 50, 100, 200, and 400 ng of Lrp protein. Lanes 7-13 represent incubations with 200 ng of Lrp protein and L-leucine at 2, 4, 8, and 16 mM, glycine at 10, 20, and 40 mM, respectively. (B). Effect of various amino acids on gel retardation assays with Lrp protein. Experiments were performed as in (A) with 200 ng of Lrp protein in all experiments except lane 2 in which no protein was added. L-leucine 1, 2, 4, and 8 mM; glycine 10, 20, and 40 mM; L-alanine 5, 10, 20, and 40; and L-isoleucine 4, 8, 16, and 32 mM were added to lanes 3-17 respectively. (C) Lrp specifically binds to the *lysU* upstream region. A 497 bp *Dra*I DNA fragment from pFU10 was digested with *Sau*3A1; 3'-end-^{32}P-radiolabeled with Klenow enzyme; and incubated with varying amounts of purified Lrp protein. Lanes 1-6 represent the addition of increasing amount of Lrp protein (0, 50, 100, 200, 300, and 400 ng, respectively). Lanes 7-9 represent incubations with 400 ng of Lrp protein and L-leucine at 5, 10, and 20 mM, respectively.

slowly than did the same fragment analyzed in the absence of Lrp protein (Fig. 5C). On the other hand, migration of the 382 bp fragment was essentially unaffected by the presence of Lrp protein. The 7 bp fragment was too small to test, and in any case ran off the gel.

It is clear that the 108 bp fragment is retarded by Lrp. To study this further, the 108 bp fragment was inserted into the *Bam*HI-*Eco*RV sites of Bluescript^{+}, yielding plasmid pFU10. When this plasmid pFU10 was cut with *Hin*dIII and Sac, a 150 bp fragment which contained the entire 108 bp sequence was formed. This 150 bp fragment was retarded by Lrp just as the 108 bp fragment had been. As shown in Figure 5A, the fraction of DNA retarded increased as the amount of Lrp protein increased from 25 to 400 ng, all the input DNA being bound when 400 ng of binding protein was added.

If the binding of Lrp protein to DNA from the upstream region of *lysU* represents a specific interaction of Lrp protein with this DNA, and if binding is directly related to regulation of gene expression by Lrp protein, one might expect that the Lrp/DNA complex would not be formed in the presence of L-leucine. In Figure 5A, lanes 7-10 show the effect of adding increasing concentrations of L-leucine to incubations of the 150 bp fragment with 200 ng of the same preparation of Lrp protein used earlier. Addition of 4 mM L-leucine decreased binding of Lrp to the DNA noticeably, and at the presence of 16 mM L-leucine, no retardation band could be seen. The concentration of L-leucine which eliminated binding of Lrp to the *lysU* DNA is comparable to the concentration which greatly

reduced binding of Lrp to DNA from the upstream region of *IlvIH* [Ricca *et al.*, 1989].

As shown above, synthesis of L-serine deaminase, the *sdaA* gene product, is also regulated by Lrp. Because L-SD is induced both by L-leucine and glycine, and the expression of *lysU* is induced by the presence of L-glycylleucine [Hirshfield *et al.*, 1981], I thought that *lysU* might also be regulated by both amino acids, and so tested the effect of glycine on the binding of Lrp protein to *lysU*. Even at 40 mM, glycine had very little effect (Fig 5B, lanes 7-9). This suggests that glycine is not an effector in regulation of the *lysU* gene by Lrp.

Two other amino acids, L-alanine and L-isoleucine, also affected retardation of the 150 bp fragment by Lrp. The effect of L-alanine is shown in Fig. 5B, lanes 10-13, and that of L-isoleucine in Fig. 5B, lanes 14-17. The fact that L-alanine decreased retardation, perhaps removing the repressing protein from the DNA correlates well with the fact that L-alanine induces lysyl-tRNA synthetase activity [Hirshfield *et al.*, 1981] and increases *lysU* expression as shown by 2-D gel electrophoresis (R. Matthews, personal communication). The effect of L-isoleucine on lysyl-tRNA synthetase activity is not known, but addition of L-isoleucine and valine to the medium does not lead to increased expression of *lysU* as analyzed by 2-D gel electrophoresis (R. Matthews, personal communication).

Since the Lrp protein binds to at least two sites upstream of *IlvIH* [Ricca *et al.*, 1989], I tested the possibility of Lrp

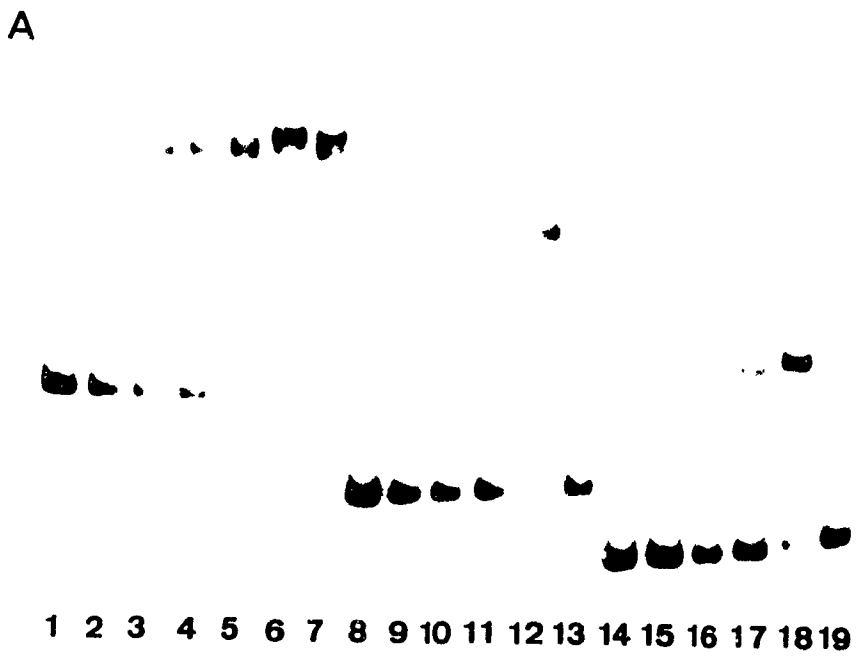


Figure 6. Binding of Lrp to the *lysU* upstream region.

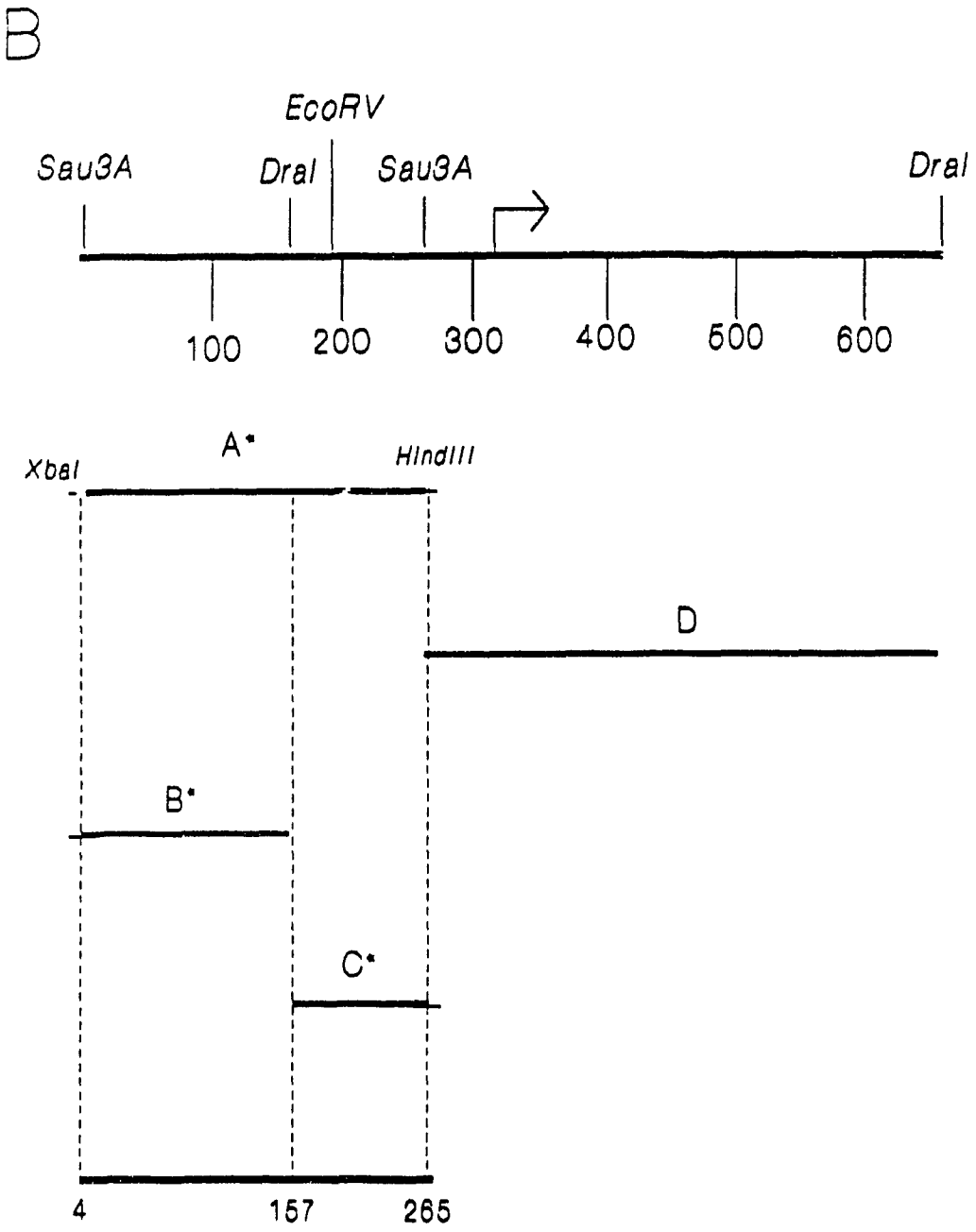


Figure 6. Binding of Lrp to the lysU upstream region.

Figure 6. Binding of Lrp to the *lysU* upstream region. (A) Gel retardation assays. Lanes 1 through 7: a 308 bp end-labelled *Xba*I-to-*Hin*dIII fragments from plasmid pFU20 (fragment A identified in panel A) were incubated with 0, 12.5, 25, 50, 100, 200 and 200 ng of purified Lrp protein, respectively. Lanes 8 through 13 represent incubation of a 168 bp end-labelled *Xba*I-to-*Dra*I fragment from plasmid pFU20 (fragment B in panel A) with 0, 25, 50, 100, 200, and 200 ng of Lrp protein, respectively. Lanes 14 through 19 represent incubation of a 140 bp *Hin*dIII-to-*Dra*I fragment from plasmid pFU20 (fragment C in panel A) with 0, 25, 50, 100 200, and 200 ng of Lrp protein respectively. In lanes 7, 13, and 19, L-leucine was added to 20 mM. (B) Schematic representation of the *lysU* upstream region, with the thick line corresponding to bacterial DNA and the thin line corresponding to polylinker in plasmid Bluescript. The positions of several restriction sites are shown.

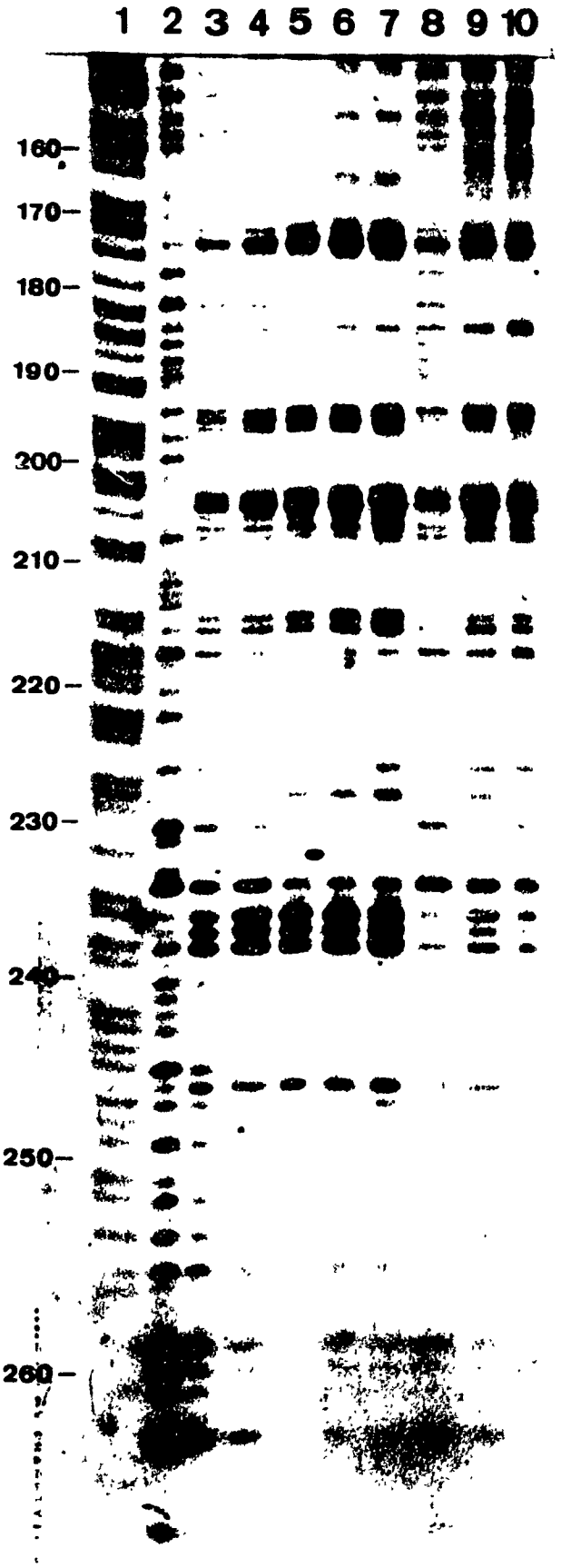
binding to other site(s) upstream of *lysU*. As shown in Fig. 6 (lanes 8-13), Lrp also binds to the DNA fragment containing a 154 bp sequence upstream of the 108 bp sequence (corresponding to nucleotides 4-157 from the *lysU* sequence [Clark and Neidhardt, 1990]. Fig. 6 also shows that the concentration dependence for binding of Lrp to the DNA fragment containing the 154 bp sequence is similar to that for binding to the fragment containing the 108 bp sequence (compare lanes 8-13 with lanes 14-19). The concentration for binding of Lrp to the fragment containing both 154 bp and 108 bp sequences is lower (Fig. 6, lanes 1-7). This suggests that the 154 bp and 108 bp DNA fragments each carry only a part of the binding site.

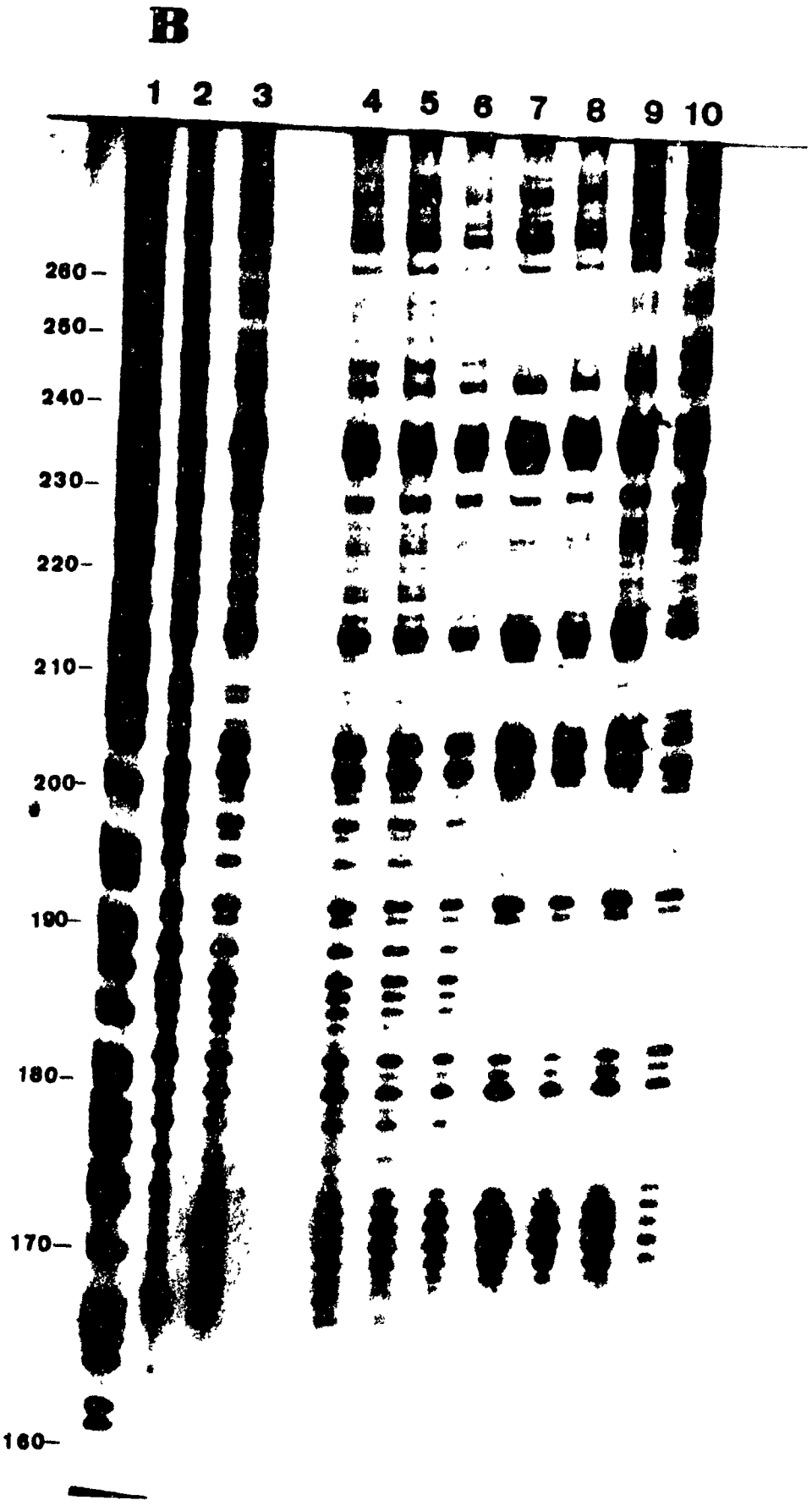
I-2. DNase I Footprint for Lrp Binding to *lysU* Upstream Region

To identify the upstream sequence of *lysU* to which Lrp protein binds, DNase I protection experiments were performed. The 3'-end labelled DNA fragments which contain the 108 bp *lysU* promoter region were saturated with increasing concentrations of Lrp protein before treatment with DNase I. Several regions from nucleotide 157 to nucleotide 265 in the *lysU* coding region were protected by Lrp protein from DNase I digestion (compare, for instance, lanes 2 and 5 in Fig. 7A, which contain 0 and 8 ng of Lrp, respectively). Concomitant addition of L-leucine substantially decreased the protection observed (Fig. 7A lanes 5, 8; Fig. 7B lanes 9, 10).

To determine whether subregions of the DNA protected by the Lrp protein from DNase I cleavage showed different affinities for Lrp, I varied the amount of Lrp protein (from

A




B
1 2 3 4 5 6 7 8 9 10
260–
250–
240–
230–
220–
210–
200–
190–
180–
170–
160–

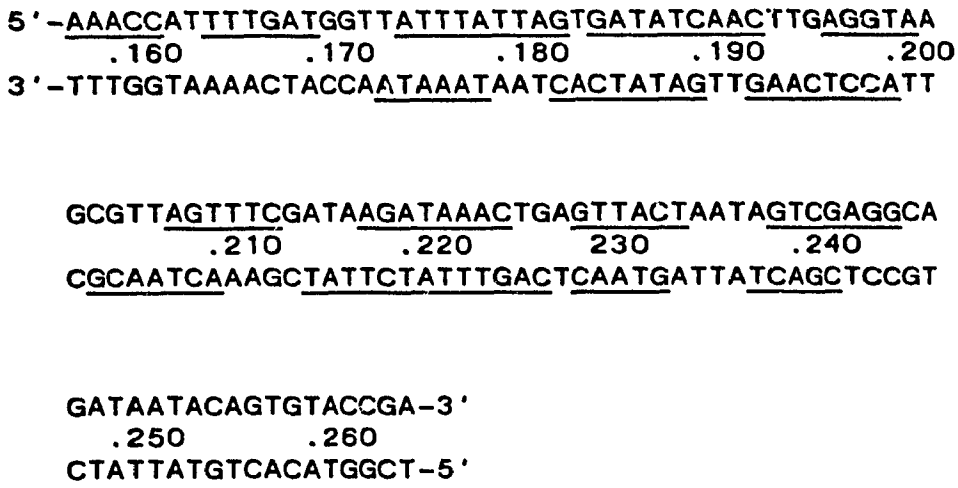


Figure 7(C). Schematic representation of DNase I footprints of Lrp bound to *lysU* upstream sequence (both strands).

Figure 7. Protection of the *lysU* upstream region by Lrp protein from DNase I digestion. (A). Digestion pattern of the coding strand. The 147 bp *Xba*I-to-*Xho*I fragment from pFU10 was labelled with ^{32}P at 3' end of the *Xba*I-cleaved site for the footprinting of the coding strand. The ladder of the Maxam-Gilbert A+G reaction of probe DNA was shown in lane 1. Varying amounts of Lrp protein were incubated with labelled DNA before treatment with DNase I. Lanes 2-7 represent incubations with 0, 2, 4, 8, 16, and 32 ng of Lrp protein, respectively. Lanes 8-10 represent incubations with 8, 16, and 32 ng Lrp protein in the presence of 20 mM L-leucine. (B). Digestion pattern of the non-coding strand. The 150 bp *Hin*dIII-to-*Sac*I DNA fragment was labelled at 3' end of the *Hin*dIII-cleaved site. Lane 1, A+G sequencing ladder. Lanes 2-9, addition of increasing amounts of Lrp protein (0, 2, 4, 6, 8, 16, 32, and 64 ng, respectively). Lane 10 represents an incubation with 64 ng of Lrp protein and 16 mM L-leucine. (C). A comparison of sites protected by Lrp protein on the two strands of *lysU* upstream sequence. Solid lines are drawn under protected nucleotides. This fragment runs from nucleotide 158 to 265 of the original clone [Clark and Neidhardt, 1990] on which the translation startsite is at 316.

0 to 64 ng) incubated with the DNA fragment which was labelled in the non-coding strand (Fig. 7B, 2-9).There was no evidence for differential affinity of Lrp for subsites within the region. Instead I observed a concentration-dependent increase in protection over the entire region.

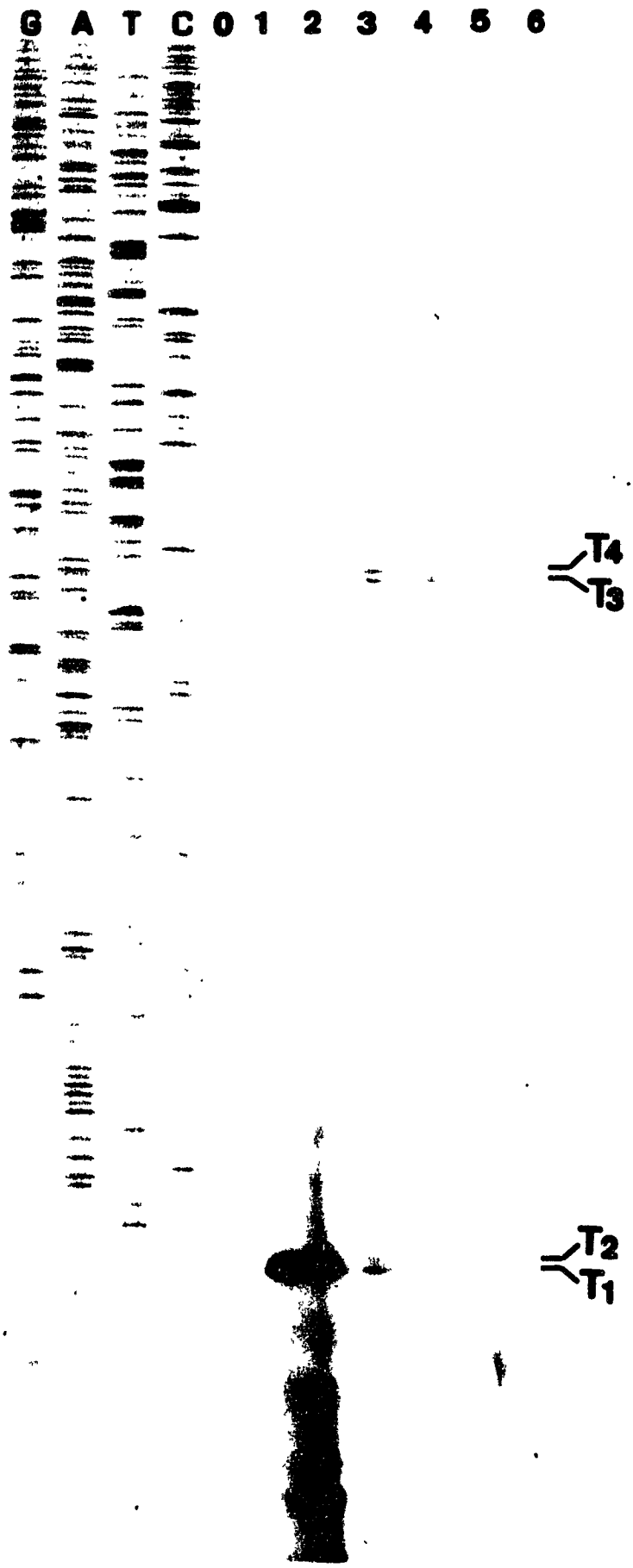
Fig. 7C compares the patterns of protection seen on the coding and non-coding strands. The protected areas do not correspond exactly on the two strands, but overlap for most of their length. The regions separating the protected areas showed a regular pattern of alternately increased and decreased sensitivity to DNase I digestion, as was reported for the MalT activator binding to regions upstream of *malE* and *malK* [Raibaud et *al.*, 1989].

II. The Interaction between Lrp Protein and *serA* Regulatory Region

It seems then that Lrp binds to the upstream regions of both *lysU* and *IlvIH*, Since *In vivo* data indicated that the transcription of *serA*, like *IlvIH*, is positively regulated by Lrp and reduced by the presence of exogenous L-leucine, it seemed likely that Lrp would also bind to the *serA* upstream region. In this section, I examine the possibility of protein-DNA interaction between Lrp and the *serA* upstream region.

II-1. Mapping of the *serA* Promoter

The initiation sites of *serA* transcription had not been determined. I therefore identified them by primer extension with reverse transcriptase and a 22-mer 5'-end-labelled


G A T C 0 1 2 3 4 5 6
T4
T3
T2
T1

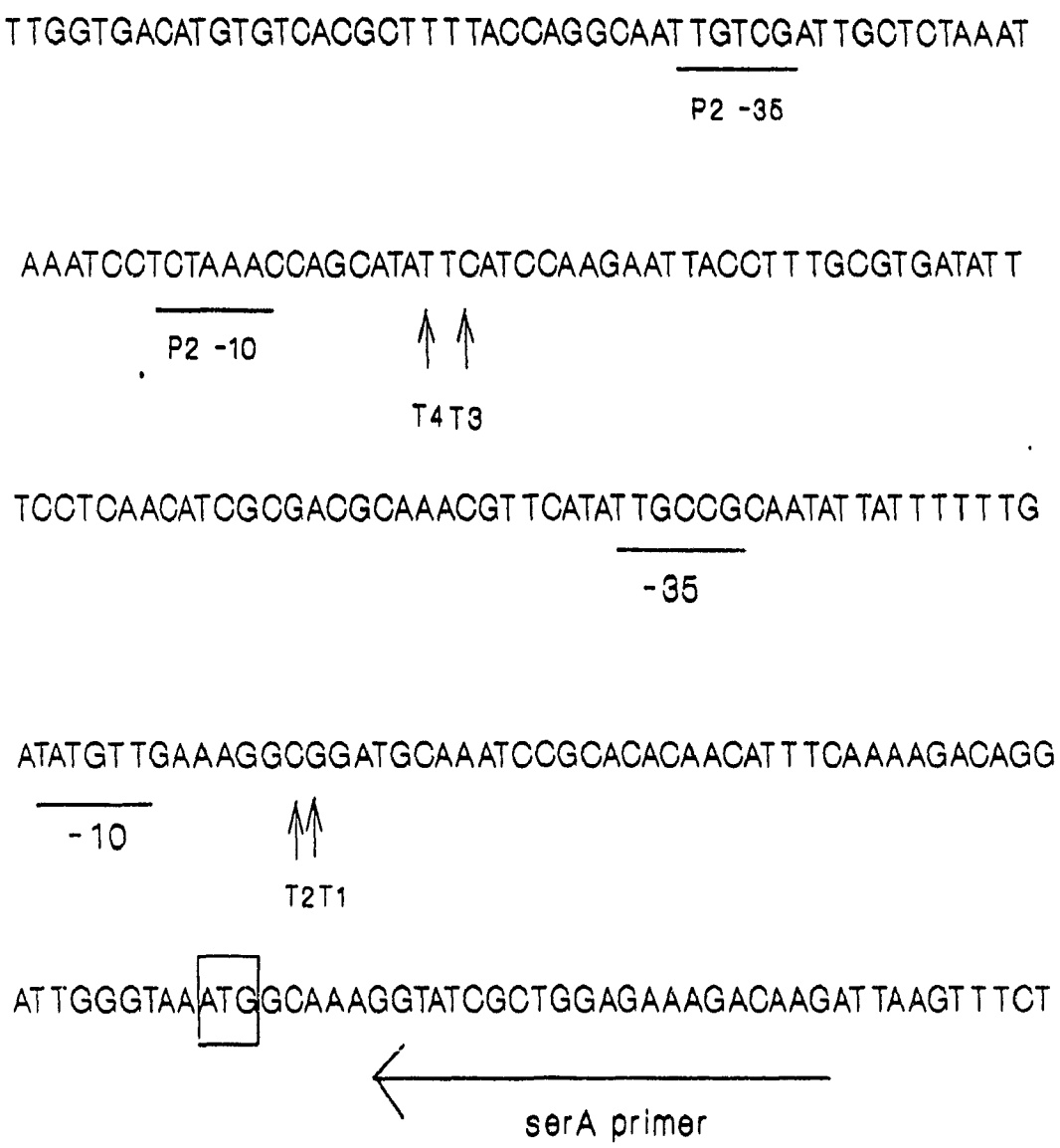


Figure 8. Mapping of the 5' terminus of serA mRNAs.

Figure 8. Mapping of the 5' terminus of *serA* mRNAs. Total cellular RNA from strains MEW1/pGT17 (*lrp*$^+$) or MEW26/pGT17 (*lrp*$^-$) was hybridized with a 5'-end-labelled single stranded oligonucleotide primer corresponding to bases 9 to 30 from the translation start site (Tobey and Grant, 1986). After extension of the primer with AMV reverse transcriptase, the products were resolved by electrophoresis on a 6% polyacrylamide 8 M urea gel. The size of the transcripts was determined by comparison with the products of a sequencing reaction performed with the same oligonucleotide primer. Lanes G, A, T, C correspond to the sequencing pattern of *serA* obtained through the dideoxy chain termination method [Sanger *et al.*, 1977] with a single-stranded DNA template from plasmid pGU2. Lane 1-6, various extension products obtained from RNAs of strain MEW1/pGT17 or strain MEW26/pGT17: lane 1-3, strain MEW1/pGT17 grown in glucose minimal medium without or with 100 µg/ml L-leucine, or grown in LB medium, respectively; lanes 4-6, strain MEW26/pGT17 grown in glucose minimal medium without or with 100 µg/ml L-leucine, or grown in LB medium, respectively. The extension reaction of lane 0 was the same as that of lane 1 except that reverse transcriptase was not added.

oligonucleotide primer (The primer was kindly provided by J. M. Calvo) which hybridized near the AUG start codon of the *serA* gene (Fig. 8). Extension products were analyzed by polyacrylamide gel electrophoresis, in parallel with the products of a dideoxy sequencing reaction performed with the same oligonucleotide primer. Several extension products appeared on the electrophoretograms of the extension reactions performed with the RNAs isolated from strain MEW1 (*lrp*$^+$) and strain MEW26 (*lrp*$^-$) carrying plasmid pGT17 (Fig 8). Transcript T1 is the major product in the wild-type strain MEW1 grown in minimal medium whether with or without L-leucine (Fig. 8, lanes 1 and 2). This transcript is present in much lower amount either in the wildtype strain grown in LB medium or in the *lrp* mutant MEW26 grown in both minimal medium and LB medium.

The size of this extension product indicated that the initiation site of transcription was located 45 bp upstream from the *serA* start codon AUG (Fig. 8). The G residue of the T1 initiation site was designated +1 (Fig. 8). The DNA sequence upstream from this start site (Fig. 8) reveals a -10 region showing 4 of 6 bp homology with the consensus Pribnow box sequence and a -35 region showing 4 of 6 bp homology with the -35 box consensus sequence for other *E. coli* promoters [Hawley and McClue, 1983]. Another transcript, T2, which is only one bp longer than transcript T1, seems to be transcribed from the same promoter, P_1. It is clear that the transcription from this promoter is activated by *lrp* gene product.

Transcripts T3 and T4 appeared on the electrophoretograms

of the extension reactions performed with the RNAs isolated from strain MEW26 (*lrp*$^-$) carrying plasmid pGT17, or from wild-type strain grown in LB medium (Fig. 8, lanes 3-6). Since transcript T4 is only two bp longer than T3, it seems that these two transcripts are transcribed from the same promoter, P_2. The sizes of these two extension product indicated that the initiation site of transcription from P_2 is located 138 and 140 bp upstream from the *serA* start codon (Fig. 8). Upstream from these transcriptional start points, both -10 and -35 region homology with the consensus sequence could be recognized (Fig. 8). The primer extension analysis also demonstrates that the transcription from promoter P_2 is repressed by the *lrp* gene product. In both promoters, the sequence resembles standard *E. coli* σ^{70} promoters rather than sequences recognized by other σ factors.

In summary, it seems that transcription from *serA* can start at two different promoters. One of these, P_1, is used in the presence of Lrp. Initiation at the other, P_2, is repressed by Lrp.

II-2. Lrp protein Binds to at least Two Sites in the *serA* upstream region

Since the *serA* upstream region showed two promoters, I wished to know whether Lrp might bind near either or both. To study this, several plasmids which contain the *serA* promoter and upstream region cloned into the polylinker of vector Bluescript were constructed as described in Materials and Methods. To determine the interaction of Lrp protein with the

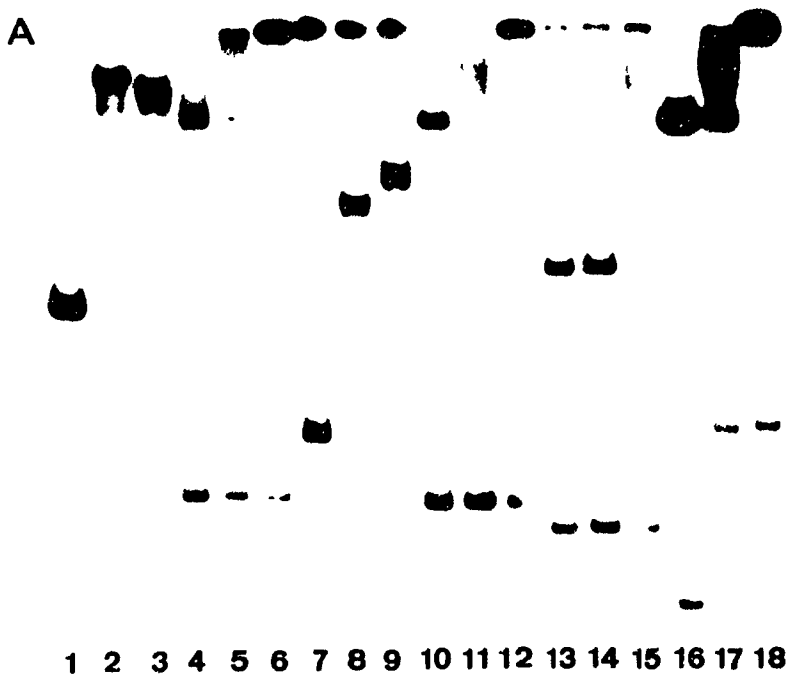


Figure 9. DNA binding assay of Lrp to the fragments derived from the *serA* upstream region.

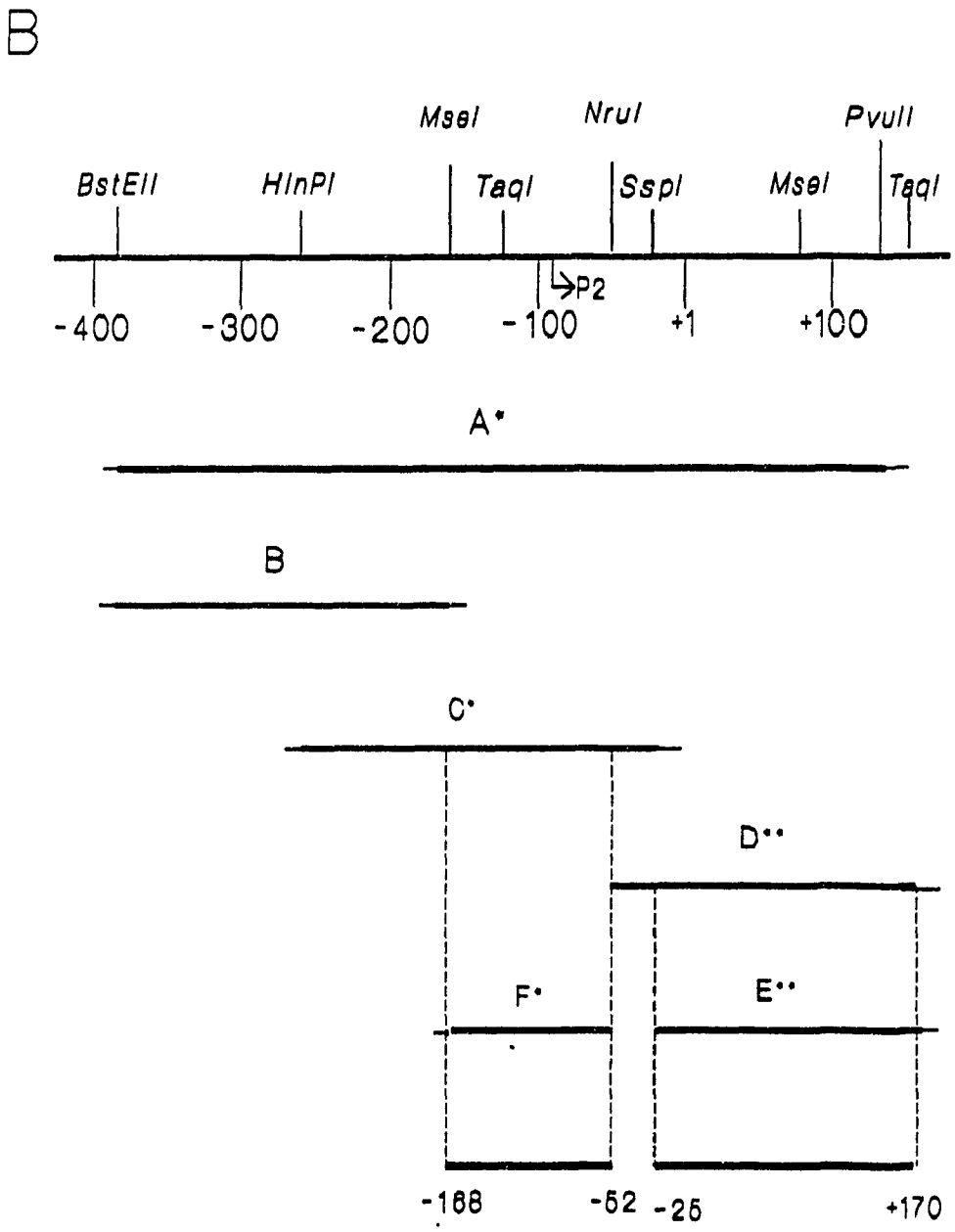


Figure 9. Binding of Lrp to the serA upstream region.

Figure 9. DNA binding assay of Lrp to the fragments derived from the *serA* upstream region. (A) End-labelled fragments identified in panel B were incubated without (lanes 1, 4, 7, 10, 13, and 16), with 25 ng (lanes 2, 5, 8, 11, 14, and 17), or 250 ng (lanes 3, 6, 9, 12, 15, and 18) of purified Lrp protein and analyzed by the gel retardation assay. The DNA fragments used are: A (lanes 1-3), 554 bp *Hin*dIII-to-*Bam*HI fragment isolated from pGU3; B (lanes 4-6), pGU4 digested with the *Eco*RI and *Hin*dIII; C (lanes 7-9), 306 bp *Xba*I-to-*Kpn*I fragment isolated from pGU5; D (lanes 10-12), pGU6 digested with the *Eco*RI and *Nru*I; E (lanes 13-15), pGU6 digested with the *Eco*RI and *Ssp*I; F (lanes 16-18), pGU7 digested with the *Eco*RI and *Nru*I. In B, D, E, and F, the Larger (top) fragment is from Bluescript$^+$, which has at least one Lrp binding site. (B) The top line represents a physical map of the *serA* upstream region. The thick lines correspond to bacterial DNA and the thin lines correspond to polylinker in plasmid Bluescript$^+$. +1, the transcriptional start site of $serAP_1$; *, fragments that bind Lrp protein; **, same as * but with low binding affinity, as shown in panel A. The thick bars at the bottom represent regions within which binding occurs, as deduced from the results of this experiment.

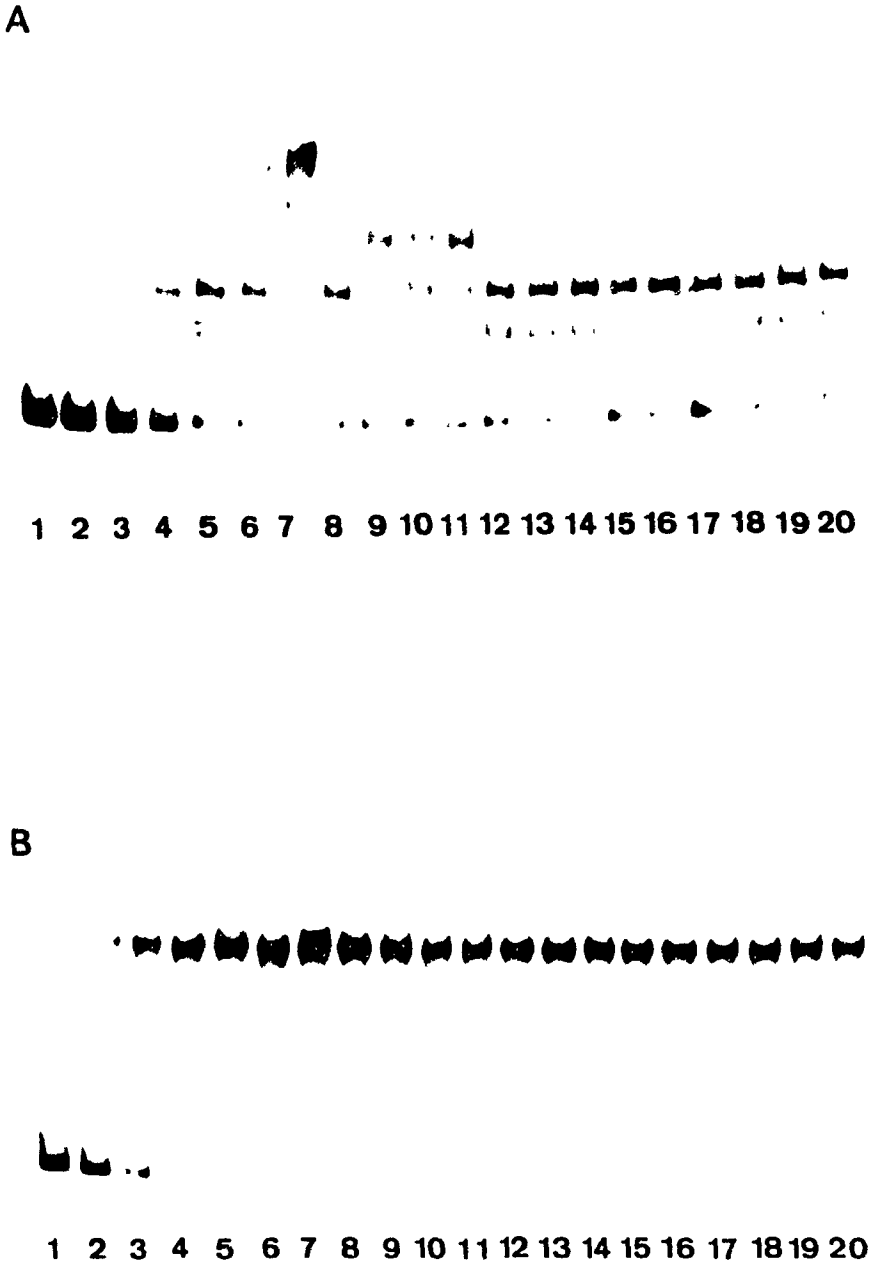


Figure 10. Gel retardation assays performed with *serA* DNA fragments.

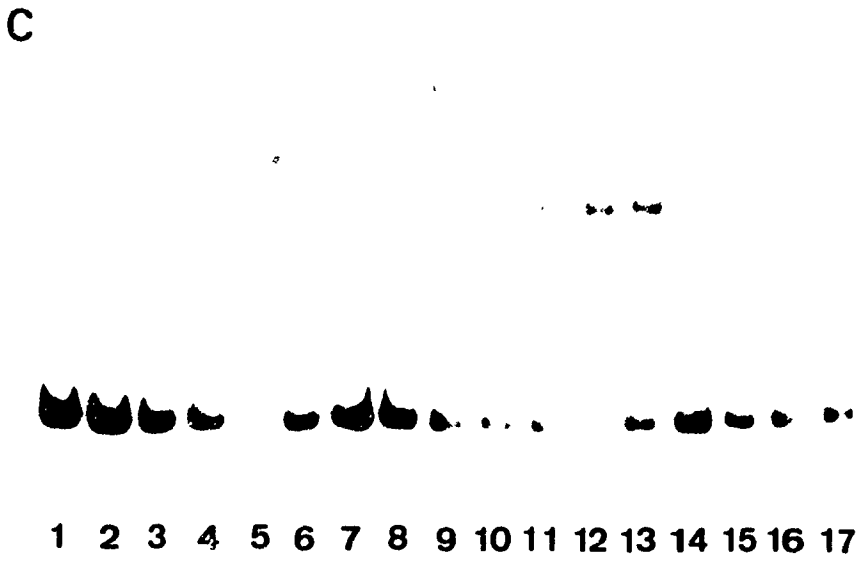


Figure 10. Gel retardation assays performed with *serA* DNA fragments.

Figure 10. Gel retardation assays performed with *serA* DNA fragments. (A) Binding of Lrp to the DNA fragment carrying both high affinity and low affinity binding sites. A 3'-end labelled *Eco*RI-to-*Xho*I fragment from pGU6 was incubated with various amounts of purified Lrp protein in the absence or presence of different concentrations of amino acids. Lanes 1 through 7 represent the addition of 0, 6.3, 12.5, 25, 50, 100, and 200 ng Lrp protein, respectively. 50 ng of Lrp were added to lanes 8 through 20. L-leucine 5, 10, 20 mM, glycine 10, 20, 40 mM, L-alanine 10, 20, 40 mM, and L-isoleucine 5, 10, 20 mM were added to lanes 9 through 20, respectively. (B) Binding of Lrp to the DNA fragment carrying only the high affinity binding site. Experiments were performed as in (A) except that the 3'-end labelled *Xba*I-to-*Kpn*I fragment from pGU5 was used in the binding assay. (C) Binding of Lrp to the *serA* low affinity binding site. Binding reaction containing 3'-end labelled *Eco*RI-to-*Nru*I DNA fragment prepared from plasmid pGU6. Each reaction contains 400 ng Lrp protein except lanes 1, 2, 3, and 4, which contains 0, 50, 100, and 200 ng Lrp protein, respectively.

serA upstream region, a variety of end-labelled DNA fragments containing different *serA* upstream regions were derived from these plasmids and tested for their ability to bind to Lrp protein and be retarded by it.

The results of these gel retardation assays indicated that the Lrp protein binds to at least two different regions, one located between -168 and -52 and the other between -25 and +170 relative to the transcription startpoint (P_1) (Fig. 9). As shown in Fig. 9, the binding affinity of the upstream site (-168 to -52) for Lrp protein is much higher than that of the downstream site (-25 to +170).

The results of Lrp binding to a DNA fragment containing both binding sites are shown in Fig. 10A. After electrophoresis, DNA was found in six positions: the free DNA band with the fastest mobility, and five shifted bands corresponding to protein-DNA complexes. Four bound-DNA bands were observed in the gel retardation assay performed with the DNA fragment carrying only the low affinity binding site (Fig. 10C). However, only one major protein-DNA complex was formed in the interaction between Lrp and the DNA fragment carrying the high affinity binding site (Fig. 10B). These observation suggest that there must be more than one site in the low affinity binding region.

To determine whether the presence of L-leucine affects the binding of Lrp, the DNA fragments which contain either a single binding site or both binding sites were used in gel retardation assays in the presence of L-leucine or other amino acids. As shown in Fig. 10A (lanes 9-11), the presence of L-leucine did affect the binding of Lrp to DNA fragments

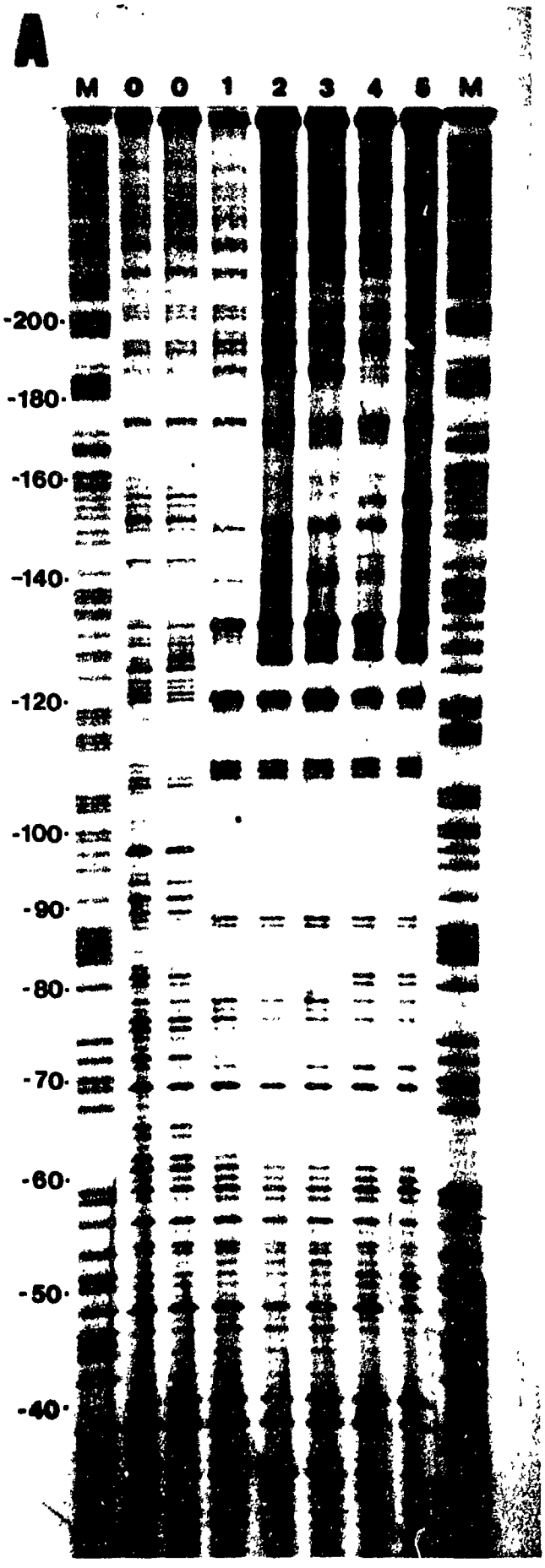
containing both binding sites. However, whereas at *lysU*, L-leucine greatly reduced Lrp binding, at *serA*, the addition of L-leucine resulted in the formation of a further retarded DNA band not seen in its absence.

These results can be explained in at least two ways. First, Lrp with L-leucine bound to it may bind to a new site of this DNA to which unliganded Lrp does not bind. Alternatively, binding of L-leucine may change the conformation of the DNA-protein complex. The addition of L-alanine only slightly affected the retardation pattern (Fig. 10A, lanes 15-17). Neither glycine nor L-isoleucine affected the retardation pattern (Fig. 10A, lanes 12-14, 18-20).

The results were different when DNA carrying only one of the two sites was tested. Either L-leucine or L-alanine reduced the binding of Lrp to the DNA fragment carrying the low affinity binding region (Fig. 10C, lanes 7-9), but neither affected binding of Lrp to the DNA fragment carrying only high affinity binding site (Fig. 10B).

DNase I footprinting assays were performed to further define the high affinity Lrp binding site. The 3'-end-labelled *Xba*I-to-*Kpn*I fragment from plasmid pGU5 (coding strand) was saturated with increasing concentrations of Lrp protein before treatment with DNase I. The results demonstrate that Lrp protein protects this fragment from DNase I action in the region -155 to -81 (Fig. 11).

It seems clear from the combination of retardation and DNase I protection, that Lrp binds with high affinity in the -155 to -81 region of *serA*. This is consistent with the idea that Lrp represses transcription from this promoter.


A
M
0
0
1
2
3
4
5
M
-200-
-180-
-160-
-140-
-120-
-100-
-90-
-80-
-70-
-60-
-50-
-40-

```
      .-260      .-250      .-240      .-230      .-220
5'-CGCGGACGTTGCGTGATTGGCACACCTGACGGTGTCAAAACCATT

           .-210      .-200      .-190      .-180
GTGAAACTGACTCTGACGGGGGAACCTCCCNNTAAAAAAATTCTC

     .-170      .-160      .-150      .-140      .-130
TTCATTAAATTTGGTGACATGTGTCACGCTTTTACCAGGCAATTG

           .-120      .-110      .-100      .-90
TCGATTGCTCTAAATAAATCCTCTAAACCAGCATATTCATCCAAG

     .-80       .-70       .-60       .-50       .-40
AATTACCTTTGCGTGATATTTCCTCAACATCGCGACGCAAACGTT

           .-30
CATATTGCCGCAAT
```

Figure 11(B). Schematic representation of the DNase I footprint of Lrp bound to the *serA* upstream sequence (coding strand).

Figure 11. DNase I footprint analysis on the *serA* sequence. (A). Digestion pattern of the coding strand. The 306 bp *Xba*I-to-*Kpn*I fragment from pGU5, labelled at the *Xba*I end with α-^{32}P-dATP, was incubated with increasing amounts of purified Lrp protein: lane 0, no Lrp; 1, 4 ng; 2, 8 ng; 3, 16 ng; 4, 32 ng; 5, 64 ng. After DNase I treatment, the samples were fractionated by electrophoresis through an acrylamide-urea gel. The same labelled fragment was subjected to the A+G sequencing reaction and served as a size standard (lane M). (B). Boundaries of DNase I protection on the *serA* sequence are shown from -155 to -81. Solid lines are drawn under protected nucleotides. The G residue of the T1 initiation site was designated +1 (Fig. 8).

III. The interaction between Lrp protein and the *sdaA* upstream sequence

The preceding data showed that Lrp binds to the upstream regions of both *lysU* and *serA*. Since in the case of *sdaA*, as in the other two cases, *in vivo* experiments demonstrate that the expression of *sdaA* gene is negatively regulated by Lrp, it seemed likely that Lrp would also bind to the *sdaA* upstream region. In this section, I examine the possibility of protein-DNA interaction between Lrp and the *sdaA* upstream region.

This question was again approached with gel retardation assays. For this, DNA fragments of various sizes were subcloned from the *sdaA* upstream region into the *Eco*RV site of the Bluescript^{+} polylinker (see Materials and Methods). Several 3'-end labelled DNA fragments containing parts of the *sdaA* upstream sequence were used for binding experiments. As shown in Fig. 12, at the 5 ng/µl concentration, Lrp could bind only to fragments containing the 171 bp *Ssp*I-to-*Eco*RV sequence. However, at higher concentration, Lrp was shown to bind to a new site downstream of the high affinity site (Fig. 13, lanes 11 through 15).

The expression of *sdaA* is known to be induced by the presence of glycine, L-leucine, or L-alanine. If the effect of these amino acids on *sdaA* expression is mediated via Lrp protein, one would expect, by the same reasoning as used earlier, that the presence of the amino acid would affect Lrp binding. To determine this, the binding assays were performed with DNA fragments containing either the high affinity site,

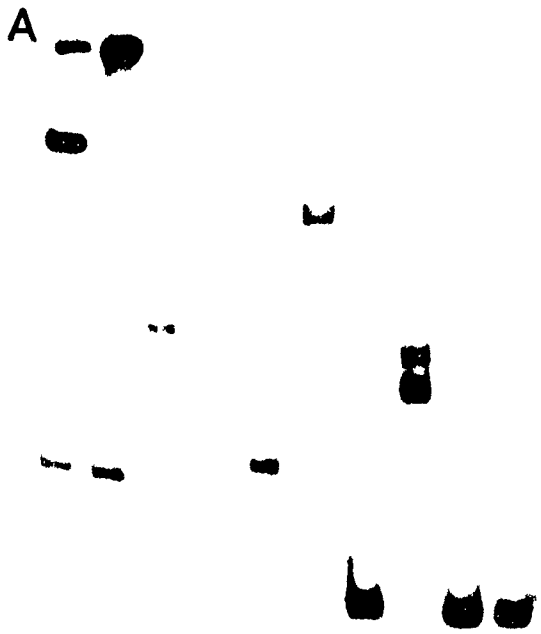


Figure 12, Binding of Lrp to the *sdaA* upstream region.

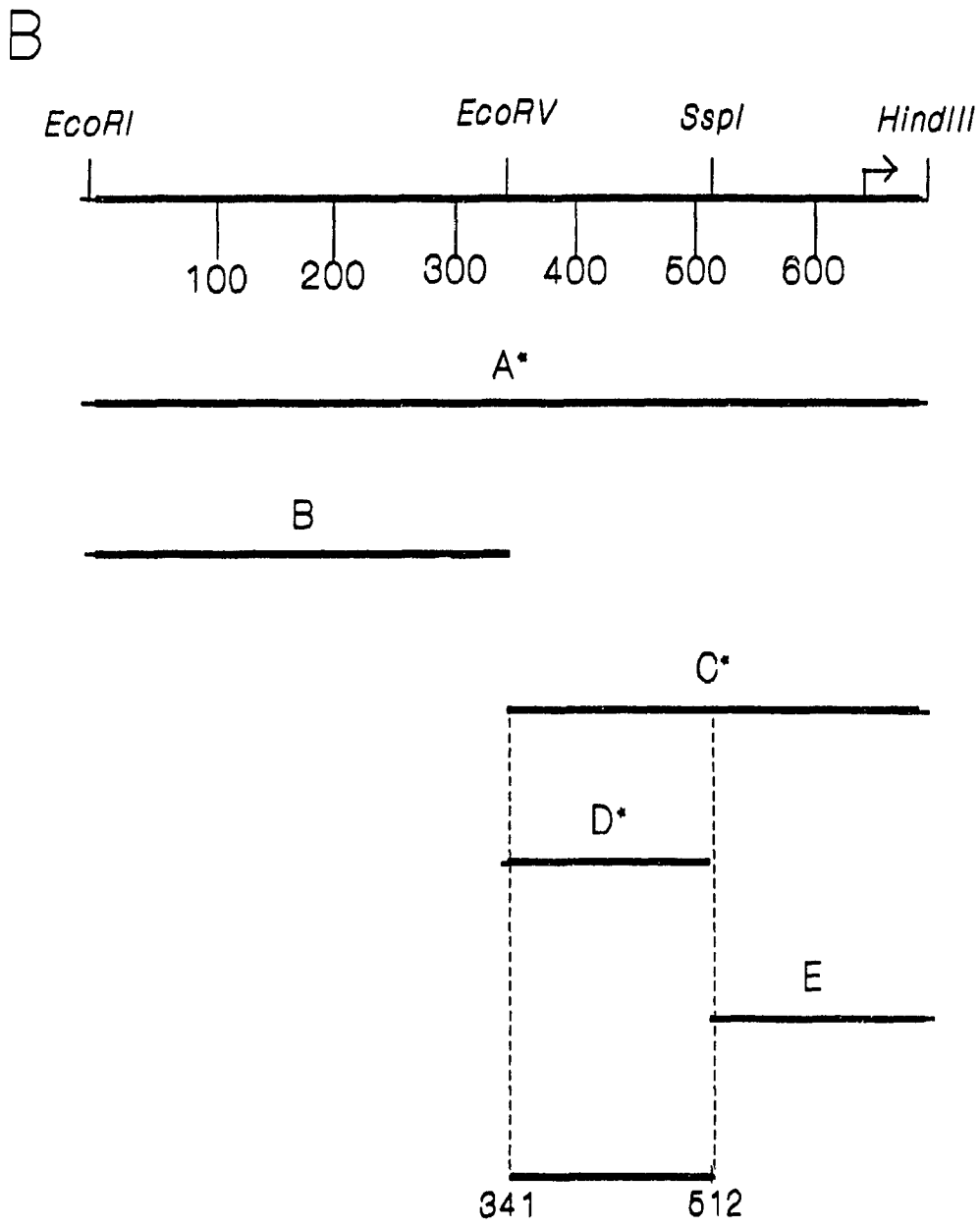


Figure 12. Binding of Lrp to the sdaA upstream region

Figure 12, Binding of Lrp to the *sdaA* upstream region. (A) End-labelled fragments identified in panel B were incubated without (lanes 1, 3, 5, 7, 9) or with (lanes 2, 4, 6, 8, 10) 100 ng purified Lrp protein in a total volume of 20 µl and gel retardation assays were carried out. In lanes 1 and 2, the larger (top band) fragment is from plasmid Bluescript$^+$. lanes 1 and 2 (fragment B): pSU1 digests with the *Eco*RI and *Eco*RV; Lanes 3 and 4 (A): 681 bp *Eco*RI-to-*Hin*dIII fragment from pSU1; Lanes 5 and 6 (C): 361 bp *Hin*dIII-to-*Pst*I fragment from pSU2; Lanes 7 and 8 (D): 189 bp *Hin*dIII-to-*Pst*I fragment from pSU4; Lanes 9 and 10 (E): 190 bp *Hin*dIII-to-*Pst*I fragment from pSU3. (B) The top line represents a physical map of the *sdaA* upstream region. The thick lines correspond to bacterial DNA and the thin lines correspond to polylinker in plasmid Bluescript$^+$. *, fragments that bind Lrp protein, as shown in panel A. The thick bars at the bottom represent the region within which binding occurs, as judged from the results of this experiment.

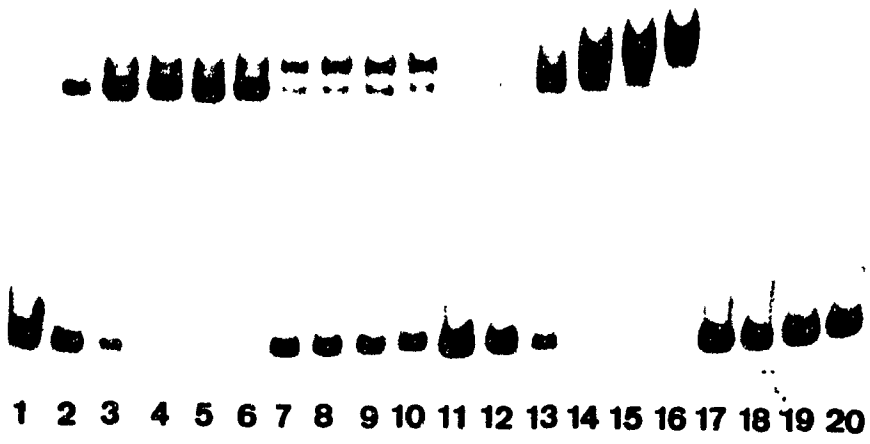


Figure 13. Lrp binds to at least two sites upstream of *sdaA*.

Figure 13. Lrp binds to at least two sites upstream of *sdaA*. End labelled *Hin*dIII-to-*Pst*I fragment from pSU4 (lanes 1-10) or from pSU3 (lanes 11-20) were used in gel retardation assays. Lanes 1 through 20 represent incubations with 0, 20, 40, 60, 80, 100, 40, 60, 80, 100, 0, 200, 400, 600, 800, 1000, 400, 600, 800, and 1000 ng, respectively. Lanes 7 -10, 17-20 the incubation mix also contained L-leucine at 20 mM.

the low affinity site, or both. As shown in Fig. 13 (lanes 7 through 10, 17 through 20) and Fig. 14 (14A, lanes 8 through 12; 14B, lanes 3 through 7; 14C, lanes 3 through 7), in all cases, the presence of L-leucine greatly reduced Lrp binding activity, particularly with the DNA fragment containing the low affinity site or both sites.

The presence of glycine had no significant effect on Lrp binding activity (Fig. 14A, lanes 13 through 16; 14B, lanes 8 through 11; 14C, lanes 8 through 11). However glycine is known to be an inducer of L-SD synthesis. This seems to suggest that the effect of glycine on *sdaA* expression is not mediated by Lrp.

Two other amino acids, L-alanine and L-isoleucine, did affect binding of Lrp to the *sdaA* upstream region. The effect of L-alanine is shown in Fig. 14 (14A, lanes 17 through 20; 14B and 14C, lanes 12 through 16) and that of L-isoleucine is shown in Fig. 14B and 14C, lanes 17 through 20. The fact that L-alanine interferes with Lrp binding to *sdaA*, perhaps removing the repressor from the DNA, correlates well with the *in vivo* result that L-alanine induced *sdaA* expression. Whether L-isoleucine affects *sdaA* expression is not known.

To define Lrp binding sites in a second way, DNase I footprinting experiments were performed with the DNA fragments carrying the high affinity or low affinity binding site. To delineate the high affinity binding site, a 190 bp *Hin*dIII-to-*Pst*I fragment from plasmid pSU4 was 3'-end labelled at *Hin*dIII site (coding strand), incubated with increasing concentrations of Lrp, and then subjected to DNase I digestion. As shown in

Fig. 15A, a large region of the DNA from nucleotides 432 to 510 was protected by Lrp protein against DNase I cleavage. In the absence of L-leucine, the protection against DNase I could be seen at as low as 8 ng of Lrp protein (Fig. 15A, lanes 3, 4, 5). However, in the presence of L-leucine, the protection required a higher concentration of Lrp (Fig. 15A, lanes 9, 10).

Similar studies were done of DNase I cleavage protection patterns in the low affinity binding site (coding strand; Fig. 15B). In this site, protection was weak and the site protected was much shorter. In the absence of L-leucine, protection by Lrp was seen with 20 ng of Lrp protein (Fig. 15B, lanes 2, 3, 4, 5), while in the presence of L-leucine, no protection was shown even with 320 ng of Lrp (Fig. 15B, lanes 6, 7, 8, 9). There was no evidence for differential affinity of Lrp for subsites within the region. The boundaries for DNase I protection are indicated under the sequence in Fig. 15C.

The transcription start site of *sdaA* is as yet unidentified. To locate the promoter and regulatory region of *sdaA*, I constructed *sdaA-lacZ* operon fusions carrying different parts of the *sdaA* upstream region in plasmid pBR322 (see materials and methods). As shown in Fig. 16, the fusion plasmid pSUZ3 had no significant ß-galactosidase activity. However, plasmids pSUZ2 and pSUZ4 both showed a high level of ß-galactosidase activity suggesting that they carried the *sdaA* promoter. This indicates that the promoter may be located within the 171 bp *Ssp*I-to-*Eco*RV region.

The expression of *lacZ* from plasmid pSUZ2 was shown to be

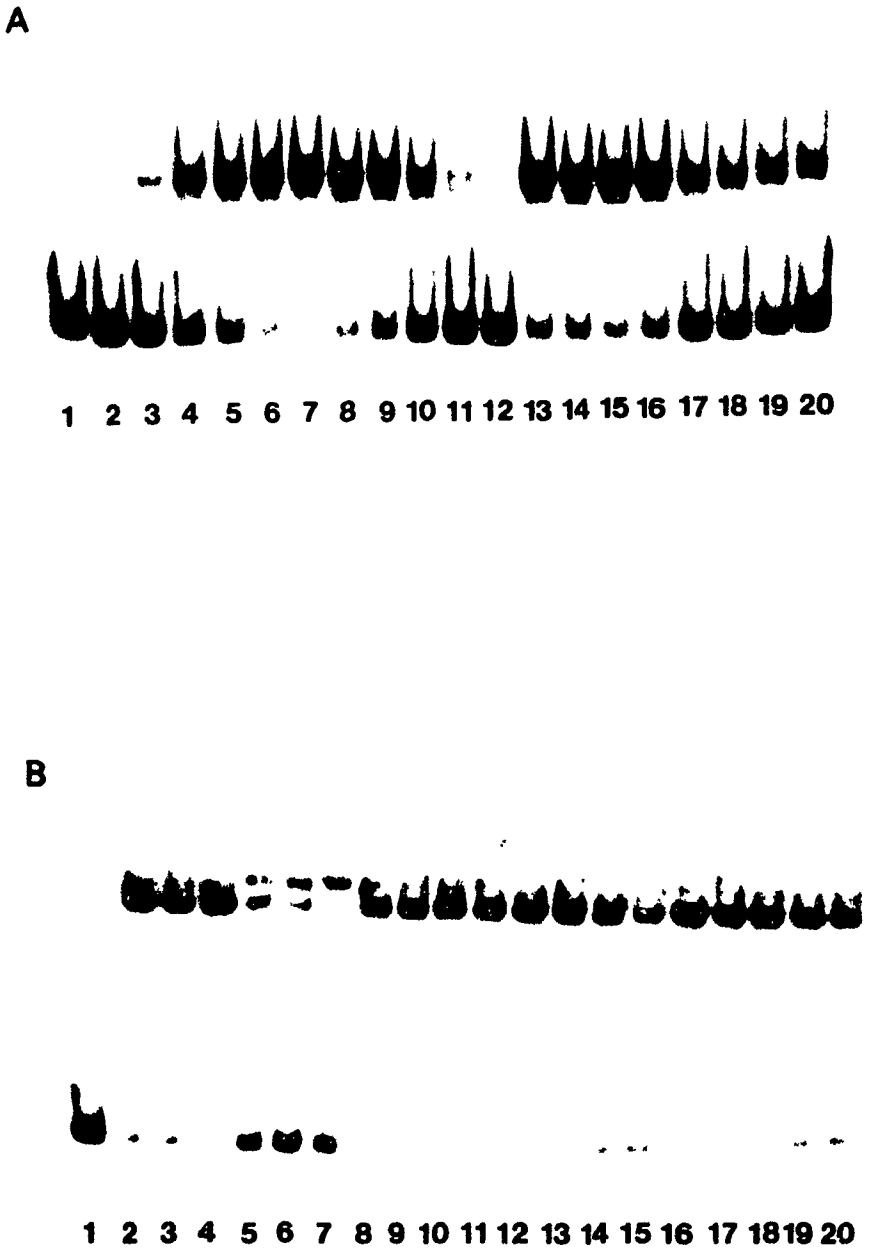


Figure 14. Effect of various amino acids on the binding of Lrp to *sdaA* upstream region.

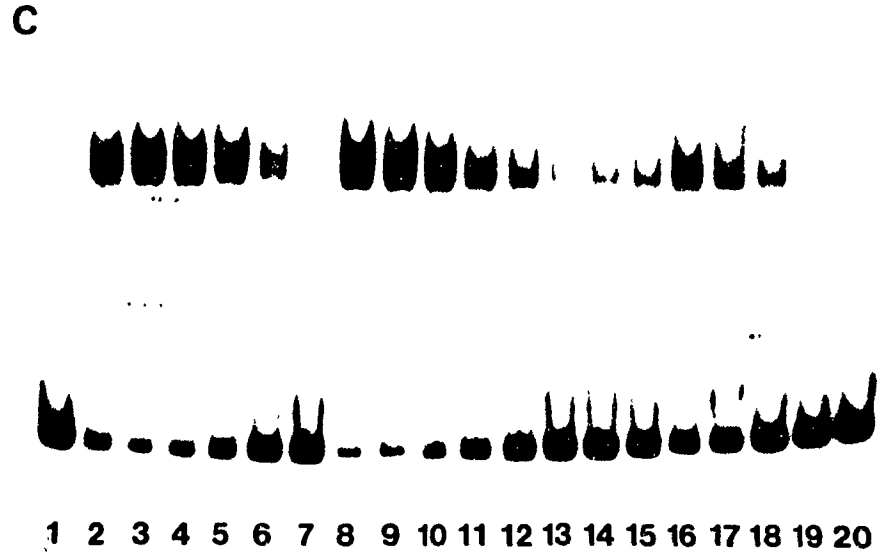
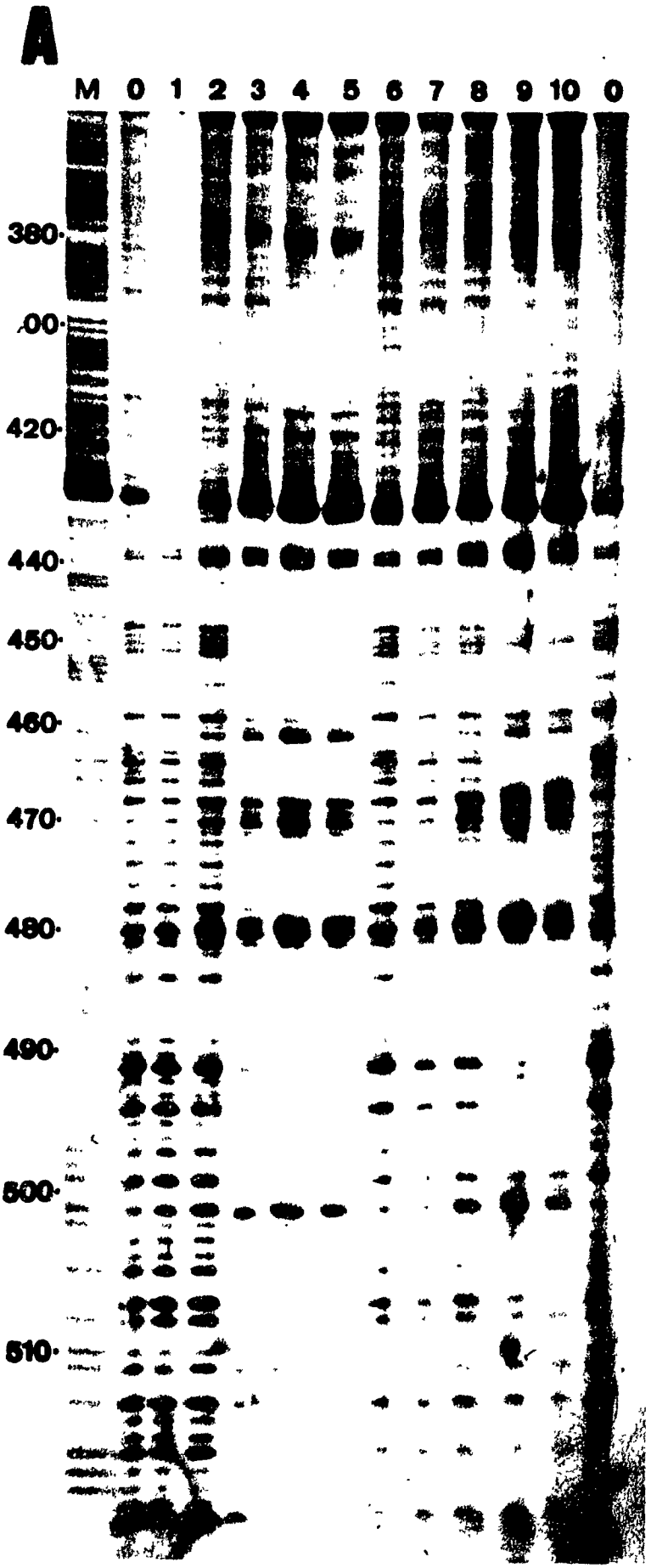
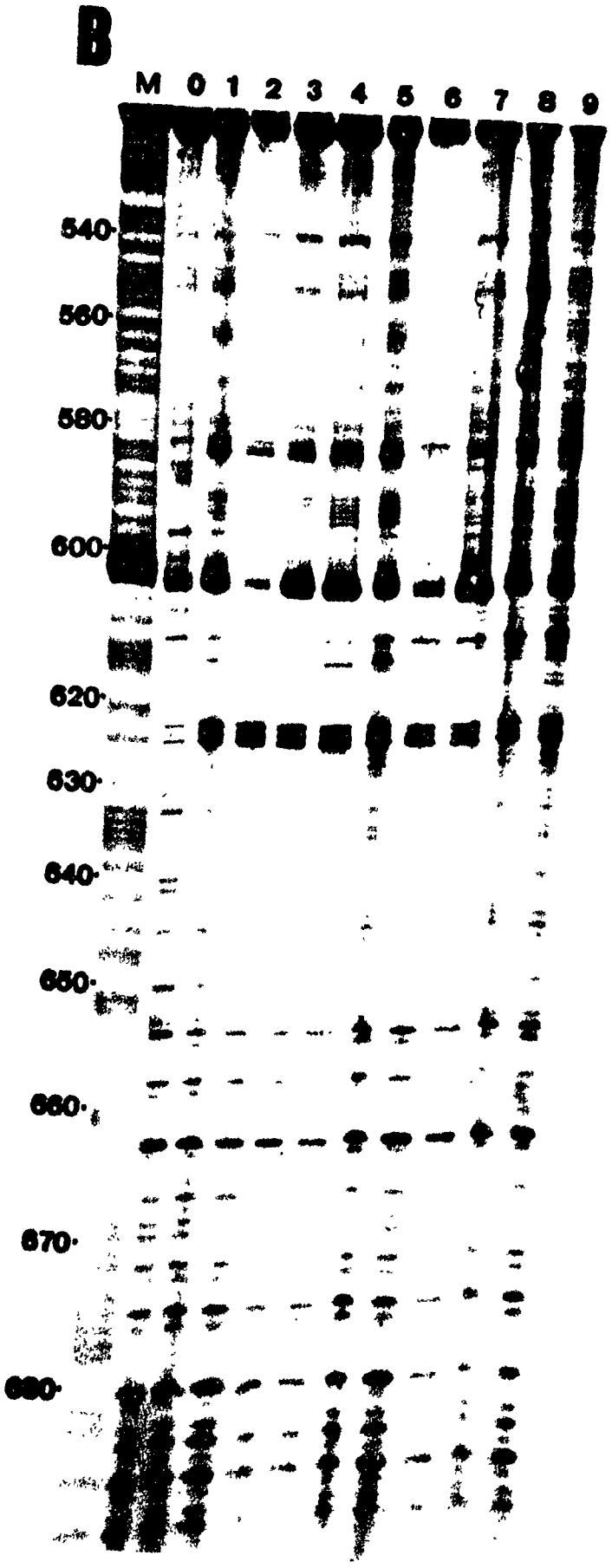


Figure 14. Effect of various amino acids on the binding of Lrp to the *sdaA* upstream region.

Figure 14. Effect of various amino acids on the binding of Lrp to the *sdaA* upstream region. (A) A 361 bp end labelled *Hin*dIII-to-*Pst*I fragment from plasmid pSU2 was incubated with various amounts of Lrp protein and amino acids and analyzed by the gel retardation assay. 0, 5, 10, 20, 40, 80, and 160 ng of Lrp protein were added to lanes 1 through 7, respectively. 80 ng of Lrp protein, and L-leucine 1, 2, 4, 8, 16 mM, glycine 5, 10, 20, 40 mM, L-alanine 5, 10, 20, 40 mM were added to lanes 8 through 20, respectively. (B) Experiments were performed with a 189 bp *Hin*dIII-to-*Pst*I fragment from plasmid pSU4. 50 ng Lrp protein was added to all experiments except lanes 1 in which no protein was added. L-leucine 2.5, 5, 10, 20, 40 mM, glycine 10, 20, 40, 80 mM, L-alanine 10, 20, 40, 80 mM, and L-isoleucine 1.9, 3.8, 7.5, 15, 30 mM were added to lanes 3 through 20 respectively. (C) A 190 bp *Hin*dIII-to-*Pst*I fragment from plasmid pSU3 was used for binding assay. Experiments were performed as in (B) with 500 ng of Lrp protein (in lanes 2 through 20) except that the concentration of L-leucine is 10-times lower, that is, 0.25, 0.5, 1, 2, 4 mM L-leucine was added to lanes 3 through 7, respectively.


A
M 0 1 2 3 4 5 6 7 8 9 10 0
380
00
420
440
450
460
470
480
490
500
510


B
M 0 1 2 3 4 5 6 7 8 9
540
560
580
600
620
630
640
650
660
670
680

```
            .350       .360       .370       .380
5'-ATCTACCGCCGTGGTGATTCACATCGGGTATGGCTGTCCTGGTACGAACA

   .390       .400       .410       .420       .430
   GTATTTTGTATGGGGAATGACCGCAGGCATAATTCGTGAGCTGGCGCTGC

   .440       .450       .460       .470       .480
   AAATTGGTGTGAAACCCTGACTATACTTATCTTTACATCTACAAAACACT

   .490       .500       .510
   ACTTGAGACAATCATCGCAAT-3'


            .520       .530       .540       .550       .560
5'-ATTAGTTAAATCGCGGTTTTTGATTAGTTTAATTCATGTGAATAGTTAAG

            .570       .580       .590       .600       .610
   CCAGTCGCCGCGTTCCCTCTTACACTATGCGCTGTTATTAGTTCGTTACT

            .620       .630       .640       .650       .660
   GGAAGTCCAGTCACCTTGTCAGGAGTATTATCGTGATTAGTCTATTCGAC

            .670       .680
   ATGTTTAAGGTGGGGATTGGTC-3'
```

Figure 15(C). Schematic representation of the DNase I footprints of Lrp bound to the sdaA upstream sequence (coding strand).

Figure 15. DNase I footprint analysis on the *sdaA* upstream region. (A) Digestion pattern of the high affinity binding site. The 206 bp *Xho*I-to-*Eco*RI fragment from pSU4 was labelled at the *Xho*I site, incubated with increasing concentrations of Lrp protein, and subjected to DNase I digestion. Lane M, G+A sequence marker. Lane 0, no Lrp protein was added. Lanes 1 through 5 represent incubations with 2, 4, 8, 16, and 32 ng of Lrp protein, respectively. Lanes 6 through 10 represent incubations with 2, 4, 8, 16, and 32 ng Lrp in the presence of 20 mM L-leucine. (B) Digestion pattern of the low affinity binding site. The 207 bp *Xho*I-to-*Eco*RI fragment from pSU3 was used for the footprinting analysis as in A).. Lane M, A+G sequence marker. Increasing amounts of Lrp protein were added to the incubations shown in lanes numbered as follows: 0, no Lrp; 1, 20 ng; 2 and 6, 40 ng; 3 and 7, 80 ng; 4 and 8, 160 ng; 5 and 9, 320 ng. L-leucine (20 mM) was added to those in lanes 6 through 9. (C) Boundaries of DNase I protection on the two sites upstream of *sdaA* (high affinity binding sequence shown from nucleotides 432 to 510, and the low affinity binding site shown from nucleotides 584 to 622). Solid lines are drawn under protected nucleotides. Numbering is based on the original clone (Su *et al.*, 1989) on which the translation start site is at 645.

Plasmid	Fragment fused to lacZ	B-galactosidase activity in		
		wild type	*lrp*	fold
pSUZ1	*EcoRV* *SspI* *HindIII*	2010	27830	13.8
pSUZ2	*EcoRV* *SspI*	8570	21130	2.5
pSUZ3	*SspI* *HindIII*	5	5	1

Figure 16. Multiple-copy expression of repressed (wild-typ)and de-repressed (lrp) sdaA-lacZ fusions. Plasmids was construct as described in Materials and methods. The location of fragments was shown in Fig. 12.

subject to a 10-fold repression in the *lrp*$^+$ strain. This is consistent with the data described earlier for the chromosomal *sdaA-lacZ* translational fusion in strain MEW22. This observation demonstrates that Lrp repressed the transcription of *sdaA*. In plasmid pSUZ4, the presence of Lrp protein only showed a 2- to 3- fold repression. This suggests that the downstream low affinity Lrp binding site may interact with the upstream region, so that the regulation of *sdaA* expression involves both sites.

IV. Binding of Lrp to *gcv* and *lrp* upstream regions

Both *gcv* and *lrp* are regulated by Lrp but not affected by L-leucine *in vivo*. The expression of *gcv* operon is under the positive control of Lrp while the expression of *lrp* is subjected to negative autoregulation. Since binding of Lrp to particular sites in the upstream regions of several genes regulated by leucine and Lrp together had been demonstrated, I was interested to see whether Lrp also bound to genes not regulated by L-leucine, using gel shift assays as before.

For this purpose, a *Sal*I-to-*Sma*I fragment containing the *gcv* upstream region from pGS146 was cloned into the *Sal*I, *Eco*RV sites of the Bluescript$^+$ polylinker, and the resulting plasmid, pGSU1, used in binding experiments. Plasmid pGSU1 was digested with *Eco*RI and *Pst*I, end labelled, and incubated with purified Lrp protein. The results showed that Lrp could bind to the *gcv* upstream region (Fig. 17, lanes 3-9). The

presence of 20 mM of L-leucine has no effect on the Lrp binding activity (compare lanes 3, 4 with lanes 6, 7). This *in vitro* binding activity is consistent with the *in vivo* data that L-leucine has no effect on the *gcv* expression. Since the expression of *gcv* is repressed by adenine, the effect of adenine on Lrp binding activity was tested. However, the presence of 2 µg/µl of adenine also had no effect on the binding activity (Fig. 17, lanes 8 and 9).

A similar study was made on the *lrp* upstream region carried on plasmid pLR4. As expected, Lrp protein was shown to bind to a 500 bp *EcoR*I-to-*Bgl*II fragment containing the *lrp* upstream region (Fig. 17, lanes 11-16). Again, the presence of 20 mM of L-leucine had no effect on binding activity (Fig. 17, lanes 14-16).

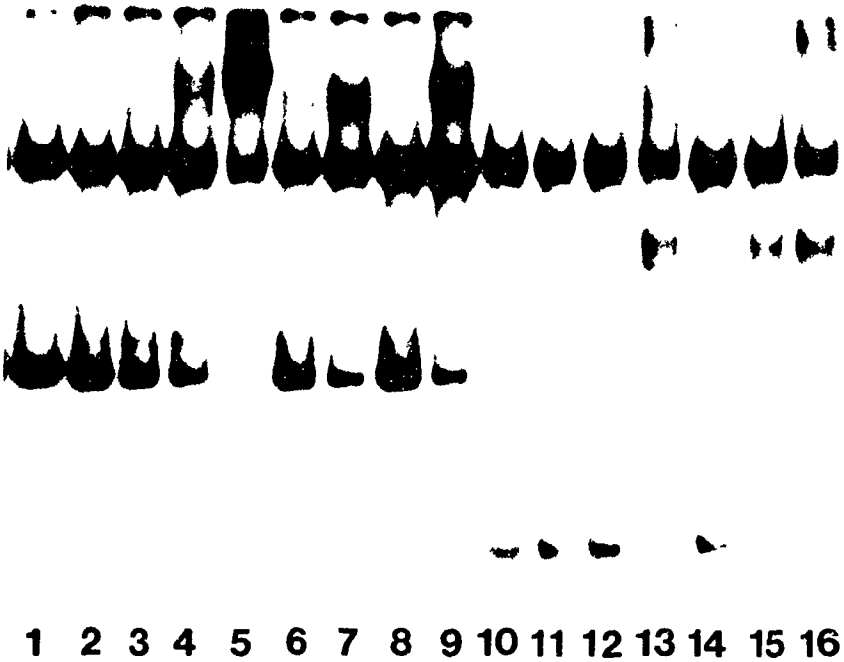


Figure 17. Binding of Lrp to the upstream region of *gcv* and *lrp*.

Figure 17. Binding of Lrp to the upstream region of *gcv* and *lrp*. For the experiments represented in lanes 1-9: plasmid pGSU1 carrying *gcv* was digested with *Eco*RI and *Pst*I, 3'-end labelled with α-^{32}P-dATP by Klenow enzyme, and incubated with 0, 5, 10, 20, 40, 10, 20, 10, and 20 ng of purified Lrp protein respectively. L-leucine (20mM) was added to incubations shown in lanes 6 and 7, and adenine (2 ug/ml) was added to those in lanes 8 and 9. For the experiments shown in lanes 10-16: plasmid pLR4 carrying *lrp* was digested with *Eco*RI and *Bgl*II, end labelled, and incubated with 0, 5, 10, 40, 10, 20, and 40 ng of purified Lrp protein, respectively. L-leucine (20mM) was added in lanes 14-16.

DISCUSSION

The work reported in this thesis, as well as observations from other labs, has resulted in the discovery of a new global response in *E. coli*, the leucine/Lrp regulon. The experiments reported here include the isolation and characterization of an *lrp* mutation, the identification of members of the leucine/Lrp regulon, the beginning of the study of the physiological effects of leucine/Lrp regulon, and studies of the *in vitro* interactions between Lrp and some of its target promoters. In this discussion, I will consider the techniques used for this study, the possible physiological role of Lrp and the regulatory mechanism(s) by which Lrp acts.

I. The Study of the *lrp* Gene

I-1. Insertional Inactivation of the *lrp* Gene

I began this work by trying to isolate an insertion mutation resulting in overexpression of L-SD. It was known that L-SD activity was increased in the presence of L-leucine and/or glycine in the growth medium. This increase was considered to be (probably) due to the induction of expression of the structural gene for L-SD, *sdaA* [Su *et al.*, 1989]. For instance, one might consider that the transcription of *sdaA* is under negative control, and that this repression could be released by L-leucine or glycine. If this were true, a mutation which caused a deficiency of repressor should result in the constitutive expression of L-SD and allow the mutant to use L-serine as sole carbon source.

Several mutants which could use L-serine as sole carbon source had been isolated in this laboratory prior to the start of this work. One of them, the *ssd* mutant, showed high L-SD activity, and permitted growth on L-serine. It has been suggested that the *ssd* gene is identical to *cpxA*. Were this true, the fact that a strain carrying a *cpxA*::Cm^r mutation could not grow on L-serine would exclude the possibility that the *ssd* gene product is a repressor of *sdaA* gene expression.

It seemed worth looking further for the *sdaA* repressor. I chose to do this by insertion mutagenesis, which has been widely used for isolation of null mutations. Previous work in this laboratory took advantage of Mud [Casadaban and Cohen, 1979] and λp*lac*Mu [Bremer *et al.*, 1984; 1985] technology to isolate insertion mutations. These insertions provided a powerful tool to study the regulation of gene expression since they put the *lacZ* gene under the control of the promoter of the particular gene of interest.

The transposon Tn*10* has also been used for isolating insertion mutations in regulatory genes. It has advantages of small size, high insertion rate and stability of its insertions. Because of these characteristics, an insertion mutation which allowed growth on serine could be easily moved into a strain carrying an *sdaA*::*lacZ* mutation (created by λp*lac*Mu insertion) and the effect of the Tn*10* insertion on L-SD expression could be studied via ß-galactosidase assays.

The selection for mutations affecting *sdaA* regulation is complicated by the fact that other mutations apparently

unrelated to L-SD may also permit the use of L-serine as carbon source. I isolated 70 independent tetracycline-resistant mutants able to grow on serine, and most showed high L-SD activity. To be sure that this high activity was due to the insertion, and not due to a concomitant selection of mutations in *ssd*, [Newman *et al.*, 1982b], the insertion from one of strains was transduced into MEW1 by selecting tetracycline resistance. The tetracycline-resistant recipient had the same phenotype as the donor strain. It grew on L-serine, and showed high L-SD in glucose-minimal medium grown cells. This mutation was termed *rbl-1*::Tn*10*, for regulation by leucine. It has since been renamed *lrp*::Tn*10*.

I-2. The *rbl* Gene Product Might Be the Regulator of a Global Response to L-Leucine

As shown in Table 3 and 4, the *rbl* mutation increased the activity of L-SD and TDH sevenfold. The presence of L-leucine in growth medium stimulated the wild-type activities five- and three-fold, respectively, but had essentially no effect on the mutant levels (Table 9). This suggests that the mutant lacks the factor through which the L-leucine effect is mediated.

The *rbl* mutation resulted in increased expression of the gene coding for L-SD. This was shown by studies of ß-galactosidase synthesis from an *sdaA*::*lacZ* fusion, which places ß-galactosidase synthesis under the control of the promoter of *sdaA*, the structural gene of L-SD [Su, 1991]. The ß-galactosidase activity was increased seven-fold in an *rbl*

mutant (Table 3), demonstrating that Rbl (Lrp) decreases expression of that gene.

Similar studies showed that the *rbl* mutation strongly decreased transcription of the *serA* and *IlvIH* operons, coding for phosphoglycerate dehydrogenase and acetolactate synthase III, respectively (Table 8 and 9). Both of these operons are known to be repressed in the wild-type strain by the presence of exogenous L-leucine [McKitrick and Pizer, 1980; Squires *et al.*, 1981].

In summary, the *rbl*::Tn*10* mutant, though selected for its effects on L-serine degradation, was also affected in the activities of enzymes involved in several other metabolic pathways. In all of the cases first studied, the enzyme activities were similarly affected in the wild-type strains by the presence of L-leucine, and indeed the phenotype of the mutant was similar to that of wild-type strains cultivated with L-leucine. These observation led us to name the gene *rbl* (for regulation by leucine) and to suggest that it codes for the regulatory protein of a new global response, the leucine regulon. The fact that the mutant studied carried a Tn*10* insertion, presumably resulting in the total absence of any active product, suggests that in these cases the effect of exogenous L-leucine is to inactivate the *rbl* gene product.

It was previously suggested that L-leucine might trigger increased expression of degradative functions in rich medium, with a concomitant decrease in biosynthetic functions [Newman *et al.*, 1976]. The scheme is clearly less general than this,

since not all catabolic functions are affected by the *rbl* mutation or by the presence of exogenous L-leucine. Activity of at least two amino acid degradative enzymes, L-tryptophanase and L-proline oxidase, were unaffected by either the presence of the *rbl* mutation or by the presence of L-leucine in the growth medium. Thus the *rbl* can control at least a subset of degradative enzymes.

I-3. Map Location and Function Indicate that *rbl* is Identical to *lhb*, *oppl*, and *livR*

The *rbl*::Tn*10* mutation was located near 20 min on the *E. coli* genetic map, first by conjugation, and then by P1-mediated transduction. No gene then known to be located near 20 min seemed likely to be identical to this new gene, i.e. to regulate L-SD synthesis. It seemed then that *rbl* might be a new member of the group of genes affecting L-serine metabolism. However it soon became clear that several other mutations in this gene have been isolated, under a variety of names and functions.

I located the *rbl* gene within the Kohara phage collection. A subclone from Kohara phage λ1F10 complemented the deficiencies of the *rbl*::Tn*10* mutation. Southern blot analysis demonstrated that the 2.6 Kb *Eco*RI-*Hin*dIII fragment from this plasmid hybridized to an interrupted gene in the *rbl*::Tn*10* mutant. This indicated that the Tn*10* element was inserted within this fragment. Since subclones carrying this 2.6 Kb fragment (or less) complemented the *rbl* mutation, it is clear that the *rbl* gene is located within this region.

In the same region is found the *lhb* gene, identified by its effects on expression of the *IlvIH* operon [Ricca *et al.*, 1990]. J.M. Calvo suggested to us that the *lhb* gene might encode a general mediator of L-leucine effects. Several observations presented here indicate that *rbl* and *lhb* refer to the same gene. The two mutations *rbl* and *lhb* , have quantitatively similar effects on L-SD and TDH activity and on *sdaA*, *serA*, and *IlvIH* transcription (Table 13). The hybridization analysis also showed that the 2.6 Kb *Eco*RI-*Hin*dIII fragment from Kohara phage λ1F10 hybridized to an interrupted gene in strain CV1008 (*lhb*::Tn*10* mutant).

Several other genes in the 20 min region have been identified as regulators of L-leucine-sensitive operons and have been shown to be identical to *rbl*/*lhb*. These include *IlvR* [Anderson *et al.*, 1976] and *oppI* [Andrews *et al.*, 1986; Andrews and Short, 1986]. The *IlvR* gene codes for a regulatory protein affecting high-affinity transport of branched chain amino acids through the LIV-I transport system, encoded by the *IlvJ* and *IlvKHMGF* operons [Nazos *et al.*, 1984]. The expression of both operons is repressed in the presence of L-leucine [Anderson and Oxender, 1977]. It was recently reported that *IlvR* and *lrp* (*rbl*/*lhb*) are allelic and the suggestion was made that Lrp acts directly or indirectly to repress *IlvJ* and *IlvK* expression, and that L-leucine is required for this repression [J. M. Calvo, personal communication].

The *oppI* gene in *E. coli* represses the expression of the *oppABCDF* operon, which encodes a transport system primarily for tripeptides [Andrews and Short, 1985]. The expression of

this operon was also shown to be induced by the presence of L-leucine or L-alanine [Andrews *et al.*, 1986; Andrews and Short, 1986b]. The identity of the *lhb* and *oppI* genes was demonstrated by DNA sequencing studies [Platko *et al.*, 1990; Austin *et al.*, 1989]. It is clear that *lhb*, *livR*, *oppI*, and *rbl* are all the same gene. We have agreed that the designation *lrp* (leucine-responsive regulatory protein) is to be used in the future for all of these.

II. Isolation of Insertion Mutants in Genes Controlled by L-Leucine/Lrp Using λp*lac*Mu-Mediated *lacZ* Fusion

The presence of L-leucine in growth medium has long been known to affect the expression of a number of enzymes not directly involved in L-leucine metabolism. As mentioned above, The addition of L-leucine to the culture medium repressed expression from *livIH*, *serA*, *livJ*, and *livK* genes and stimulated expression of other genes including *sdaA*, *tdh*, *lysU*, and *oppABCDEF*. In all these cases, the effect of L-leucine is mediated by Lrp.

It seemed interesting to find out how extensive the set of L-leucine-regulated genes in *E. coli* might be, and whether in each case the effect of L-leucine is mediated by Lrp. This question stimulated a systematic search for L-leucine-controlled genes.

II-1. The Techniques Used for This Study

Genes whose expression is affected by any particular factor in the environment can be identified either by

examining the gene products directly on two-dimensional gels or by studying expression of reporter genes with any one of a variety of gene fusion techniques. These two approaches have been used successfully to study a variety of stimuli which result in global changes in protein synthesis.

Two-dimensional gels, which allow separation of more than 1,000 proteins [O'Farrel, 1975], have been used to particular advantage in defining the nature of the heat-shock response in *E. coli* [Neidhardt *et al.*, 1984]. This is the most direct way to determine the changes in the pattern of protein synthesis, and can be applied to both essential and nonessential proteins. The technique has also been used to study changes in protein synthesis in response to glucose starvation [Jenkins *et al.*, 1988], shifts in available carbon sources, or changes between aerobic and anaerobic growth [Pedersen *et al.*, 1978; Smith and Neidhardt, 1983a; 1983b].

Use of *lacZ* gene fusions is equally valuable in the study of regulation of gene expression. Based on the exact location of the fusion joint, two types of fusion, protein fusion and operon fusion, can be constructed. In the protein fusion, a *lacZ* gene lacking both transcription and translation initiation signals is fused in frame to the coding sequence of a target gene; while in a operon fusion, a promoterless *lacZ* gene containing its own translation start site is fused to an exogenous promoter. The expression of ß-galactosidase from both types of fusion is under the control of the promoter of target gene. Both types of fusion can be used to study transcriptional regulation of any gene. The protein fusion

also can be used to study the translational control. Taking the advantage of transposable genetic elements, it is possible to move the *lac* genes to any position on the chromosome. Since the insertion in a gene will lead to its inactivation, this method cannot be used to identify essential genes, nor is it useful to study genes subject to autoregulation.

The defective Mu phage, Mu d1(Ap *lac*) was first used for genetic identification of genes subject to a common regulatory signal by Kenyon and Walker [Kenyon and Walker, 1980]. The Mu d phage carried a promoterless *lac* gene and a selectable antibiotic resistance gene replacing many of the Mu lytic functions [Casadaban and Cohen, 1979]. When the phage inserts downstream from a target promoter, the ß-galactosidase activity will reflect the transcriptional-regulating properties of that promoter. Using this technology, a set of genes whose expression is induced by DNA-damaging agents and regulated by the *recA* and *lexA* gene products were identified [Kenyon and Walker, 1980]. A similar approach was used for study the cells response to phosphate starvation and resulted in the identification a number of new genes regulated by phosphate [Wanner and McSharry, 1982].

I decided to use λp*lac*Mu phage for identification of L-leucine-controlled genes since it has several advantages compared with Mu d phage [Bremer *et al.*,1984]. Efficient Mu transposition requires four elements including both ends (s and c) of the Mu phage and the products of genes A and B. λp*lac*Mu carries the two Mu ends and a promoterless *lac* gene

next to the s end. It can insert like Mu into genes nonspecifically using the Mu transposition machinery provided by a helper phage carrying the Mu A and B genes. However, since λp*lac*Mu prophages are stable and are not induced at high temperature, the fusions isolated with λp*lac*Mu can be subjected to a variety of genetic manipulations which the Mud system cannot support.

II-2. The Leucine/Lrp Regulon

There are a large number of proteins whose synthesis is regulated by Lrp. In the screening of random *lacZ* fusions, 27 were found to respond to exogenous L-leucine. All of these L-leucine-controlled genes were also shown to be regulated by Lrp. Furthermore, L-leucine had no effect on gene expression in the absence of the *lrp* gene product in all these insertions except for those inserted in L-leucine-biosynthetic genes. These observations suggested that in most cases the regulation of L-leucine is mediated by Lrp. The effects of L-leucine and Lrp on the expression of L-leucine-biosynthetic genes were additive; in these insertions L-leucine limitation stimulated expression even in absence of Lrp. This might be due to control by transcription attenuation of *leu* operon, or to local control by another leucine sensitive regulatory protein specific to this operon.

Among genes identified later, there were some whose expression was controlled by Lrp but not affected by L-leucine. To include these, the regulon was renamed and is now

called the leucine/Lrp regulon. The system used for isolating insertion mutants in this study would not allow us to isolate fusions to promoters regulated in this way, nor could we isolate those which produced auxotrophs (with the exception of those requiring L-leucine).

How large might the leucine/Lrp regulon be? Since several known members of the regulon (such as *sdaA*) were not isolated, I probably have not even isolated insertions in all the genes which could be identified by the screening I used. Based on the two dimensional gel analysis, Ernsting and coworkers demonstrated that the expression of more than 30 polypeptides were affected by the absence of Lrp, and reported that many of those were regulated by Lrp but not by leucine [Ernsting *et al.*, 1992]. This number is likely to be a low estimate. The expression of many of the fusions varied only 2- to 3-fold with Lrp, and these would probably not be seen by comparison of 2-D gels. Moreover some proteins were not resolved on 2-D gels, due to their large molecular weight or their basic isoelectric points [Ernsting *et al.*, 1992]. This suggests that the description of genes regulated by Lrp has just begun.

III. Genes Regulated by Lrp

The members of the leucine/Lrp regulon identified so far are shown in Table 23. Lrp activates the expression of some genes, and represses that of others. The known genes which are activated include *IlvIH*, *serA*, *leu*, *gltD*, *fan*, *pap*, and *gcv*. Those that are repressed include *sdaA*, *tdh*, *lysU*, *kbl*, *oppABCDF*, *IlvJ*, *IlvKHMGF* and *lrp* itself.

III-1. Genes Regulated by Lrp and L-Leucine

As mentioned above, the genes regulated by Lrp code for proteins catalyzing a wide variety of reactions. The gene products of leucine/Lrp regulon identified so far are involved in amino acid biosynthesis (*serA*, *leu*, *IlvIH*, and *gltD*), amino acid and oligopeptide transport (*IlvJ*, *IlvKHMGF*, and *oppABCDF*), amino acid degradation (*sdaA*, *tdh* and *kbl*), amino acyl-tRNA synthesis (*lysU*), C_1 metabolism (*gcv*) and pili formation (*fan* and *pap*) [D. A. Low, personal communication].

For most of the Lrp-regulated genes, the presence of L-leucine partially alleviated the effect of Lrp. Some genes (class 1, Table 15) whose expression is activated by Lrp are turned off by L-leucine. In most cases, e.g. *serA*, *fan* and *IlvIH*, the effect of L-leucine is less severe than the effect of an absence of Lrp protein. Conversely, genes repressed by Lrp (class 2, Table 15) are induced by L-leucine, but again in most cases, e.g. *sdaA* and *tdh*, the effect of L-leucine is not as great as that of an *lrp* mutation.

In some genes, regulation by Lrp required the presence of L-leucine, rather than being reversed by it. This joint action results in activation of several genes (class 3, Table 15), and repression of at least one. L-leucine is required for Lrp to repress *IlvJ* and *IlvK* expression (class 4 of Table 15) [J. M. Calvo, personal communication].

These observation suggest that Lrp may exist in two conformations, one in association with L-leucine, and the other without it. Either conformation could act as an

TABLE 23. Genes regulated by Lrp

Gene	Map location (min)	Function	Effect of Lrp	Effect of L-leucine
IlvIH	2	Ile, Val	Activator	Antagonist
serA	63	Ser	"	"
gltD	70	Glu	"	"
fan	?	pili formation	"	"
sdaA	41	L-ser deaminase	Repressor	Antagonist
tdh-kbl	81	Thr deaminase	"	"
oppABCDF	28	Oligopept. uptake	"	"
lysU	94	Lysyl-tRNA synth.	"	"
sdaB	60	L-ser deaminase	Activator	Required
IlvJ/K,	76	Leu, Val, & Ile uptake	Repressor	Required
gcv	63	Gly cleavage	Activator	None
pap	?	pili formation	"	"
lrp	20	Regulator	Repressor	None
ompF	21	Porin	Activator	?
leu	2	Leu	"	?
ompC	48	Porin	Repressor	?

activator at some genes and as a repressor at others. For genes in class 3 and 4, free Lrp seems to be inactive. When L-leucine is bound to it, it assumes an state and can either act as an activator or a repressor, depending on the particular structure of the gene. On the other hand, for genes in class 1 and 2, the free Lrp conformation is active.

It is not known whether there is any gene of the leucine/Lrp regulon for which one Lrp conformation activates and the other represses, analogous to AraC function at the *araBAD* promoter [Lobell and Schleif, 1991]. However, in one of the unidentified class 2 mutants, CP25, the ß-galactosidase level in the absence of Lrp is intermediate between those in the *lrp*$^+$ strain growing with and without L-leucine. This suggests that CP25 might be regulated in this way. That is, in the absence of L-leucine, free Lrp would repress expression, while with L-leucine, the leucine-Lrp complex would activate.

III-2. Genes Regulated by Lrp but not by L-Leucine

The system used in this study for screening Lrp-regulated genes depends on a response to L-leucine. No fusion of *lacZ* to L-leucine-independent promoters would be isolated by this screen. However, studies of λp*lac*Mu insertions in *gcv* operon and *lrp* gene indicated that both of these, and doubtless other genes, are regulated by Lrp with no effect of exogenous L-leucine.

Transcription of the gcv operon, whose products are involved in the conversion of glycine to C1-THF, is totally

dependent on Lrp, but essentially unaffected by L-leucine. In this operon, the activation of Lrp is physiologically significant, so much so that a strain carrying a *lrp* mutation is phenotypically *gcv*$^-$.

It has been recently shown that the methylation blocking factor gene (*mbf*) is identical to *lrp* [D. A. Low, personal communication]. This gene product has been shown to be required for the transcription of *pap* operon [Braaten *et al.*, 1991], and this activation is not affected by L-leucine.

Transcription of *pap* depends on the methylation of 2 upstream GATC sites by deoxyadenosine methylase (Dam). Removal of the methyl group from the upstream GATC site is required for the transcription of this operon [Braaten *et al.* 1991). Lrp binding to this site is thought to block methylase access, particularly since it was shown that Lrp binds to a *pap* DNA fragment containing both GATC sites. The suggestion was then made that binding of Lrp inhibits Dam methylation by steric hindrance [D. A. Low, personal communication].

The remaining class of Lrp-regulated genes are those which are negatively regulated by Lrp in a L-leucine-independent fashion, e.g. the *lrp* gene itself. The expression of the *lrp* gene is repressed by its product, Lrp, with no effect of exogenous L-leucine.

Analysis by 2-D gels indicated that the synthesis of glutamine synthetase (GlnA), the small subunit of glutamate synthetase (GltD), and OmpF are positively regulated by Lrp, and synthesis of OmpC negatively regulated, all in a L-leucine-independent fashion [Ernsting *et al.*, 1992]. Since

the 2-D gel analysis probably can not detect less than 2- to 3-fold regulation by L-leucine, the expression of some of these genes may in fact also be regulated by L-leucine. This is clearly true for the *gltD* gene, where no L-leucine effect is noticed on the 2-D gels, but a 2-fold decrease was seen when glutamate synthase was assayed directly in L-leucine-grown cells. On the other hand, this effect is very minor, when compared to the 1000-fold lower level seen in the *lrp* mutant [Ernsting *et al.*, 1992].

IV. Metabolic Effects of the Leucine/Lrp Regulon

Lrp regulates the expression of a large number of genes, by a factor of between 2- and 1000-fold. Inspection of the nature of the genes regulated by Lrp suggests that, with some exceptions, Lrp activates expression from biosynthetic genes and represses expression from genes whose products are involved in degradation of amino acids and transportation of amino acids and oligopeptides. This section will examine the physiological and metabolic consequences of a loss of *lrp* function.

IV-1. A Mutation in *lrp* Suppresses the Metabolic Deficiency in a *metK* Mutant

It was long believed that the *metK* mutations, which result in the loss of almost all S-adenosylmethionine synthase activity, had very little physiological effect on the cells. However, the work in this thesis demonstrates that the *metK* mutations have more profound effects on cell metabolism than

previously expected, and that these effects can be greatly reduced by inactivation of the *lrp* gene product. In fact, the effects of *metK* mutations are so severe that the commonly studied *metK* mutants carry a second mutation, often an *lrp* allele, which suppresses the extreme effects of *metK*.

The *metK* strain used in this study, RG62, is the prototype *metK* strain, isolated by virtue of its ethionine resistance and shown to have less than 3% of wild-type SAM synthetase activity when grown in the absence of L-leucine [Greene *et al.*, 1973]. It was recently reported that this strain, RG62, has an unusual ability to degrade L-serine, and this was attributed to the *metK* mutation [Matthews and Neidhardt, 1989]. Our results showed that strain RG62 indeed had unusual L-serine degradation properties: it could grow with L-serine as sole carbon source and had high levels of L-SD. However, this was not due to the *metK* mutation; since these properties did not follow when the *metK* mutation from RG62 was transduced into another strain.

The *metK* transductants, unlike strain RG62, grew very slowly in glucose minimal medium unless provided with L-leucine or suppressed by an *lrp* mutation. In fact, it was shown that RG62 harbours a *lrp* mutation as judged by its location near 20 min, and the facts that it increases expression of L-SD, decreases expression of *serA*, and permits growth on L-serine. Furthermore, during growth in glucose, the slow-growing *metK* transductants rapidly accumulated fast-growing derivatives, many of which had acquired the ability to grow on L-serine as a result of the overproduction of L-SD,

consequent on a mutation shown to be cotransducible with *serC*. These observations suggested that during the purification of the original ethionine-resistant mutant, RG62 accumulated an *lrp* mutation, allowing it to grow much faster than the original *metK* strain. It is clear then that the RG62, previously characterized as a *metK* mutant, is in fact a *metK lrp* double mutant.

This is not the only mutant of methionine metabolism which shows alterations in serine metabolism. Other strains derepressed for methionine biosynthesis, including *metJ* mutant, were shown to grow slowly with L-serine as carbon source [Brown *et al.*, 1990]. However, this ability depended on overproduction of the *metC* gene product, L-cystathionase, and was not accompanied by an increase in L-SD [Brown *et al.*, 1990].

Other phenotypes ascribed to the *metK* mutation may also have to be reconsidered. It was reported that strain RG62 required L-serine and L-isoleucine for growth at 44^0C [Matthews and Neidhardt, 1989]. In this study, we showed that the *lrp*::Tn*10* mutant required L-serine for growth at high temperature. This suggests that the temperature-sensitivity of RG62 is due to the *lrp* mutation. L-SD was known to be overproduced at high temperature [Newman *et al.*, 1985]. This effect is additive with the stimulation due to the *lrp* mutation, so that the rate of L-serine degradation should be quite high in the *lrp* mutant when grown at high temperature. This, coupled with the observation that the mutation in *lrp* resulted in decreased *serA* expression, probably is sufficient

to explain the high-temperature L-serine requirement. The L-isoleucine requirement could result from the known sensitivity of *E. coli* to L-serine, which is relieved by L-isoleucine [Cosloy and McFall, 1970; Uzan and Danchin, 1978].

Another intriguing mutant is the temperature-sensitive *htrA*::Tn*10* described by Lipinska *et al.* [Lipinska *et al.*, 1989], which, required either L-serine or L-leucine for growth in minimal medium at temperatures above 42^{0}C. Whether this phenotype related to *lrp* is unclear.

The lysyl-tRNA synthetase II (product of the *lysU* gene), a heat shock protein that is normally expressed only at elevated temperature, was reported to be constitutively expressed in strain RG62, but this property was not corrected by the presence of a cloned *metK*$^+$ gene [Matthews and Neidhardt, 1988]. Since the synthesis of lysyl-tRNA synthetase II is induced by the addition of L-leucine to the growth medium, it seemed likely that this constitutivity resulted from the *lrp* mutation in RG62 rather than the *metK* allele. A direct assay of enzyme activity showed that the synthesis of lysyl-tRNA synthetase II is derepressed in an *lrp* mutant but not in a *metK* mutant [I. N. Hirshfield]. The 2-D gel analysis also indicated that the expression of *lysU* at 37^{0}C is associated with the *lrp* allele [Ernsting *et al.*, 1992].

The extremely slow growth rate of *metK* mutants may be a result of a low pool of SAM leading to problems with methylation reactions. One might expect that Lrp normally represses the synthesis of an enzyme which can substitute for SAM synthetase, either a second synthetase, such as the

recently described *metX* gene product [Satischandran and Markham, 1990], or an alternative methyl donor. It was reported that the *metX* could be expressed only in rich medium [Satischandran and Markham, 1990]. The effect of the mutation in *lrp* could be understood if it resulted in the expression of *metX* in glucose minimal medium, or in the activation of a gene whose product permitted synthesis of methyl groups from a new source. It is also possible that the *lrp* mutation could bypass the need for some critical methylation reaction.

IV-2. Lrp as a Major Regulator of C1 Metabolism

Whatever its other, and possibly more general, metabolic roles may be, Lrp has a profound effect on C1 metabolism. One carbon (C1) metabolism in *E. coli* is based on the synthesis of single carbon tetrahydrofolate-derivatives. When *E. coli* is grown in glucose minimal medium, C1 units are derived mainly from L-serine, either directly, by serine hydroxymethyl transferase (SHMT, the gene product of *glyA*) with the concomitant synthesis of glycine, or indirectly from glycine via the glycine cleavage enzymes (the products of the *gcv* operon) [Newman *et al.*, 1974]. These routes are balanced according to the ratio of C1 to glycine required by the cell.

Lrp is clearly involved in maintaining this balance. The presence of L-leucine in the growth medium would decrease the amount of L-serine available, both by increasing deamination of L-serine, and decreasing the biosynthesis of L-serine. Despite the decrease in L-serine, the supply of glycine could still be maintained. Though the glycine coming from L-serine

via SHMT may be reduced, additional glycine might be derived from L-threonine via threonine dehydrogenase (TDH) and 2-amino-3-ketobutyrate CoA ligase (KBL), whose expression is induced by exogenous L-leucine and inhibited by Lrp. Glycine would then be made from both serine and threonine, and C1 units would be made from both L-serine and glycine.

The *lrp* mutation, with its very low level of glycine cleavage enzymes, is physiologically *gcv*$^-$, unable to use glycine as nitrogen source or as source of C1 units. In the *lrp* mutant, then, the cells must become entirely dependent on SHMT for the production of C1 units, unless it has some other pathway for C1 production. However, the fact that an *lrp glyA* double mutant could not grow in glucose minimal medium with glycine without the addition of the end-products of C1 metabolism suggests that no other pathway contributes much C1 in an *lrp* background. Since the cell requires much more C1 units than glycine molecules, and the high level of TDH in the *lrp* mutant suggests that it may derive much of its glycine from L-threonine, the mutation in *lrp* must encounter difficulty in balancing its C1 units and glycine production. How it solves this problem is still unclear.

IV-3. Lrp Regulated L-Serine Metabolism

L-Serine plays an important role in cell metabolism. Apart from its obvious use in protein biosynthesis, L-serine is the major source of C1 units, and the direct or indirect precursor of glycine, L-cysteine, L-methionine, L-tryptophan and purines [Newman and Magasanik, 1963; Kredich and Tomkins, 1966; Tran

et al., 1983; Yanofsky, 1960]. However, L-serine is also known to be toxic to *E. coli* and other organisms [Cosloy and McFall, 1970; Isenberg and Newman, 1974; Hama *et al.*, 1990, Uzan and Danchin, 1978]. Therefore, there must be some mechanism(s) for cell to control its internal level of L-serine carefully.

In wild-type *E. coli* grown in glucose minimal medium, the synthesis of L-serine starts from 3-phosphoglycerate, an intermediate of the Embden-Meyrerhof pathway. The three reactions involved in L-serine biosynthesis are catalyzed by 3-phosphoglycerate dehydrogenase, 3-phosphoserine aminotransferase, and 3-phosphoserine phosphatase, coded by *serA*, *serC* and *serB* respectively [Stauffer, 1987]. However, some mutants could derive their L-serine from L-threonine [Fraser and Newman, 1975], using the conversion of threonine to glycine and acetyl coenzyme A catalyzed by threonine dehydrogenase (TDH) and α-amino-ß-ketobutyrate ligase (KBL) [Chan and Newman, 1981]. In these mutants, L-serine was synthesized from glycine via the serine hydroxymethyltransferase reaction [Scrimgeour and Huennekens, 1962].

Both biosynthesis and degradation of L-serine are regulated by leucine/Lrp. The mutation in *lrp* decreased the transcription of *serA* 6-fold and increased the L-SD activity 7- to 10-fold. The *lrp* mutant also has a 7-fold increase in TDH activity. However, since the *lrp* mutant is physiologically *gcv*$^-$, it cannot derive its L-serine from glycine. One might expect then that the *lrp* mutant, grown in glucose-minimal medium, has a low level of internal L-serine, and that the

cells are partially starved for L-serine. Indeed, the addition of L-serine to the growth medium actually increased the growth rate of the *lrp* mutant at 37°C and exogenous L-serine was absolutely required for this mutant to grow at elevated temperature. This L-serine starvation is probably caused more by the decrease in 3-phosphoglycerate dehydrogenase than by the increase in L-SD activity, since a strains producing higher levels of L-SD, due to the presence of the *sdaA*$^+$ gene on a high-copy plasmid or due to a mutation in *ssd*, grew well at 42°C without the addition of L-serine.

Deamination of L-serine by *E. coli* involves not only one but two L-serine deaminases encoded by two different genes, *sdaA* and *sdaB* [Su and Newman, 1991]. The expression from both *sdaA* and *sdaB* promoters is regulated by Lrp and L-leucine, but in different way. The expression of *sdaA* is repressed by Lrp and this repression is partially alleviated by L-leucine; as shown by the fact that the ß-galactosidase level from the *sdaA*::*lac* fusion was increased 7-fold in the *lrp* mutant strain. On the other hand, full expression of *sdaB* in glucose minimal medium required both Lrp and L-leucine [Su, 1991]. The expression of a gene(s) near *sdaB*, mutated in strains CP41 and CP52, and perhaps a regulatory gene essential for *sdaB* expression, is regulated by Lrp in a similar way. It is possible that the regulation of *sdaB* by Lrp is a secondary effect of this regulatory protein. However, Lrp had almost no effect on the high level expression of either *sdaB* or this regulatory gene. It has since become clear in any case that the major influence on the expression of these genes is

mediated by cyclic-AMP and catabolite activator protein [Shao, personal communication].

IV-4. Lrp-Regulated Ammonia Assimilation

E. coli can use a variety of organic nitrogen-containing compounds as the sole nitrogen source, though growth with these organic nitrogen sources is slower than with ammonium sulfate. Cells can derive nitrogen for biosynthesis from the amido group of glutamine, the amino group of glutamate, or directly from incorporation of ammonium ion [Reitzer and Magasanik, 1987]. Based on the product of degradation, nitrogen sources can be divided to two classes: the ammonia-generating (such as D- and L-serine) and the glutamate-generating nitrogen sources (such as L-proline and L-aspartate).

When cells are grown in medium with a low concentration of ammonium ion or with a poor nitrogen source, glutamate synthase and glutamine synthetase are essential for ammonia assimilation and important for the regulation of nitrogen metabolism [Pahel *et al.*, 1978]. It has been reported that a strain carrying a mutation in glutamate synthase cannot use any of the ammonia-generating nitrogen sources except D-serine [Pahel *et al.*, 1978].

On the other hand, when cells are grown with a glutamate-generating nitrogen source, glutamate synthase is required only for the establishment of steady-state nitrogen limited growth; once the nitrogen source is degraded, the glutamate synthase is not required for the synthesis of glutamate [Reitzer and Magasanik, 1987]. However, a high level of

glutamine synthetase is required for ammonia assimilation whether with an ammonia-generating or a glutamate-generating nitrogen source [Reitzer and Magasanik, 1987]. Mutation in other genes also can result in a pleiotropic inability to use a variety of poor nitrogen sources. Such as the deficiency of the glutamine-dependent asparagine synthetase in *K. aerogenes* or certain amidotransferase in *S. typhimurium* results in the inability to grow with ,a wide variety of a single nitrogen sources [reviewed in Reitzer and Magasanik, 1987].

As shown in this study, a strain carrying a mutation in *lrp* is unable to use glycine, L-alanine, L-proline, or L-arginine as nitrogen source. This suggests that Lrp may regulate the expression of one or more genes whose product is involved in ammonia assimilation. The inability of this mutant to grow with glycine as nitrogen source may result from its low level of glycine cleavage enzymes. Inability to use other nitrogen sources might be due to the low level of the glutamate synthase, since it was reported that in *lrp* mutants, glutamate synthase activity was decreased 1000-fold [Ernsting *et al.*, 1992], This is consistent with the fact that a mutation in *gltB*, the gene which encodes the large subunit of glutamate synthase, resulted in inability to use glycine, L-proline, and L-arginine as nitrogen source.

On the other hand, the *lrp* mutant is able to use L-serine, one of the ammonia-generating nitrogen sources, as sole nitrogen source. That might be due to the fact that L-serine is degraded so rapidly in the *lrp* mutant that enough

ammonia is produced for the synthesis of glutamate by glutamate dehydrogenase, as was proposed for another ammonia-generating nitrogen source, D-serine [Pahel *et al.*, 1978].

Several of the Lrp-regulated insertion mutants were also shown to be unable to use a group of amino acids as nitrogen source. These include glycine, L-serine, L-proline, and L-arginine. The genes affected have not yet been identified. The fact that these insertion strains are found in two classes suggests that more than one gene of this type is affected. Since the effect of Lrp on the expression of these insertions is no more than 40-fold, it seems that none of the inserts were in *gltB*. Thus, Lrp must regulate two other genes which are involved in ammonia assimilation.

IV-5. Lrp Regulates the Biosynthesis and Transport of Branched-Chain Amino Acids

The synthesis of the L-leucine biosynthetic enzymes is known to be decreased in the presence of L-leucine. If this regulation is mediated by the *lrp* gene product, then a mutation in *lrp* may result in such low levels of these enzymes that L-leucine is limiting for growth and cells are partially starved. This could explain the observation that although the *lrp* mutant behaves as though it were constantly in the presence of excess L-leucine, the addition of L-leucine to the growth medium actually increases the growth rate. The partial characterization of insertion mutations showed that Lrp positively controls the expression of as yet unidentified genes involved in L-leucine biosynthesis. These mutations were

mapped at 2 min on the *E. coli* chromosome. The map location and the fact that the mutants require only L-leucine for growth suggests that these strains carried inserts in the L-leucine biosynthetic operon. The expression of these L-leucine biosynthetic genes is also affected by L-leucine independently of Lrp. This is probably due to the transcriptional attenuation mechanism found in the *leu* operon [Wessler and Calvo, 1981].

Acetohydroxy acid synthetase III (coded by *IlvIH* operon) is one of two isoenzymes synthesized by *E. coli* which catalyses the first step unique to the biosynthesis of L-isoleucine, L-valine, and L-leucine. The transcription of this operon is repressed 5- to 10-fold when cells are grown in the presence of L-leucine, but not L-isoleucine or L-valine [Squires *et al.*, 1981]. Lrp is involved in L-leucine-mediated regulation of *IlvIH* operon expression [Ricca *et al.*, 1989]. The *lrp* mutation decreased expression from this operon more than 30-fold. It was suggested that Lrp activated the transcription from *IlvIH* operon and that L-leucine repressed expression by interfering with the action of Lrp protein [Platko *et al.*, 1990]. Despite the low expression from the *IlvIH* operon, the *IlvIH* mutant is not noticeably starved for any branched-chain amino acid. This suggests that the isoenzyme acetohydroxy acid synthetase I (the gene product of *IlvBN* operon) is sufficiently expressed in the *lrp* mutant [Ricca *et al.*, 1989].

The *IlvJ* and *IlvKHMGF* operons, which are located near 76 min on the *E. coli* chromosome and whose products are involved

in high affinity transport of the branched-chain amino acids, are similar to *livJH* in their regulation. The expression of each of these operons is repressed in the presence of exogenous L-leucine [Quay and Oxender, 1976]. It was shown that the effect of L-leucine on the expression of *livJ* and *livKHMGF* also is mediated by Lrp, but the mechanism is different from that used for the *livJH* genes. Lrp represses the expression from the *livJ* and *livKHMGF* promoters and L-leucine acts as a corepressor for the repression [J. M. Calvo, personal communication].

Two-dimensional gel analysis indicated that other genes are regulated by Lrp [Ernsting *et al.*, 1992]. These include genes coding for glutamine synthetase (*glnA*), and for the outer membrane porins (OmpC and OmpF). Considering that so many genes are regulated by Lrp, it is surprising that mutations in *lrp*, even insertions, result in such a limited phenotype. As mentioned above, *lrp* mutants are physiologically gcv^- and glt^-; they are able to use L-serine as sole carbon source and require L-serine for growth at elevated temperature, and are partially starved for L-leucine. No extreme physiological deficit is seen in glucose-minimal medium grown cells.

V. Some considerations as to the probable physiological role of Lrp

V-1. The regulation of *lrp* gene expression

Since Lrp is involved in regulating the expression of such a large and diverse array of other genes, it is interesting to

know how the *lrp* gene itself is regulated. The studies on the expression of *lrp-lacZ* fusions indicated that expression from the *lrp* promoter is regulated by Lrp itself, and by growth in rich medium.

The ß-galactosidase activity from a chromosomal *lrp::lacZ* insertion was repressed 10-fold when the strain carried a lrp^+ gene in plasmid pACYC184. However, this did not involve L-leucine since the presence of L-leucine had no significant effect on expression of *lrp::lacZ*.

The expression of Lrp was decreased more than 12-fold when cells were grown in LB medium, again as shown by ß-galacotosidase activity from the chromosomal *lrp::lacZ* insertion. This suggests that autoregulation by Lrp must be overridden by other controls. Since the addition of 0.5% of casamino acids to the glucose minimal medium can repress Lrp expression over 6-fold, while the presence of any single amino acids had essentially no effect (data not shown), it seems that the regulation in LB medium is mediated by a combination of several amino acids, or by some general measure of nitrogen availability.

V-2. The Lrp may be used by *E. coli* for adapting between the free living state and the easy life inside the host

The results of this study, combined with those of other laboratories, indicated that genes regulated by Lrp encode proteins involved in amino acid biosynthesis, amino acid and oligopeptide transport, amino acid degradation, ammonia

assimilation, amino acyl-tRNA synthesis, C1 metabolism, and in methylation reactions.

Lrp activates the expression of genes whose products are involved in the biosynthesis of amino acids (*IlvIH*, *leu*, *serA*, *glnA*, *gltD*, and *gltB*), in ammonia assimilation when cells are grown in nitrogen-limiting conditions (*gltD*, and *glnA*), in C1 metabolism (*gcv*), and in the adaptation of the cells to grow in conditions of low osmolarity and low temperature (*ompF*).

On the other hand, Lrp represses expression of genes whose products are involved in transport of small molecules into the cell (*IlvJ*, *IlvKHMGF*, and *oppABCDF*), in the degradation of amino acids (*sdaA*, *tdh*, and *kbl*), and in the adaptation of cells to grow in high osmolarity and high temperature conditions (*ompC* and *lysU*).

One could summarize this, with the idea that Lrp increases biosynthetic functions and decreases degradative ones, a tendency which is partially alleviated by exogenous L-leucine. This would be even more evident in LB medium, where there would be a considerable concentration of L-leucine, which, combined with the great decrease in Lrp concentration, would have the result that the degradative functions would be even more favoured in LB than in minimal medium with L-leucine. The fact that the expression of *lrp* is decreased more than 12-fold in LB medium suggests that the role of Lrp may be less important in conditions in which the cell is surrounded by food.

These observations suggests that the leucine/Lrp regulon serves to decrease the synthesis of proteins required

primarily in rich medium and to increase the synthesis of biosynthetic enzymes required in poor medium. Since the intestinal tract of the host is usually richer than most external environments, it seems that Lrp plays a role in controlling a switch of *E. coli* cells between the rich, intestinal-type environments and the poor environments of the outside world. Lrp would repress the genes involved in the degradation of exogenous compounds and the adaptation of the cells to grow in intestinal tract of the host, while activating the genes involved in the biosynthesis and the adaptation of cells to grow in free living environments.

A high concentration of L-leucine probably occurs in nature only as a result of massive proteolysis. This could make L-leucine, which is a metabolic dead-end in *E. coli*, a convenient indicator of a rich environment. As studied so far, the member genes of the leucine/Lrp regulon which are induced by the presence of L-leucine tend to favour catabolism, whereas those repressed by exogenous L-leucine seem to favour biosynthesis.

VI. The *lrp* gene product

Lrp is the DNA binding protein which governs expression of the leucine/Lrp regulon [Platko *et al.*, 1990]. It is a moderately abundant, basic protein, existing in cells as a homodimer of molecular weight 38 kDa [Willins *et al.*, 1991]. DNA binding activity of Lrp was first shown at the *IlvIH* promoter, under the name of Ihb (*IlvIH* binding protein) [Ricca *et al.*, 1989].

Both *in vivo* and *in vitro* studies demonstrate that Lrp acts directly to activate the expression of *lrp*-regulated operons [Q. Wang and J.M. Calvo, personal communication]. *In vitro*, it was shown that Lrp binds to two sites upstream of the *IlvIH* promoter and that mutations in each site prevent binding. The mutations that decrease binding activity decrease the *in vivo* transcription from *IlvIH* operon. In addition, *in vitro* studies show that purified Lrp stimulates transcription from the *IlvIH* operon.

Apart from the fact that it contains no tryptophan residues, Lrp does not seem to have particular structural features differentiating it from other DNA-binding proteins. Lrp has a region centred at amino acid 40 that shows modest similarity to a known helix-turn-helix DNA binding motifs [Willins *et al.*, 1991].

Comparison of the complete amino acid sequences of Lrp and AsnC showed 25% of amino acid residues are identical between the two proteins [Willins *et al.*, 1991]. I showed in this work that the similarity between Lrp and AsnC is enough for AsnC to substitute for Lrp to stimulate the expression of *gcv* operon and regulate some other members of leucine/Lrp regulon.

Whether Lrp can replace AsnC in the activation of transcription from the *asnA* promoter or repression of the transcription of *asnC* is unclear. The effects of AsnC on the expression of different members of leucine/Lrp regulon are different (Table 26). The expression of some members of the leucine/Lrp regulon is not affected by AsnC at all (data not shown). Whether the cross-specificities of these regulators

has physiological significance is not known. In any case, the fact that the amino acid sequence homology is uniformly distributed throughout the two proteins strongly argues that the structural genes for AsnC and Lrp have evolved from a common ancestral gene.

Other examples of cross-regulation are known. It has been shown that the protein kinases CheA and NR_{II} have cross-specificities [Ninfa *et al.*, 1988). As mentioned in the introduction, NR_{II} is a histidine kinase of the Ntr regulon. CheA is a histidine kinase in the bacterial chemotaxis system that acts to phosphorylate two response regulators: CheY, which directly interacts with the flagellar motor to control swimming behaviour [Wylie *et al.*, 1988], and CheB, a methylesterase that controls receptor methylation and thus sensitivity of the chemotactic sensory system [Ninfa *et al.*, 1988]. CheA can catalyze the ATP-dependent phosphorylation of NR_I and thereby activate transcription from $glnAp_2$ promoter, and NR_{II} can phosphorylate CheY [Ninfa *et al.*,1988]. It was suggested that, to some degree, any histidine kinase can act as a phosphodonor for any response regulator [Stock *et al.*, 1989].

There are several cases in which functional similarities like those of AsnC and Lrp are thought to be due to the fact that they are coded by homologous genes which have a common origin, either through duplication of an ancestral gene or translocation from a related prokaryote [Mizuno *et al.*, 1983; Ferrara *et al.*, 1984; Belfaiza *et al.*, 1986; Leveque *et al.*, 1990]. The proteins that are highly homologous usually show

functional similarities [Craigen *et al.*, 1985; Bollinger *et al.*, 1984; Mizuno *et al.*, 1983; Belfaiza *et al.*, 1986]. In some cases, the proteins encoded by the homologous genes have exactly the same function, such as lysyl-tRNA synthetases (encoded by the *lysU* and *lysS* genes) [Clark and Neidhardt, 1990; Leveque *et al.*, 1990], elongation factors EF-Tu (encoded by the *tufA* and *tufB* genes) [An and Friesen, 1980; Yokota *et al.*, 1980], aspartokinase-homoserine dehydrogenases (encoded by the *thrA* and *metL* genes) [Ferrara *et al.*, 1984], ornithine carbamoyltransferases (encoded by the *argF* and *argI* genes) [Von Vliet *et al.*, 1984], and SAM synthetases (encoded by the *metK* and *metX* genes) [Satishchandran *et al.*, 1990].

In some cases, DNA-binding proteins resemble one another even more extensively than AsnC and Lrp, and have been organized into different conceptual families [reviewed in Henikoff *et al.*, 1988]. For example, in two-component regulatory systems, gene expression is controlled by 2 proteins, member of one of two homologous families of proteins: one belonging to the histidine protein kinase family and the other to the response regulator family.

VII. Techniques for the Study of Protein-DNA Interactions

A number of methodologies have been used to study the interactions of regulatory proteins with the specific DNA sequences which they recognize and bind. A technique which was extensively used in early work is the filter binding assay [Riggs *et al.*, 1970], in which the DNA-protein complex is retained on a nitrocellulose filter while free DNA passed

through the filter. One of the limitations of this method is that it does not allow analysis of the detailed composition of the products of the binding reaction. Since the binding of a single protein to a DNA molecule is usually sufficient cause for filter retention, it is impossible to determine directly the extent to which a single DNA species binds several proteins.

A rapid and simple technique, gel retardation, which involves separation of free DNA from DNA-protein complexes based on differences in their electrophoretic mobilities [Garner and Revzin, 1981; Fried and Crothers, 1981], has been widely used in the study of protein-DNA interactions.

The gel retardation assay has several advantages compared to the filter binding assay and other common techniques used to study protein-nucleic acid interactions. The gel retardation method can be used in study of the simultaneous interactions of more than one proteins with DNA [Garner and Revzin, 1981] and to study the interaction of a given binding protein with a mixed population of DNA fragments [Fried and Crothers, 1981]. This method is probably more sensitive in attempts to show the existence of kinetically labile protein-DNA complexes than is the filter binding assay, because of the tendency for such complexes to dissociate during washing of the filter. One example of this it that the presence of IPTG, LacR-operator complexes can be detected by the gel retardation method but not by the filter binding assay [Fried and Crothers, 1981].

Further, the gel retardation method has two important and

unique advantages. Since the protein-DNA complexes can be separated from the free proteins, this method can be used for purifying new proteins that binds to a defined target DNA. The other unique advantage is that the protein-DNA complexes can be resolved even if they differ only in their stoichiometries or in the physical arrangement of their components [Carey, 1990].

Whatever their advantages, neither the filter binding assay nor the gel retardation methods define the protein binding sites on DNA. This can be done by DNA footprinting. The current popularity of technique dates from the original reports of DNase I footprinting [Galas and Schmitz, 1978]. There are two kinds of footprinting experiments, the protection methods and interference methods [Tullius, 1989].

The protection experiment is based on the fact that bound protein blocks some cleavage or modification reaction performed on the DNA-protein complex. The reagent used for protection footprinting studies can be a nuclease or a small molecule that reacts with unprotected DNA, for instance dimethyl sulfate, hydroxyl radical, MPE, and copper phenanthroline. DNase I is the most common reagent for protection studies because of its relatively low sequence specificity in cleavage of DNA [Drew and Travers, 1984] and the footprint is usually easy to obtain and simple to interpret.

In the interference experiments, the DNA is specifically modified at different sites. A study as to which of the

specifically chemically modified nucleotides interfered with the binding of protein to DNA then indicates which nucleotides were important for protein binding.

VIII. The Interaction between Lrp and the Upstream Region of Genes Regulated by Lrp

Lrp, under the name of Ihb, was known to bind to two regions upstream of the *IlvIH* promoter [Ricca *et et.*, 1989]. It is possible then that Lrp binds to the promoter regions or the regions upstream of the promoters of all genes it regulates. Using the gel retardation technique, I show that Lrp binds to promoter/regulatory regions of *serA*, *sdaA*, *lysU*, *gcv* and *lrp* in vitro.

VIII-1. Lrp Binding to the *lysU* Upstream Region

The expression of the *lysU* gene is regulated by the *lrp* gene product. This was indicated by the fact that lysyl-tRNA synthetase activity increased 4-fold in an *lrp*::Tn*10* mutant, and that this corresponded to an actual increase in the *lysU* gene product as identified by 2-D gel analysis.

Gel retardation and DNase I footprinting experiments demonstrate that purified Lrp protein binds to the DNA sequence immediately upstream of the *lysU* gene, and may exert its effect on *lysU* directly. The pattern of expression of *lysU* seen by 2-D gel analysis and measured by enzyme activity is consistent with Lrp acting as a repressor of *lysU* expression, and L-leucine counteracting this repression. In agreement with

the pattern of expression observed in *In vivo* experiments, the gel retardation experiments showed that the binding of Lrp protein to the *lysU* upstream region was inhibited by the addition of L-leucine. These observations are consistent with a direct and specific effect of Lrp protein on *lysU* expression, although further experiments will be required to conclusively prove that the regulation of *lysU* expression by Lrp is mediated by the binding of Lrp protein to this region of *E. coli* DNA.

The observation that L-leucine reduces the binding of Lrp to the *lysU* upstream region is also consistent with the observations on the effect of L-leucine on the regulation of *IlvIH* by Lrp [Ricca *et al.*, 1989]. Although *IlvIH* expression is positively regulated by Lrp, and *lysU* expression is negatively regulated, in both cases the presence of high concentrations of L-leucine reduces the binding of Lrp to the upstream region of the genes. These observations are consistent with an allosteric regulation of Lrp protein by L-leucine, with the active DNA-binding form of Lrp being unliganded.

Not only L-leucine, but also L-isoleucine and L-alanine were observed to reduce the binding of Lrp protein to *lysU* upstream region in *In vitro* experiments. *In vivo*, the presence of L-leucine or L-alanine, but not L-isoleucine, resulted in the increase of *lysU* expression as measured by enzyme activity. The addition of L-alanine to the growth medium had already been shown to increase lysyl-tRNA synthetase activity

[Hirshfield *et al.*, 1981], and to elevate the synthesis of *lysU* gene product as judged by 2-D gel analysis [R.G. Matthews, personal communication]. A role for L-alanine in the regulation of protein synthesis in the leucine/Lrp regulon is also suggested by earlier studies of regulation of genes now known to be regulated by Lrp. It has been reported that the addition of L-alanine to growth medium led to the increased transcription of the *oppABCDF* operon [Andrews and Short, 1986]. Transcription of the oligopeptide permease operon is known to be negatively regulated by the *oppI* gene [Andrews and Short, 1986; Andrews *et al.*, 1986]. The fact that *oppI* was recently shown to be identical to *lrp* [Willins *et al.*, 1991] and the fact that the expression of several Lrp regulated genes was affected by the presence of L-alanine suggests that the effect of L-alanine on the binding of Lrp protein to the *lysU* upstream region is specific and relevant to the regulation of *lysU* expression *in vivo*.

As shown in Fig. 6, Lrp could bind to two sequences upstream of *lysU*. The concentration dependence for binding of Lrp to two separated sequences was much higher than that for binding to the fragment containing both sequences. This might suggest that the binding of Lrp to two sequences is cooperative, or that each of the sequences contained only part of the Lrp binding site.

It was recently shown that two transcriptional initiation sites are seen upstream of *lysU*, these being located 80 and 88 bp upstream from the start codon AUG [Leveque *et al.*, 1991].

The result of DNase I footprinting experiments identify one region of binding of Lrp between nucleotides 160 and 240. This binding site overlaps both -10 and -35 boxes. It is possible that binding of Lrp to this site can potentially occlude both -10 and -35 elements and block the binding of RNA polymerase.

VIII-2. Binding of Lrp to the *serA* Promoter Region

A. Dual promoter control the expression of *serA*.

Two extension products were shown on the electrophoretograms of the primer extension reaction performed with the RNAs isolated from the *lrp* mutant. In contrast, only one initiation site of *serA* transcription ($serAP_1$) was observed in the wild-type strain. It is surprising that transcription from $serAP_2$ was repressed by Lrp, since the transcription of *serA* was known to be positively controlled by Lrp. However, in several cases, coordinate regulation of transcription from different promoters was observed in genes controlled by dual or multiple promoters [Musso *et al.*, 1977; Malan and McClure, 1984; Huang *et al.*, 1990; 1992], such that increased transcription from one promoter can be brought about by blocking another.

The occurrence of two or several promoters in the regulatory regions of genes or operons in *E. coli* is not a new observation. In some of CRP-cAMP regulated genes, clusters of overlapping promoters are controlled through a single CRP-cAMP target site. In the *gal* operon, two overlapping promoter are separated by 5 bp [Musso *et al.*, 1977; Alba *et al.*, 1981]. It was suggested that the presence of CRP-cAMP resulted in

transcription initiation being shifted from the upstream promoter (*galP*$_2$) to the downstream promoter (*galP*$_1$) and that the regulation of transcription through promoter selection might be a common feature of CRP-cAMP regulated promoters [Malan and McClure, 1984]. In the *lac* operon, two weak promoters, *lacP*$_2$ and *lacP*$_3$, are located 22 and 15 bp upstream of the strong *lacP*$_1$ promoter [Malan and McClure, 1984; Xiong *et al.*, 1990]. All three promoters are controlled through one common CRP-cAMP target site: *lacP*$_2$ and *lacP*$_3$ are repressed by CRP-cAMP, while *lacP*$_1$ is stimulated by CRP-cAMP [Xiong *et al.*, 1990]. It was recently reported that the positive effect of CRP-cAMP on *ompB* expression *in vivo* is also a composite of activation and inhibition by CRP-cAMP of multiple *ompB* promoters [Huang *et al.*, 1992].

B. Binding of Lrp to the *serA* promoter region may alter transcription.

The transcription of *serA* in *E. coli* is under positive control by Lrp, but the mechanism of this effect is unknown. The gel retardation assay indicated that Lrp binds to at least two sites of *serA* upstream region. This observation suggest that Lrp is a direct effector of *serA* *in vivo* and *in vitro*. The overall positive effect of Lrp on *serA* transcription *in vivo* seems to be a composite of activation and repression by Lrp on multiple *serA* promoters. Primer extension analysis shows that one *serA* transcript is activated by Lrp, concomitant with the inhibition of synthesis of a second

transcript. The result of DNase I footprinting identified one region of high affinity binding between -82 and -157. This region is overlaps the proposed -10 and -35 boxes of *serAP*$_2$. Lrp protein negatively controls the promoter that overlaps this Lrp binding site and positively regulates the promoter that is located further downstream from this site. Binding of Lrp to this region then might well hinder the binding of RNA polymerase to *serAP*$_2$, and inhibit transcription from this promoter.

Lrp is both a positive and a negative regulator of transcription of numerous operons in *E. coli*. The dual control of Lrp protein on *serA* expression is unusual, in that this is the first indication that Lrp both activates and represses transcription from an individual operon.

In this mechanism, Lrp is similar to CRP-cAMP. In the extensively studied *gal* operon, a single Lrp binding site in the *gal* promoter region positively regulates the *galP*$_1$ promoter and negatively regulates a second promoter, *galP*$_2$ [Adhya, 1987]. It was recently shown that binding of CRP-cAMP to the *ompB* promoter region activated transcription from two promoters that overlap the CRP binding site, and inhibited transcription from the other two that are located further downstream from this site [Huang *et al.*, 1992].

VIII-3. Lrp Binds to the *sdaA* Upstream Region

Similar to the case of *lysU*, the expression of *sdaA* is under the negative control of Lrp. The gel retardation assay

demonstrated that there is a high affinity Lrp binding site and at least one low affinity binding site located in the upstream region of *sdaA*.

The transcription start site of *sdaA* is still unknown, despite considerable effort. However, ß-galactosidase activity from plasmid pSUZ4 (promoterless *lacZ* fused to the *sdaA* high affinity binding site) is similar to that from plasmid pSUZ2 (containing both binding sites) in *lrp* background. This observation suggests that the promoter of *sdaA* is located within or upstream of high affinity binding site.

On the other hand, the expression of ß-galactosidase from plasmid pSUZ4 is also repressed by Lrp protein, but only 2- to 3-fold, a much lower effect than in seen in the intact gene. With plasmid pSUZ2, which carries both sites, the presence of Lrp protein decreases the expression of ß-galactosidase more than 10 fold- similar to regulation of the chromosomal gene. These observations suggest that both two binding sites are required for full repression of *sdaA* expression by Lrp.

IX. Possible Mechanisms by which Lrp Regulates Gene Expression

IX-1. The Mechanisms by Which Lrp Interacts with Its Target Promoters

Several *E. coli* regulatory proteins, such as CRP and AraC, can activate transcription of some operons and inhibit transcription of others. CRP-cAMP can both positively and negatively regulate transcription through different promoters

with a single operon. The ligand-free CRP has no function in gene regulation. Both activation and repression require the presence of the ligand cAMP.

Lrp, like CRP, can positively regulate the expression of some operons and negatively regulate the expression of others. The interaction between Lrp and its effector, L-leucine, is very complex. At least six patterns of interaction between Lrp protein, L-leucine, and Lrp-regulated promoters have been demonstrated (Table 24).

As shown in Table 24, Lrp can stimulate or repress the expression of its target genes both with and without involvement of L-leucine. In both cases, there are three types of interaction known: Lrp acts alone, Lrp action requires L-leucine, or Lrp action is reversed by L-leucine.

As shown in this study, Lrp protein greatly stimulates the expression of the *gcv* operon, but the presence of L-leucine has little or no effect. Consistent with this is the fact that the gel retardation assay shows that the purified Lrp protein binds to the upstream region of the *gcv* operon; the presence of L-leucine does not affect the binding activity. On the other hand, the expression of *lrp* is subject to negative autoregulation, and is very little effected by L-leucine. Again the *in vitro* binding study shows that L-leucine has no effect on the binding of Lrp to the *lrp* upstream region.

In the one of the most extensively studied Lrp regulated-operon, transcription from *IlvIH* promoter is activated by Lrp and repressed by L-leucine [Platko *et al.*, 1990]. Both *in vivo* and *in vitro* data suggested that L-leucine represses *IlvIH*

TABLE 24. The interaction between Lrp, L-leucine, and the target promoters.

Action	Prototype gene
Lrp activates	
L-leucine reduces the effect of Lrp	*IlvIH*
L-leucine is required for Lrp effect	*sdaB*[a]
L-leucine has no effect	*gcv*
Lrp represses	
L-leucine reduces the effect of Lrp	*sdaA*
L-leucine is required for Lrp effect	*IlvJ/K*
L-leucine has no effect	*lrp*

[a]: Or gene regulating *sdaB*.

expression by interfering with the action of Lrp protein [Platko *et al.*, 1990]. This is also true for *serA*. The expression from *sdaA* and *lysU* promoter, on the other hand, is repressed by Lrp and this repression is reduced by L-leucine. The results of *In vitro* binding assays are consistent with the *In vivo* observations, that is, the presence of L-leucine can reduce the binding of Lrp protein to upstream region of *sdaA* and *lysU*. However, it has not been shown clearly that binding of Lrp to these regions is the factor that in fact controls their expression of *sdaA* and *lysU*.

In other cases, Lrp can acts as either repressor (*IlvJ/K*) or activator (CP41, CP52 and CP57) only in combination with L-leucine. For this type of regulation, the *In vitro* binding study which would demonstrate whether Lrp does bind to these promoters and whether binding of Lrp requires the presence of L-leucine has not yet been done.

In the case of the *araBAD* operon, both ligand-free and ligand-bound AraC protein can bind to the *araBAD* regulatory region. However, in the absence of arabinose, ligand-free AraC protein binds to $ara I_1$ only and represses transcription from the *araBAD* promoter; in the presence of inducer arabinose, the ligand-bound AraC protein binds to both $ara I_1$ and $ara I_2$ sites and activates the transcription of *araBAD* [Lee *et al.*, 1987].

There is no real evidence that such a switch mechanism occurs in Lrp/leucine regulon. However, in insertion mutant CP25, in the absence of L-leucine, Lrp represses expression, whereas in the presence of L-leucine, Lrp stimulates

expression. This could suggest that both the repressor and the activator forms of Lrp protein are involved in the regulation of an individual gene or promoter.

What can be said in general of the role of L-leucine in gene regulation? It is possible that L-leucine binds to Lrp protein and the interaction between L-leucine and Lrp changes the protein conformation. Both ligand-free and ligand-bound Lrp proteins have binding activity, but in the presence of the ligand L-leucine, Lrp protein recognizes and binds to different sequences. Alternately, L-leucine may not affect lrp binding activity, but may affect its regulatory property. The gel retardation experiments show that at least in some cases, L-leucine did affect Lrp binding activity (as discussed above).

In the case of *lrp* and *gcv* promoters, the presence of L-leucine has essentially no effect on the gene expression. The gel retardation experiments also demonstrated that the addition of L-leucine has no effect on the Lrp binding activity. If there are two conformations of Lrp, with and without L-leucine, these observations suggest that either both forms of Lrp protein bind to a single binding site or they bind to different sites but have similar effects on transcription. A footprinting study might answer this question. In other types of promoters (except CP25), it seems that only one forms of Lrp protein can activate or repress transcription.

IX-2. Interaction of Lrp with Its Target Promoters

In vitro experiments have shown that Lrp can bind to the regulatory region of at least some of its target genes. The surprise is that DNase I footprinting experiments indicate that the Lrp protection site is unusually large: 70 bp in *IlvIH* [Ricca *et al.*, 1989], 88 bp in *lysU*, 75 bp in *sdaA*, and 70 bp in *serA*, compared with the 21 bp protected by LacR [Galas and Schmitz, 1978] and the 20-24 bp protected by PurR at its various sites [He *et al.*, 1990]. Since Lrp is a small homodimer protein, with a monomer molecular mass of 18.8kDa [Willins *et al.*, 1991], such a long binding site must bind more than one molecule.

Moreover, the DNase I footprinting experiments demonstrated that the subsites of the DNA protected by Lrp show similar affinity throughout the whole region. The subregions separating the protected areas showed a regular pattern of alternately increased and decreased sensitivity to DNase I cleavage, as in the cases of MalT activator binding to *malE* and *malK* regulatory region [Raibaud *et al.*, 1989] and *pulAp* and *pulCp* regulatory region [Vidal-Ingigliardi *et al.*, 1991]. It is possible that Lrp, like MalT, forms a nucleoprotein complex at the regulatory region of its target promoters.

Based on the sequence date from the upstream of target genes, it has been proposed that TTTATTCtNaAT is a consensus sequence of a Lrp binding site [Rex *et al.*, 1991]. This sequence was found within the Lrp binding sites of *lysU*.

However, it is not found within the Lrp binding site of *serA* and *sdaA*, and even can not be found in the upstream region of the *serA*, *sdaA*, and *IlvJIK* operon. Furthermore, this sequence is rather common in both prokaryotic and eucaryotic genes, and can also be found within the coding region or downstream region of genes. No other consensus sequence has been identified so far.

It is also possible that Lrp recognizes DNA secondary structure rather than primary sequence, as proposed for OxyR [Tartaglia *et al.*, 1989]. In the case of OxyR, five large DNA binding sites (about 45 bp) do not show any similarity. [Tartaglia *et al.*, 1989].

IX-3. The Possible Mechanism of Lrp action

Lrp, like CRP-cAMP, activates expression from some promoters and inhibits expression from others. DNase I protection experiments demonstrated that at least in some Lrp repressed promoters, such as $serAP_2$ and *lysU*, the Lrp binding site overlaps with the RNA polymerase binding site. These observations suggest that the role of Lrp in negative control may due to Lrp competing with RNA polymerase for the DNA binding site. Binding of Lrp to the operator site then would block the interaction between RNA polymerase and the promoter.

In the Lrp-regulated promoters studied so far, most of them containing multiple Lrp binding sites in the upstream region. In the case of *sdaA*, it seems that both binding sites are required for full repression. Deletion of the downstream

low affinity binding site largely reduced repression by Lrp. Whether a repression loop is formed between these two binding sites would be worth testing.

How then does Lrp activate the expression of its target promoters? It is possible that Lrp's mechanism of action will be similar to that of CRP, since there are already some similarities between these two global regulatory proteins. Both of them are small, basic, relatively abundant proteins that function as homodimers [Anderson *et al.*, 1971; Willins *et al.*, 1991]. That is, Lrp may activate transcription initiation through direct protein-protein contact with the RNA polymerase or through the bending of DNA, as proposed for the mechanism of CRP-cAMP activation of *lac* expression [Reznikoff, 1992]. The DNase I footprinting experiments showed that repeating DNase I hypersensitive sites are present within all Lrp binding sites studied so far. This suggests that binding of Lrp might loop or bend DNA.

Considering Lrp binding to an unusually long stretch of DNA and the large number of Lrp binding site in the *E. coli* chromosome, at least 4 Lrp binding site in pBR322, the binding of Lrp may effect DNA structure and may be one determinant of the specificity for folding DNA in the cell.

X. Summary

This thesis described the isolation and characterization of a highly pleiotropic *Escherichia coli* mutant, *lrp*. The *lrp* gene was located near 20 min of *E. coli* chromosome and shown

to affect the activity of a number of enzymes involved in different metabolic pathways, most of which are regulated by L-leucine.

Selected for its ability to grow with L-serine as sole carbon source, the *lrp*::Tn*10* mutant showed high levels of L-serine deaminase activity (due to increased transcription of *sdaA*, one of the structural genes for L-serine deaminase), and of another amino acid-degrading enzyme, L-threonine dehydrogenase, and decreased the transcription of *serA* and *IlvIH* operons, coding for amino acid biosynthetic enzymes. The effect of the *lrp* mutation on the expression of these genes could also be observed in lrp^+ strains grown in the presence of L-leucine, whereas exogenous L-leucine had essentially no effect on the *lrp* mutant strains. The *lrp* mutation or the addition of exogenous L-leucine also suppressed the slow growth of a *metK* mutant, deficient in S-adenosylmethionine synthetase. We propose that the *lrp* gene product is the regulator of a global response to L-leucine.

To determine whether other genes whose expression is also regulated by Lrp, 27 strains with λp*lac*Mu9 insertion in L-leucine-regulated genes were isolated and characterized. The expression of *lacZ* from all of these insertion mutants was shown to be regulated by Lrp. On the other hand, L-leucine had no effect on gene expression in *lrp* mutants of all these insertions except for the insertions within the *leu* operon.

Lrp and L-leucine influenced gene expression in a surprising variety of ways. In most cases, gene expression was

either positively or negatively regulated by Lrp, and L-leucine could alleviate the effect of Lrp. In other cases, Lrp either activated or repressed only in combination with L-leucine (*sdaB* and *IlvJIK*).

However, not all genes regulated by Lrp were affected by L-leucine. The expression of *gcv* and *lrp* are under the control of Lrp but not affected by L-leucine.

The expression of *lrp* is repressed during growth in rich medium. Together with the facts that Lrp tends to activate the expression of biosynthetic genes and repress the expression of degradative and transport genes, this suggests that Lrp increases the synthesis of proteins required for growth in poor environments and decreased the synthesis of proteins required in rich medium.

The primer extension analysis demonstrated that the overall positive effect of Lrp on *serA* expression *in vivo* seems to be a composite of activation and repression by Lrp on multiple *serA* promoters.

Gel retardation assays indicated that Lrp binds to upstream region of *lysU*, *sdaA*, *serA*, *gcv*, and *lrp*. The effect of L-leucine on binding activity is consistent with the *in vivo* results. DNA footprinting analysis showed that Lrp protein protected upstream region of *lysU*, *serA*, and *sdaA* from DNase I digestion in a large region. However, no consensus sequence for Lrp binding was found.

REFERENCES

Adhya, S. 1987. The galactose operon, p. 1503-1512. *In Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology. F.C. Neidhardt (ed.). American Society for Microbiology, Washington.

Adhya, S. 1989. Multipartite genetic control elements: communication by DNA loop. ANN. Rev. Genet. *23*:227-250.

Adhya, S., and S. Garges. 1990. Positive control. J. Biol. Chem. *265*:10797-10800.

Alba, H., S. Adhya, and B. de Crombrugghe. 1981. Evidence for two functional *gal* promoters in *Escherichia coli* cells. J. Biol. Chem. *256*:11905-11910.

Amouyal, M., L. Mortensen, H. Buc, and K. Hammer. 1989. Single and double loop formation when deoR repressor binds to its natural operator sites. Cell *58*:545-551.

An, G., and J.D. Friesen. 1980. The nucleotide sequence of *tufB* and four nearby tRNA structural genes of *Escherichia coli*. Gene *12*:33-39.

Anderson, J.J., and D.L. Oxender. 1977. *Escherichia coli* transport mutants lacking binding protein and other components of the branched-chain amino acid transport systems. J. Bacteriol. *130*:384-492.

Anderson, J.J., S.C. Quay, and D.L. Oxender. 1976. Mapping of two loci affecting the regulation of branched-chain amino acid transport in *Escherichia coli* K-12. J. Bacteriol. *126*:80-90.

Anderson, W.B., A.B. Schneider, M. Emmer, R.L. Perlman, and I. Pastan. 1971. Purification of and properties of the cyclic adenosine 3', 5'-monophosphate receptor protein which mediated cyclic adenosine 3', 5'-monophosphate-dependent gene transcription in *Escherichia coli*. J. Biol. Chem. *246*:5929-5973.

Andrews, J.C., T.C. Blevins, and S.A. Short. 1986. Regulation of peptide transport in *Escherichia coli*: induction of tne *trp*-linked operon encoding the oligopeptide permease. J. Bacteriol. *165*:428-433.

Andrews, J.C., and S.A. Short. 1985. Genetic analysis of *Escherichia coli* oligopeptide transport mutants. J. Bacteriol. *161*:484-492.

Andrews, J.C., and S.A. Short. 1986. *Opp-lac* operon fusions and transcriptional regulation of the *Escherichia coli trp*-linked oligopeptide permease. J. Bacteriol. *165*:434-442.

Aronson. D., P.D. Ravnikar, and R.L. Somerville. 1988. The sequence of the 2-amino-3-ketobutyrate coenzyme A ligase(*kbl*) gene of *E. coli*. Nucleic. Acids Res. *16*:3586.

Austin, E.A., J.C. Andrews, and S.A. Short. 1989. Abstr. Mol. Genet. Bacteria Phages, p153.

Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.A. Smith, J.G. Seidman, and K. Struhl. 1987. Current Protocols in Molecular Biology, p. 4.8.1-4.8.3. John Wiley & Sons, Inc., New York.

Bachmann, B.J. 1990. Linkage map of *Escherichia coli* K-12, edition 8. Microbiol. Rev. *54*:130-197.

Bahl, H., H. Echols, D.B. Straus, D. Court, R. Crowl, and C.P. Georgeopoulos. 1987. Induction of the heat shock response of *E. coli* through stabilization of σ^{32} by the phage λ cIII protein. Genes Dev. *1*:57-64.

Bauer, B.F., E.G. Kas, R.M. Elford, and W.M. Holmes. 1988. Sequence determinants for promoter strength in the *leuV* operon of *Escherichia coli*. Gene *63*:123-134.

Belfaiza, J., C. Parsot, A. Martel, C. Bouthier de la Tour, D. Margarita, G.N. Cohen, and I. Saint-Girons. 1986. Evolution in biosynthetic pathways: Two enzymes catalyzing consecutive steps in methionone biosyhthesis originate from a common ancestor and possess a similar regulatory region. Proc. Natl. Acad. Sci. USA *83*:867-871.

Beutin, L., and M. Achtman. 1979. Two *Escherichia coli* chromosomal cistrons, *sfrA* and *sfrB*, which are needed for expression of F factor *tra* functions. J. Bacteriol. *139*:730-737.

Birkmann,. A., and A. Bock. 1989. Characterization of a cis regulatory DNA elements is necessary for formate induction of the formate dehydrogenase gene (*fdhF*) of *Escherichia coli*. Mol. Microbiol. *3*:187-195.

Bolivar, F., R.L. Rodriguez, P.J. Greene, M.C. Betlach, H.L. Heyneker, H.W. Boyer, J.H. Cross, and S. Falkow. 1977. Construction and characterization of new cloning vehicles. II. A multi-purpose cloning system. Gene *2*:95-113.

Bossi, L., and D.M. Smith. 1984. Conformational change in the DNA associated with an unusual promoter mutation in a tRNA operon of *Salmonella*. Cell *39*:643-652.

Botsford, J.L. 1981. Cyclic nucleotides in procarryotes. Microbiol. Rev. *45*:620-642.

Botsford, J.L., and R.D. DeMoss. 1971. Catabolite repression of tryptophanase in *Escherichia coli*. J. Bacteriol. *105*:303-312.

Bourgaige, D.B., T.A. Phillips, R.A. VanBogelen, P.G. Jones, F.C. Neidhardt, and M.J. Fournier. 1990. Lossof 4.5S RNA induces the heat shock response and lambda prophage in *Escherichia coli*. J. Bacteriol. *172*:1151-1154.

Braaten, B.A., L.B. Blyn, B.S. Skinner, and D.A. Low. 1991. Evidence for a methylation-blocking factor (*mbf*)locus involved in *pap* pilus expression and phase variation in *Escherichia coli*. J. Bacteriol. *173*:1789-1800.

Bracco, L., D. Kotlarz, A. Kolb, S. Diekmann, and H. Buc. 1989. Synthetic curved sequences can act as transcriptional activators in *Escherichia coli*. EMBO J. *8*:4289-4296.

Bremer, E., T.J. Silhavy, J.M. Wiesemann, and G.M. Wienstock. 1984. λplacMu: A transposable derivative of phage λ for creating *lacZ* protein fusions in a single step. J. Bacteriol. *158*:1084-1093.

Bremer, E., T.J. Silhavy, and G.M. Weinstock. 1985. Transposable λplacMu bacteriophages for creating *lacZ* operon fusions and kanamycin resistance insertions in *Escherichia coli*. J. Bacteriol. *162*:1092-1099.

Brennan, R.G., and B.W. Matthews. 1989. The helix-turn-helix DNA binding motive. J. Biol. Chem. *264*:1903-1906.

Brenowitz, M., A. Pickar, and E. Jamison. 1991. Stability of a Lac repressor mediated "looped complex". Biochemistry *30*:5986-5998.

Brent, R., and M. Ptashne. 1981. Mechanism of action of the *lexA* gene product. Proc. Natl. Acad. Sci. USA *78*:4204-4208.

Brown, E.A., R. D'Ari, and E.B. Newman. 1990. A relationship between L-serine degradation and methionine biosynthesis in *Escherichia coli* K-12. J. Gen. Microbiol. *136*:1017-1023.

Buck, M., S. Miller, M. Drummond, and R. Dixon. 1986. Upstream activator sequences are present in the promoter on nitrogen fixation genes. Nature (London) *320*:374-378.

Burgess R.R., and A.A. Travers. 1970. *Escherichia coli* RNA polymerase: purification, subunit structure, and factor requirements. Fed. Proc. *29*:1164-1169.

Burns, R.O., J. Calvo, P. Margolin, and H.E. Umbarger. 1966. Expression of the leucine operon. J. Bacteriol. *91*:1570-1576.

Buxton, R.S., and L.S. Drury. 1983. Cloning and insertional inactivation of the *dye* (*sfrA*) gene. Mutation of which affects sex factor F expression and dye sensitivity of *Escherichia coli* K-12. J. Bacteriol. *154*:1309-1314.

Carey, J. 1991. Gel retardation. Methods in Enzymology *208*:103-117.

Carter, P.W., D.L. Weiss, H.L. Weith, and J.M. Calvo. 1985. Mutations that convert the four leucine codons of the *Salmonella typhimurium leu* leader to four theronine codons. J. Bacteriol. *162*:943-949.

Casadaban, M.J., and J. Chou. 1984. *In vivo* formation of gene fusions encoding hybrid ß-galactosidase protein in one step with a transposable Mu-*lac* transducing phage. Proc. Natl. Acad. Sci. U.S.A. *81*:535-539.

Casadaban, M.J., and S.N. Cohen. 1979. Lactose genes fused to exogenous promotors in one step using a Mu-*lac* bacteriophage: *In vivo* probe for transcriptional control sequences. Proc. Natl. Acad. Sci. U.S.A. *76*:4530-4533.

Chamberlin, M.J. 1974. The selectivity of transcription. Annu. Rev. Biochem. *43*:721-775.

Chan, T.T.K., and E.B. Newman. 1981. Threonine as a carbon source for *Escherichia coli*. J. Bacteriol. *145*:1150-1153.

Chang A.C.Y. and S.N. Cohen. 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the p15A cryptic miniplasmid. J. Bacteriol. *134*:1141-1156.

Chippaux, M., V. Bonnefoy-Orth, J. Ratouchniak, and M.C. pascal. 1981. Operon fusions in the nitrate reductase operon and study of the the control gene *nirR* in *Escherichia coli*. Mol. Gen. Genet. *182*:477-479.

Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. *162*:156-159.

Christman, M.F., R.W. Morgan, F.S. Jacbson, and B.N. Ames. 1985. Positive control of a regulon for defenses against oxidative stress and some heat-shock proteins in *Salmonella typhimurium*. Cell *41*:753-762.

Christman, M.F., G. Storz, and B.N. Ames. 1989. OxyR, a positive regulator of hydrogen peroxide-inducible genes in *Escherichia coli* and *Salmonella typhimurium*, is homologous to a family of bacterial regulatory proteins. Proc. Natl. Acad. Sci. USA *86*:3484-3488.

Clark, R.L. and F.C. Neidhardt. 1990. Roles of the two lysyl-tRNA synthetases of *Escherichia coli*: Analysis of nucleotide sequence and mutant behaviour. J. Bacteriol. *172*:3237-3243.

Collado-Vides, J., B. magasanik, and J. Gralla. 1991. Control site location and transcriptional regulation in *Escherichia coli*. Microbiol. Rev. *55*:371-394.

Cosloy, S.D. and E. McFall. 1970. L-Serine-sensitive mutants of *Escherichia coli* K-12. J. Bacteriol. *103*:840-841.

Cotter, P A. and R.P. Gunsalus. 1989. Oxygen, nitrate, and molybdenum regulation of *dmsABC* gene expression in *Escherichia coli*. J. Bacteriol. *171*:3817-3823.

Cotter, P.A., V. Chepuri, R.B. Gennis, and R.P., Gunsalus. 1990. Cytochrome o (*cyoABCDE*) and d (*cydAB*) oxidase gene expression in *Escherichia coli* is regulated by oxygen, pH, and the *fnr* gene product. J. Bacteriol. *172*:6333-6338.

Cowing, D.W., J.C.A. Bardwell, E.A.Craig, C. Woolford, R.W. Hendrix, and C.A. Gross. 1985. Consensus sequence for *Escherichia coli* heat shock promotors. Proc. Natl. Acad. Sci. U.S.A. *82*:2679-2683.

Craig, N.L., and H.A. Nash. 1984. *E. coli* integration host factor binds to specific sites in DNA. Cell *39*:707-716.

Craigen, W.J., R.G. Cook, W.P. Tate, and C.T. Caskey. 1985. Bacterial peptide chain release factors: Conserved primary structure and possible frameshift regulation of release factor 2. Proc. Natl. Acad. Sci. USA *82*:3616-3620.

Dandanell, G., and K. Hammer. 1985. Two operator sites separated by 599 base pairs are required for *deoR* repression of the *deo* operon of *Escherichia coli*. EMBO J. *4*:3333-3338.

Dandanell, G., P. Valentin-Hansen, J.E. Love-larsen, and K. Hammer. 1987. Long-range cooperativity between gene regulatory sequences in a prokaryote. Nature *325*:823-826.

Demple, B. and J. Halbrook. 1983. Inducible repair of oxidative damage in *Escherichia coli*. Nature *304*:466-468.

Dendinger, S., and W.J. Brill. 1970. Regulation of proline degradation in *Salmonella typhimurium*. J. Bacteriol. *103*:144-152.

Drahos, D.J., and R.W. Hendrix. 1982. Effect of bacteriophage lambda infection on the synthesis of *groE* protein and other *Escherichia coli* proteins. J. Bacteriol. *149*:1050-1063.

Drlica, K. and J. Rouviere-Yaniv. 1987. Histonelike proteins of bacteria. Microbiol. Rev. *51*:301-319.

Echols, H. 1986. Multiple DNA-protein interactions governing high-precision DNA transactions. Science *233*:1050-1055.

Egan, S.M., and V. Stewart. 1990. Nitrate regulation of anaerobic respiratory gene expression in *narX* deletion mutants of *Escherichia coli* K-12. J. Bacteriol. *172*:5020-5029.

Emmerich, R.V. and I.N. Hirshfield. 1987. Mapping of the constitutive lysyl-tRNA synthetase gene of *Escherichia coli* K-12. J. Bacteriol. *169*:5311-5313.

Ernsting, B.R., M.R. Atkinson, A.J. Ninfa. and R.G. Matthews. 1992. Characterization of the regulation controlled by the leucine-responsive regulatory protein in *Escherichia coli*. J. Bacteriol. *174*:1109-1118.

Ferrara, P., N. Duchange, M.M. Zakin, and G.N. Cohen. 1984. Internal homologies in the two aspartokinase-homoserine dehydrogenases of *Escherichia coli* K-12. Proc. Natl. Acad. Sci. USA *81*:3019-3023.

Flashner, Y., and Gralla. 1988. Dual mechanism of repression at a distance in the *lac* operon. Proc. Natl. Acad. Sci. USA *85*:8968-8972.

Fraser, J., and E.B. Newman. 1975. Derivation of glycine from threonine in *Escherichia coli* K-12 mutants. J. Bacteriol. *122*:810-817.

Freundlich, M., R.O. Burns, and H.E. Umbarger. 1962. Control of isoleucine, valine and leucine biosynthesis. I. Multivalent repression. Proc. Natl. Acad. Sci. USA *48*:1804-1808.

Fried, M. and D.M. Crothers. 1981. Equilibria and kinetics of *lac* repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. *9*:6505-6525.

Friedman, D.I. 1988. Integration host factor: a protein for all reasons. Cell *55*:545-554.

Fu, H.A., S. Iuchi, and E.C.C. Lin. 1991. The requirement of ArcA and Fnr for peak expression of the *cyd* operon in *Escherichia coli* under microaerobic conditions. Mol. Gen. Genet. *226*:209-213.

Galas, D.J. and A. Schmitz. 1978. Nucleic Acids Res. *5*:3157-3170.

Garner, M.M., A. Revzin. 1981. A gel electrophoresis methods for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. Nucleic Acids Res. *9*:3047-3060.

Gartenberg, M.R., and D.M. Crothers. 1991. Synthetic DNA bending sequences increase the rate of *in vitro* transcription initiation at the *Escherichia coli lac* promoter. J. Mol. Biol. *219*:217-230.

Gaston, K., A. Bell, A. Kolb, H. Buc, and S. Busby. 1990. Stringent spacing requirements for transcription activation by CRP. Cell *62*:733-743.

Gehring, W.J. 1987. Homeo boxes in the study of development. Science *236*:1245-1252.

Gerlach, P., L. Sogaard-Andersen, H. Pedersen, J. Martinussen, P. Valentin-Hansen, and E. Bremer. 1991. The cyclic AMP (cAMP)-cAMP receptor protein complex functions both as an activator and as a corepressor at the *tsx*-P_2 promoter of *Escherichia coli* K-12. J. Bacteriol. *173*:5419-5430.

Gober, J.W., and L. Shapiro. 1990. Integration host factor (IHF) is required for the activation of developmentally regulated genes in *Caulobacter*. Genes Dev. *4*:1494-1504.

Goodman, S.D., and H.A. Nash. 1989. Functional replacement of a protein-induced bend in a DNA recombination site. Nature *341*:251-254.

Gottesman, S. 1984. Bacterial regulation: global regulatory networks. Ann. Rev. Genet. *18*:415-441.

Gourse, R.L., H.A. de Boer, and M. Nomura. 1986. DNA Determinants of rRNA synthesis in *E. coli*: growth rate dependent regulation, Feedback inhibition, upstream activation, antitermination. Cell *44*:197-205

Gralla, J.D. 1989. Bacterial gene regulation from distant DNA sites. Cell *57*:193-195.

Gralla, J.d. 1990. Promoter recognition and mRNA initiation by *Escherichia coli* E-sigma-70. Methods Enzymol. *185*:37-54.

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

Greenberg, J.T., and B. Demple. 1989. A global response induced in *Escherichia coli* by redox-cycling agents overlaps with that induced by peroxide stress. J. Bacteriol. *171*:3933-3939.

Greenberg, J.T., P. Mouach, J. Chou, P.D. Josephy, and B. Demple. 1990. Positive control of a multilevel antioxidant defense regulon activated by superoxide-generating agents in *Escherichia coli*. Proc. Natl. Acad. Sci. USA *87*:6181-6185.

Greene, R.C., J.V.S. Hunter, and E.H. Coch. 1973. Properties of *metK* mutants of *Escherichia coli*. J. Bacteriol. *115*:57-67.

Grossman, A.D., J.W. Erickson, and C.A. Gross. 1984. The *htpR* gene product of *E. coli* is a sigma factor for heat shock promoters. Cell *38*:383-390.

Gross, C.A., D.B. Straus, J.W. Erickson, and T. Yura. 1990. The function and regulation of heat shock proteins in *Escherichia coli*. p. 167-189. *In* R.I. Morimoto, A. Tissieres, and C. Georgeopoulos (ed.) Stress proteins in biology and medicine. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Guan, C.D., B. Wanner, and H. Inouye. 1983. Analysis of regulation of *phoB* expression using a *phoB-cat* fusion. J. Bacteriol. *156*:710-717.

Haener, E.W., C.W. Tabor, and H. Tabor. 1977. Isolation of a *metK* mutant with a temperature-sensitive S-adenosylmethionine synthetase. J. Bacteriol. *132*:832-840.

Hama, H., T. Shimamoto, M. Tsuda, and T. Tsuchiya. 1988. Characterization of a novel L-serine transport system in *Escherichia coli*. J. Bacteriol. *170*:2236-2239.

Hama, H., Y. Sumita, Y. Katutagi, M. Tsuda, and T. Tsuchiya. 1990. Target of serine inhibition in *Escherichia coli*. Biochem. Biophys. Res. Comm. *168*:1211-1216.

Hammer-Jespersen, K. 1983. Nucleoside catabolism. p.203-258. In Metabolism of nucleotides, nucleosides and nucleobases in microorganisms. A. Munch-Petersen (ed.). Academic Press, Inc. (London), Ltd., London.

Hawley, D.K., A.D. Johnson, and W.R. McClure. 1985. Functional and physical characterization of transcriptioninitiation complexes in the bacteriophage λ O_R region. J. Biol. Chem. *260*:8618-8620.

Hawley, D.K., and W.R. McClure, 1982. Mechanism of activation of transcription initiation from the λ P_{RM} promoter. J. Mol. Biol. *157*:493-525.

Hawley, D.K., and W.R. McClure, 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. *11*:2237-2255.

He, B., A. Shiau, K.Y. Choi, H. Zaikin, and J.M. Smith. 1990. Genes of the *Escherichia coli pur* regulon are negatively controlled by a repressor-operator interaction. J. Bacteriol. *172*:4555-4562.

Henikoff, S., G.W. Haughn, J.M. Calvo, and J.C. Wallace. 1988. A large family of bacterial activator proteins. Proc. Natl. Acad. Sci. USA *85*:6602-6606.

Herendeen, S.L., R.A. VanBogelen, and F.C. Neidhardt. 1979. Levels of major proteins of *Escherichia coli* during growth at different temperatures. J. Bacteriol. *139*:185-194.

Hirschman, J., P.K. Wong, K. Sei, J.Keener, and S. Kustu. 1985. products of nitrogen regulatory genes *ntrA* and *ntrC* of enteric bacteria activate *glnA* transcription *in vitro*: evidence that the *ntrA* product is a σ factor. Proc. Natl. Acad. Sci. USA *82*:7525-7529.

Hirshfield, I.N., P.L.Block, R.A. Van Bogelen, and F.C. Neidhardt. 1981. Multiple forms of lysyl-transfer ribonucleic acid synthetase in *Escherichia coli*. J. Bacteriol. *146*:345-351.

Hirshfield, I.N., C. Liu, and F.M. Yeh. 1977. Two models of metabolic regulation of lysyl-transfer ribonucleic acid synthetase in *Escherichia coli* K-12. J. Bacteriol. *131*:589-597.

Hobson, A.C. 1974. The regulation of methionine and S-adenosylmethionine biosynthesis and utilization in mutants of *salmonella typhimurium* with defects in S-adenosylmethionine synthetase. Mol. Gen. Genet. *131*:263-273.

Hochschid, A. and M. Ptashne. 1986. Cooperative binding of λ repressors to sites separated by integral turns of DNA helix. Cell *44*:681-687.

Hoopes, B.C., and W.R. McClure. 1987. Strategies in regulation of transcription initiation. p. 1231-1240. *In Escherichia coli* and *Salmonella typhimurium* Cellular and Molecular Biology. F.C. Neidhardt (ed.). American Society for Microbiology, Washington.

Hoover, T.R., E. Santero, S. Porter, and S. Kustu. 1990. The integration host factor stimulates interaction of RNA polymerase with NfiA, the transcriptional activator for nitrogen fixation operons. Cell *63*:11-22.

Huber, P., and W. Arber. 1989. Mutational analysis of a prokaryotic recombinational enhancer element with two functions. EMBO J. *8*:577-585.

Huber, P., S. Iida, W. Arber, and T.A. Bickle. 1985. Site-specific DNA inversion is enhanced by a DNA sequence element in cis. Proc. Natl. Acad. Sci. USA *82*:3766-3780.

Huang, L., P. Tsui, and M. Freundlich. 1990. Integration host factor is a negative effector of the in vitro expression of *ompC* in *Escherichia coli*. J. Bacteriol. *172*:5293-5298.

Huang, L., P. Tsui, and M. Freundlich. 1992. Positive and negative control of *ompB* transcription in *Escherichia coli* by cyclic AMP and the cyclic AMP receptor protein. J. Bacteriol. *174*:664-670.

Hunt, T.P., and B. Magasanik. 1985. Transcription of *glnA* by purified *Escherichia coli* components: core RNA polymerase and the products of *glnF*, *glnG* and *glnL*. Proc. Natl. Acad. Sci. USA *82*: 8453-8457.

Hunter, J.S.V., R.C. Greene, and C.H. Su. 1975. Genetic characterization of the *metK* locus in *Escherichia coli* K-12. J. Bacteriol. *122*:1144-1152.

Hwang, J., and G. Gussin. 1988. Interactions between *Escherichia coli* RNA polymerase and lambda repressor: Mutations in P_{RM} affect repression of P_R. J. Mol. Biol. *200*:735-739.

Ingledew, W.J., and R.K. Poole. 1984. The respiratory chains of *Escherichia coli*. Microbiol. Rev. *48*:222-271.

Isenberg, S., and E.B. Newman. 1974. Studies on L-serine deaminase in *Escherichia coli* K-12. J. Bacteriol. *118*:53-58.

Iuchi, S., D.C. Cameron, and E.C.C. Lin. 1989. A second global regulatory gene (*arcB*) mediating repression of enzymes in aerobic pathway of *Escherichia coli*. J. Bacteriol. *171*: 868-873.

Iuchi, S., V. Chepurl, and E.C.C.Lin. 1990. Requrlement for terminal cytochromes in generation of the aerobic signal for the *arc* regulatory system in *Escherichia coli*: study utilizing deletions and *lac* fusions of *cyo* and *cyd*. J.Bacteriol. *172*:6020-6025.

Iuchi, S., and E.C.C. Lin, 1991. Adaptation of *Escherichia coli* to respiratory conditions: regulation of gene expression. Cell *66*:5-7.

Iuchi, S., and E.C.C. Lin. 1988. *acrA* (*dye*), a global regulatory gene in *Escherichia coli* mediating repression of enzymes in aerobic pathways. Proc. Natl. Acad. Sci. USA *85*:1888-1892.

Iuchi, S., and E.C.C. Lin. 1987. The *narL* gene product activates the nitrate reductase operon and represses the fumarate reductase and trimethylamine N-oxide reductase operons in *Escherichia coli*. Proc. Natl. Acad. Sci. USA *84*:3901-3905.

Jenkins, D.E., E.A. Auger, and A. Matin. 1991. Role of RpoH, a heat shock regulator protein, in *Escherichia coli* carbon starvation protein synthesis and survival. J. Bacteriol. *173*:1992-1996.

Jenkins, D.E., S.A. Chaisson, and A. Matin. 1990. Starvation-induced cross protection against osmotic challenge in *Escherichia coli*. J. Bacteriol. *172*: 2779-2781.

Jenkins, D.E., J.E. Schultz, and A. Matin. 1988. Starvation-induced cross protection against heat or H_2O_2 challenge in *Escherichia coli*. J. Bacteriol. *170*:3910-3914.

Johnson, R.C., and M.I. Simon. Hin-mediated site-specific recombination requires two 26 bp recombination sites and a 60 bp recombinational enhancer. 1985. Cell *41*:781-791.

Johnson, W.C., C.P. Moran, and R. Losick. 1983. Two RNA polymerase sigma factors from *Bacillus subtilis* discriminate between overlapping promoters for a developmentally regulated gene. Nature *302*: 800-804.

Kahmann, R., F. Rudt, C. Koch, and G. Mertens. 1985. G Inversion in Bacteriophage Mu DNA is stimulated by a site within the Invertase gene and a host factor. Cell *41*:771-780.

Kalman, L.V., and R.P. Gunsalus. 1988. The *frdR* gene *Escherichia coli* globally regulates several operons involved in anaerobic growth in response to nitrate. J. Bacteriol. *170*:623-629.

Kalman, L.V., and R.P. Gunsalus. 1989. Identification of a second gene involved in global regulation of fumarate reductase and other nitrate-controlled genes for anaerobic respiration in *Escherichia coli*. *171*:3810-3816.

Kalman, L.V., and R.P. Gunsalus. 1990. Nitrate-and molybdenum-independent signal transduction mutations in *narX* that alter regulation of anaerobic respiratory genes in *Escherichia coli*. *172*:7049-7056.

Kassevetis, G.A., and E.P. Geiduschek. 1984. Defining a bacteriophage T4 late promoter: bacteriophage T4 gene55 protein suffices for directing late promoter recognition. Proc. Natl. Acad. Sci. USA *81*:5101-5105.

Kenyon, C.J., and G.C. Walker. 1980. DNA damaging agents stimulata gene expression at specific loci in *Escherichia coli*. Proc. Natl. Acad. Sci. USA *77*:2819-2823.

Kenyon, C.J., R. Brent, M. ptashen, and G.C. walker. 1982. Regulation of damage-inducible genes in *Escherichia coli*. J. Mol. Biol. *160*:445-457.

Kochan., and H. Murialdo. 1982. Stimulation of *groE* synthesis in *Escherichia coli* by bacteriophage lambda infection. J. Bacteriol. *149*:1166-1170.

Kohara, Y., K. Akiyama, and K. Isono. 1987. The physical map of the whole *Escherichia coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell *50*:495-508.

Kraus, J., D. Soll, and K.B. Low. 1979. Glutamyl-gamma-methyl ester acts as a methionine analogue in *E. coli*: analogue resistant mutants map at the *metJ* and *metK* loci. Genet. Res. *33*:49-55.

Kredich, N.M., and G.M. Tomkins. 1966. The enzymatic synthesis of L-cysteine in *Escherichia coli* and *Salmonella typhimurium*. J. Biol. Chem. *241*:4955-4965.

Krueger, J.H., and G.C. Walker. 1984. *groEL* and *dnaK* genes of *Escherichia coli* are induced by UV irradiation and nalidixic acid in an *htpR*$^+$-depentent fashion. Proc. Natl. Acad. Sci. *81*:1499-1503.

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

Lam, H.M., and M.E. Winkler. 1990. Metabolic relationships between pyridoxine (Vitamin B_6) and serine biosynthesis in *Escherichia coli* K-12. J. Bacteriol. *172*:6518-6528.

Lambden, P.R., and J.R. Guest. 1976. Mutants of *Escherichia coli* K12 unable to use fumarate as an anaerobic electron acceptor. J. Gen. Microbiol. *97*:145-160.

Lamond, A.I., and A.A. travers. 1983. Requirement for an upstream element for optimal transcription of a bacterial tRNA gene. Nature*305*:248-250.

Lang, B.F., and G. Burger. 1990. A rapid, high resolution DNA sequencing gel system. Anal. Biochem. *188*:176-180.

Lange, R., and R. Hengge-Aronis. 1991a. Identification of a central regulator of stationary-phase gene expression in *Escherichia coli*. Mol. Microbiol. *5*:49-59.

Lange, R., and R. Hengge-Aronis. 1991. Growth phase-regulated expression of *bolA* and morphology of stationary-phase *Escherichia coli* cells are controlled by the novel sigma factor σ^S. J. Bacteriol. *173*:4474-4481.

Lee, N., C. Francklyn, and E.P. Hamilton. 1987. Arabinose-induced binding of AraC protein to $ara I_2$ activates the *araBAD* operon promoter. Proc. Natl. Acad. Sci. USA *84*:8814-8818.

Leirmo, S. and R.L. Gourse. 1991. Factor-independent activation of *Escherichia coli* rRNA transcription I. kinetic analysis of the role of the upstream activator region and supercoiling on transcription of the *rrnP1* promoter *in vitro*. J. Mol. Biol. *220*:555-568.

Lemaux, P.G., S.L. Herendeen, P.L. Bloch, and F.C. Neidhardt. 1978. Transient rates of synthesis of individual polypeptides in *E. coli* K-12 following temperature shifts. Cell *13*:427-434.

Lerner, T., and N. Zinder. 1979. Chromosomal regulation of sexual expression in *Escherichia coli*. J. Bacteriol. *137*:1063-1065.

Lenny, A.B., and P. Margolin. 1980. Location of the *opp* and *supX* genes of *Salmonella* and *Escherichia coli*. J. Bacteriol. *143*:747-752.

Lessie, T.G., and H.R. Whiteley. 1969. Properties of threonine deaminase from a bacterium able to use threonine as sole source of carbon. J. Bacteriol. *100*:878-889.

Leveque, F., M. Gazeau, M. Fromant, S. Blanquet, and P. Plateau. 1991. Control of *Escherichia coli* Lysys-tRNA synthetase expression by anaerobiosis. J. Bacteriol. *173*:7903-7910.

Leveque, F., P. Plateau, P. Dessen, and S. Blanquet. 1990. Homology of *lysS* and *lysU*, the two *Escherichia coli* genes encoding distinct lysyl-tRNA synthetase species. Nucleic Acids Res. *18*:305-312.

Li, X., and J. Krakow. 1988. Monoclonal antibodies that inhibit activation of transcription by the *Escherichia coli* cyclic AMP receptor protein. J. Biol. Chem. *263*:3448-3453.

Lindquist, S. 1986. The heat shock response. Ann.Rev. Biochem. *55*:1151-1191.

Lipinska, B., O. Fayet, L. Baird, and C. Georgopoulos. 1989. Identification, characterization, and mapping of the *Escherichia coli htrA* gene, whose product is essential for bacterial growth only at elevated temperatures. J. Bacteriol. *171*:1574-1584.

Little, J.W. 1983. The SOS regulatory system: control of its state by the level of *recA* protease. J. Mol. Biol. *167*:791-808.

Little, J.W., and D.W. Mount. 1982. The SOS regulatory system of *Escherichia coli*. Cell *29*:11-22.

Little, J.W., D.W. Mount, and C.R. Yanisch-Perron. 1981. Purified lexA protein is a repressor of the *recA* and *lexA* genes. Proc. Natl. Acad. Sci. USA *78*:4199-4203.

Liu-Johnson, H.N., M.R. Gartenbarg, and D. crothers. The DNA binding domain and bending angle of *E. coli* CAP protein. Cell *47*:995-1005.

Lobell, R.B., and R.F. Schleif. 1991. AraC-DNA looping: Orientation and distance-dependent loop breaking by the cyclic AMP receptor protein. J. Mol. Biol. *218*:45-54.

Losick, R., and J. pero. 1981. Cascades of sigma factors. Cell *25*:582-584.

Lowry, O.H., J. Rosebrough, A.L. Farr, and R. J. Randall. 1951. Protein measurement with Folin phenol reagent. J. Biol. Chem. *193*:265-275.

Lobell, R.B., and R.F. Schleif. 1991. Arac-DNA Looping: orientation and distance-dependent loop breaking by the cyclic AMP receptor protein. J. Mol. Biol. *218*:45-54.

Maas, W.K., and A.J. Clark. 1964. Studies on the mechanism of repression of arginine biosynthesis in *Escherichia coli*. II. Dominance of repressibility in diploids. J. Mol. Biol. *8*:365-370.

Magasanik, B. 1961. Catabolite repression. Cold Spring Harbor Symp. Quant. Biol. *26*:249-256.

Magasanik, B. 1970. Glucose effects: Inducer exclusion and repression. p. 189-219. In The lactose operon. J.R. Beckwith and D. Zipser (ed.). Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Magasanik, B., and F. C. Neidhardt. 1987. Regulation of carbon and nitrogen utilization. p.1318-1325. *In Escherichia coli* and *Salmonella typhimurium* Cellular and Molecular Biology. F.C. Neidhardt (ed.). American Society for Microbiology, Washington.

Majors, J. 1975. Initiation of in vitro mRNA synthesis from wild-type *lac* promoter. Proc. Natl. Acad. Sci. USA *72*:4394.

Makino, k., H. Shinagawa, M.Amemura, S. Kimura, and A. Nakata. 1988. Regulation of the phosphate regulon of *Escherichia coli*. activation of *pstS* transcription by PhoB protein *in vitro*. J. Mol. Biol. *203*:85-95.

Malan, T.P., A. Kolb, H. Buc, and W.R. McClure. 1984. Mechanism of CRP-cAMP activation of *lac* operon transcription initiation. activation of P1 promoter. J. Mol. Biol. *180*:881-909.

Malik, S., M. Dimitrov, and A. Goldfarb. 1985. Initiation of transcription by bacteriophage T4-modified RNA polymerase independently of host sigma factor. J. Mol. Biol. *185*:83-91.

Mandal, N., W. Su, R. Haber, S. Adhya, and H. Echols. 1990. DNA looping in cellular repression of transcription of the galactose operon. Genes Dev. *4*:410-418.

Maniatis, T., E.F.Fritsch, and J. Sambrook. 1982. In Molecular Cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Martin, K., L. Huo, and R. Schleif. 1986. The DNA loop model for *ara* repression: AraC protein occupies the proposed loop sites in vivo and repression-negative mutations lie in these same sites. Proc. Natl. Acad. Sci. USA *83*:3654-3658.

Martinussen, J., N.E. MΦllegaard, B. Holst, S.R. Douthwaite, and P. Valentin-Hansen. 1989. A new version of negative control. DNA sequences involved in expression and regulasion of CytR and cAMP/CRP controlled genes in *Escherichia coli*. pp. 31-41. *In* DNA-protein interactions in transcription. J.D. Gralla (ed.). Alan R. Liss, Inc. New York.

Matthews, R.G., and F.C. Neidhardt. 1988. Abnormal induction of heat shock proteins in an *Escherichia coli* mutant deficient in adenosylmethionine synthetase activity. J. Bacteriol. *170*:1582-1588.

Matthews, R.G., and F.C. Neidhardt. 1989. Elevated serine catabolism is associated with the heat shock response in *Escherichia coli*. J. Bacteriol. *171*:2619-2625.

Mauzerall, D., and S. Granick. 1956. The occurrence and determination of delta-aminolevulinic acid and porphobilinogen in urine. J. Biol. Chem. *219*:435-446.

McAllister, C.F., and E.C. Achberger. 1988. Effect of polyadenine-containing curved DNA on promoter utilization in *Bacillus subtilis*. J. Biol. Chem. *263*:11743-11749.

McFall, E. 1987. The D-serine deaminase operon. pp.1520-1526. *In Escherichia coli* and *Salmonella typhimurium* Cellular and Molecular Biology. F.C. Neidhardt (ed.). American Society for Microbiology, Washington.

McGilvray, D., and J.G. Morris. 1971. L-Threonine dehydrogenase, pp. 580-584. In H. Tabor and C.W. Tabor (ed.), Methods in enzymology, vol. 17, part B. Academic Press Inc., New York.

McKitrick,J. C., and L.I. Pizer. 1980. Regulation of phosphoglycerate dehydrogenase levels and effect on serine synthesis in *Escherichia coli* K-12. J. Bacteriol. *141*:235-245.

Mecsas, J., D.W. Cowing, and C.A. Gross. 1991. Development of RNA polymerase-promoter contacts during open complex formation. J. Mol. Biol. *220*:585-597.

Merrick, M.J. 1983. Nitrogen control of the *nif* regulon in *Klebsiella pneumoniae*: involvement of the *ntrA* gene and analogies between *ntrC* and *nifA*. EMBO J. *2*:39-44.

Miller, J.H. (ed). 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Miller, H.I., M. Kirk, and H. Echols. 1981. SOS induction and autoregulation of the *himA* gene for site-specific recombination in *Escherichia coli*. Proc. Natl. Acad. Sci. USA *78*:6754-6758.

Minchin, S.D., S. Austin, and R.A. Dixon. 1989. Transcriptional activation of the *Klebsiella pneumoniae nifLA* promoter by NTRC is face-of-the-helix dependent and activator stabilizes the interaction of sigma 54-RNA polymerase with the promoter. EMBO J. *8*:3491-3499.

Mizuno, T., M.Y. Chou, and M. Inouye. 1983. A comparative study on the genes for three porins of the *Escherichia coli* outer membrane. J. Biol. Chem. *258*:6932-6940.

Mizusawa, S., D. Court, and S. Gottesman. 1983. Transcription of the *sulA* gene and repression of *lexA*. J. Mol. Biol. *171*: 337-343.

Mitchell, P.J. and R. Tijan. 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science *245*:371-378.

Moltoso de Vargas, L., S. Kim, and A. Landy. DNA looping generated by DNA bending protein IHF and two domains of lambda integrase. Science *244*:1457-1461.

Morgan, R.W., M.F. Christman, F.S. Jacobson, G. Storz and B.N. Ames. 1986. Hydrogen peroxide-inducible proteins in *Salmonella typhimurium* ovelap with heat shock and other stress proteins. Proc. Natl. Acad. Sci. USA *83*:8059-8063.

Morris,J.F., and E.B. Newman. 1980. Map location of the *ssd* mutation in *Escherichia coli* K-12. J. Bacteriol. *143*:1504-1505.

Mortensen, L., G. Dandanell, and K. Hammer. 1989. Purification and characterization of the deoR repressor of *Escherichia coli*. EMBO J. *8*:325-331.

Moskaluk, C., and D. Bastia. 1988. DNA bending is induced in an enhancer by the DNA-binding domain of the bovine papillomavirus E2 protein. Proc. Natl. Acad. Sci. USA *81*:4692-4696.

Mossing, M.C. and M.T. Record,Jr. Upstream operators enhance repression of the *lac* promoter. Science *233*:889-892.

Mulvey, M.R., and P.C. Loewen. 1989. Nucleotide sequence of *katF* of *Escherichia coli* suggests KatF protein is a novel σ transcription factor. Nucleic Acids Res. *17*:9979-9991.

Munch-Petersen, A., and J. Messing. 1983. Transport of nucleic acid precursors, p. 259-305. In Metabolism of nucleotides, nucleosides and nucleobases in microorganisms. A. Munch-Petersen (ed.). Academic Press, Inc. (London), Ltd., London.

Musso, R.E., R. DiLauro, S. Adhya, and B. de Crombrugghe. 1977. Dual control for transcription of the *galactose* operon by cyclic AMP and its receptor protein at two interspersed promoters. Cell *12*:847-854.

Nash, H. and C. Robertson. 1981. Purification and properties of the *Escherichia coli* protein factor required for λ integrative recombination. J. Biol. Chem. *256*:9246-9253.

Nazos, P.M., T.S. Su, R.C. Landick, and D.L. Oxender. 1984. Branched-chain amino acid transport in *Escherichia coli*. P. 24, *In* Microbiology-1984. Amaerican Society of Microbiology, Washington D.C.

Neidhardt, F.C., and R.A. VanBogelen. 1981. Positive regulatory gene for temperature-controlled proteins in *E. coli*. Biochem. Biophys. Res. Commun. *100*:894-900.

Neidhardt, F.C., and R.A. VanBogelen. 1987. Heat shock response. pp.1334-1345. *In Escherichia coli* and *Salmonella typhimurium* Cellular and Molecular Biology. F.C. Neidhardt (ed.). American Society for Microbiology, Washington.

Neidhardt, F.C., R.A. VanBogelen, and E.T. Lau. 1983. Molecular cloning and expression of a gene that controls the high-temperature regulon of *Escherichia coli*. J. Bacteriol. *153*:597-603.

Neidhardt, F.C., R.A. VanBogelen, and V. Vaughn. 1984. Genetics and regulation of heat-shock proteins. Ann. Rev. Genet. *18*:295-329.

Newman, E.B., D. Ahmad, and C. Walker. 1982a. L-Serine deaminase activity is induced by exposure of *Escherichia coli* K-12 to DNA-damaging agents. J. Bacteriol. *152*:702-705.

Newman, E.B., G. Batist, J. Fraser, S. Isenberg, P. Weyman, and V. Kapoor. 1976. The use of glycine as nitrogen source by *Escherichia coli* K-12. Biochim. Biophys. Acta *421*:97-105.

Newman, E.B., D. Dumont, and C. Walker. 1985a. *In vitro* and *in vivo* activation of L-serine deaminase in *Escherichia coli* K-12. J. Bacteriol. *162*:1270-1275.

Newman, E.B., and V. Kapoor. 1980. *In vitro* studies on L-serine deaminase activity of *Escherichia coli* K-12. Can. J. Biochem. *58*:1292-1297.

Newman, E.B., V. Kapoor, and R. Potter. 1976. Role of L-threonine dehydrogenase in the catabolism of threonine and synthesis of glycine by *Escherichia coli*. J. Bacteriol. *126*:1245-1249.

Newman, E.B., and B. Magasanik. 1963. The relation of serine-glycine metabolism to the formation of single carbon units. Biochim. Biophys. Acta 76:437-448.

Newman, E.B., Malik, N., and C. Walker. 1982b. L-Serine degradation in *Escherichia coli* K-12: Directly isolated *ssd* mutants and their intragenic revertants. J. Bacteriol. *150*:710-715.

Newman, E.B., B. Miller, L.D. Colebrook, and C. Walker. 1985b. A mutation in *Escherichia coli* K-12 results in a requirement for thiamine and a decrease in L-serine deaminase activity. J. Bacteriol. *161*:272-276.

Newman, E.B., B. Miller, and V. Kapoor. 1974. Biosynthesis of single-carbon units in *Escherichia coli* K-12. Biochim. Biophys. Acta *338*:529-539.

Newman, E.B., J.F. Morris, C. Walker, and V. Kapoor. 1981. A mutation affecting L-serine and energy metabolism in *Escherichia coli* K-12. Mol. Gen. Genet. *182*:143-147.

Newman, E.B., and C. Walker. 1982. L-Serine degradation in *Escherichia coli* K-12: a combination of L-serine, glycine and leucine used as a source of carbon. J. Bacteriol. *151*:777-782.

Nilsson, L., A. Vanet, E. Vijgenboom and L. Bosch. 1990. The role of FIS in trans activation of stable RNA operons of *E. coli*. EMBO J. *9*:727-734.

Ninfa, A.J., E. Brodsky, and B. Magasanik. 1989. The role of NR_I-phosphate in the activation of transcription from the nitrogen regulated promoter *glnAp2* of *Escherichia coli*. pp.43-52. In DNA-protein interactions in transcription. J. D. Gralla (ed.). Alan R. Liss, Inc. New York.

Ninfa, A.J., E.G. Ninfa, A.N. Lupas, A. Stock, B. Magasanik, and J. Stock. 1988. Crosstalk between bacterial chemotaxis signal transduction proteins and regulators of transcription of the Ntr regulon: Evidence that nitrogen assimilation and chemotaxis are controlled by a common phosphotransfer mechanism. Proc. Natl. Acad. Sci. USA *85*:5492-5496.

Ninfa, A.J., L.J. Reitzer, and B. Magasanik. 1987. Initiation of transcription at the bacterial *glnAp2*promoter by purified *E. coli* components is facilitated by enhancers. Cell *50*:1039-1046.

Oehler, S., E.R. Eismann, H. Kramer, and B. Muller-Hill. 1990. The three operators of the *lac* operon cooperate in repression. EMBO J. *9*:973-979.

O'Farrell, P.H. 1975. High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem. *250*:4007-4021.

Oxender, D.L., J.J. Anderson, M.M. Mayo, and S.C. Quay. 1977. Leucine binding and regulation of transport in *E. coli*. J. Supramolecular Structure *6*:419-431.

Pabo, C.O., and R.T. Sauer. 1984. Protein-DNA recognition. Ann. Rev. Biochem. *53*:293-321.

Pahel, G., A.D. Zelenetz, and B.M. Tyler.1978. *gltB* gene and regulation of nitrogen metabolism by glutamate synthase in *Escherichia coli*. J. Bacteriol. *133*:139-148.

Pardee, A.B., and L.S. Prestidge. 1955. Induced formation of serine and threonine deaminase by *Escherichia coli*. J. Bacteriol. *70*:677-674.

Payne, J.W., and C. Gilvarg. 1968. Size restriction on peptides utilization in *E. coli*. J. Biol. Chem. *243*:5291-6299.

Pedersen, S., P.L. Bloch, S. Reeh, and F.C. Neidhardt. 1978. Patterns of protein synthesis in *E. coli*: A catalog of the amount of 140 individual proteins at different growth rates. Cell *14*:179-190.

Plakto, J.V., D.A. Willins, and J.M. Calvo. 1990. The *IlvIH* operon of *Escherichia coli* is positively regulated. J. Bacteriol. *170*:4563-4570.

Plate, C. 1976. Mutant of *Escherichia coli* defective in response to colicin K and in active transport. Proc. Natl. Acad. Sci. USA *125*:467-474.

Plate, C., and J. Suit. 1981. The *eup* genetics locus of *Escherichia coli* K-12 and its role in H^+/solute symport. J. Biol. Chem. *256*:12974-12980.

Platko, J.V., D.A. Willins, and J.M. Calvo. 1990. The *IlvIH* operon of *Escherichia coli* is positively regulated. J. Bacteriol. *172*:4563-4570.

Popham, D.L., D. Szeto, J. Keener, and S. Kustu. 1989. Function of a bacterial activator protein that binds to transcriptional enhancers. Science *243*:629-635.

Potter, R. V. Kapoor, and E.B. Newman. 1977. Role of threonine dehydrogenase in *Escherichia coli* threonine degradation. J. Bacteriol. *132*:385-391.

Ptashne, M. 1986. Gene regulation by proteins acting nearby and at a distance. Nature *322*: 697-701.

Ptashne, M. 1986a. A genetic switch. Blackwell Scientific Publications, Palo Alto, Calif. and Cell press, Cambridge, Mass.

Ptashne, M. 1991. Gene regulation. Cancer *67*:2422-2427.

Quay, S.C., T.E. Dick, and D.L. Oxender. 1977. Role of transport systems in amino acid metabolism: leucine toxicity and the branched-chain amino acid transport systems. J. Bacteriol. *129*:1257-1265.

Ralbaud, O., and M. Schwartz. 1984. Positive control of transcription initiation in bacteria. Ann. Rev. Genet. *18*:173-206.

Ralbaud, O., D. Vidal-Ingigliardi, and E. Richet. 1989. A complex nucleoprotein structure involved in activation of transcription of two divergent *Escherichia coli* promoters. J. Mol. Biol. *205*:471-485.

Rainwater, S., and P.M. Silverman. 1990. The cpx proteins of *Escherichia coli* K-12: evidence that *cpxA*, *ecfB*, *ssd*, and *eup* mutations all identify the same gene. J. Bacteriol. *172*:2456-2461.

Ramotar, D., and E.B. Newman. 1986. An estimate of the extent of deamination of L-serine in auxotrophs of *Escherichia coli* K-12. Can. J. Microbiol. *32*:842-846.

Ravnikar, P.D., and R.L. Somerville. 1987. Structural and functional analysis of a cloned segment of *Escherichia coli* DNA that specifies proteins of a C_4 pathway of serine biosynthesis. J. Bacteriol. *169*:4716-4721.

Ray, L., F. Claverie-Martin, P. Weglenski, and B. Magasanik. 1990. Role of the promoter in activation of transcription by nitrogen regulator I phosphate in *Escherichia coli*. J. Bacteriol. *172*:818-823.

Reitzer, L.J., and B. Magasanik. 1985. Expression of *glnA* in *Escherichia coli* is regulated at tandem promoters. Proc. Natl. Acad. Sci. USA *82*:1979-1983.

Reitzer, L.J., and B. Magasanik. 1987. Ammonia assimilation and the biosynthesis of Glutamine, glutamate, aspartate, asparagine, L-alanine, and D-alanine. pp.302-320. *In* *Escherichia coli* and *Salmonella typhimurium* Cellular and Molecular Biology. F.C. Neidhardt (ed.). American Society for Microbiology, Washington.

Ren, Y.L., S. Garges, S. Adhya, and J.S. Krakow. 1988. Cooperative DNA binding of heterologus proteins: evidence for contact between the cyclic AMP receptor protein and RNA polymerase. Proc. Natl. Acad. Sci. USA *85*:4138-4142.

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

Reznikoff, W.F. 1992. Catabolite gene activator protein activation of *lac* transcription. J. Bacteriol. *174*:655-658.

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

Richet, E., and O. Raibaud. 1991. Supercoiling is essential for the formation and stability of the initiation complex at the divergent *malEp* and *malKp* promoter. J. Mol. Biol. *218*:529-542.

Riggs, A., H. Suzuki, and S. bourgeois. 1970. *lac* repressor-operator interaction. I. Equilibrium studies. J. Mol. Biol. *48*:67-83.

Roberts, J. W. 1987. Regulation of protein synthesis and function in bacteria. pp465-502. In Molecular biology of the gene. Watson, J.D., N.H. Hopkins, J.W. Roberts, J.A. Steitz, and A.M. Weiner (eds.). The Benjamin/Cummings Publishing Company, Inc.

Roberts, J.W., and C.W. Roberts. 1975. Proteolytic cleavage of bacteriophage lambda repressor in induction. Proc. Natl. Acad. Sci. USA 72:147-151.

Rosenberg, H., R.G. Gerdes, and K. Chegwidden. 1977. Two systems for the uptake of phosphate in *Escherichia coli*. J. Bacteriol. *131*:505-511.

Ronson, C.W., B.T. Nixon,and F.M. Ausubel. 1987. Conserved domains in bacterial regulatory proteins that respond to environmental stimuli. Cell *49*:579-581.

Ross, W., J.F. Thompson, J.T. Newlands ,and R.L. Gourse. 1990. *E. coli* Fis protein activates ribosomal RNA transcription *in vitro* and *in vivo*. EMBO J. *9*:3733-3742.

Sak, B.D., A. Eisenstark, and D. Touati. 1989. Exonuclease III and the catalase hydroperoxidase II in *Escherichia coli* are both regulated by the *katF* gene product. Proc. Natl. Acad. Sci. USA *86*:3271-3275.

Sancar, A., G.B. Sancar, W.D. Rupp, J.W. Little, and D.W. Mount. 1982. LexA protein inhibits transcription of the *E. coli uvrA* gene *in vitro*. Nature *298*:96-98.

Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA. *74*:5463-5467.

Sasse-Dwight, S., and J.D. Gralla. 1988. Probing co-opretive DNA-binding *in vivo*: the *lac* O1:O3 interaction. J. Mol. Biol. *202*:107-119.

Sasse-Dwight, S., and J.D. Gralla. 1990. Role of eukaryotic-type fuctional domains found in the prokaryotic enhancer receptor factor sigma 54. Cell *62*:945-954.

Satishchandran, C., J.C. Taylor, and G.D. Markham. 1990. Novel *Escherichia coli* K-12 mutants impaired in S-adenosylmethionine synthesis. J. Bacteriol. *172*:4489-4496.

Sawers, G., and A. Bock. 1988. Anaerobic regulation of pyruvate formate-lyase from *Escherichia coli* K-12. J. Bacteriol. *170*:5330-5336.

Schleif, R. 1987. Why should DNA loop? Nature *327*:369-370.

Scrimgeour,K.G., and F.M. Huennekens. 1962. Serine hydroxymethylase. Methods in Enzymol. *5*:794-809.

Shih, M.C., and G.N. Gussin. 1984. Role of cII protein in stimulating transcription initiation at the λ P_{RE} promoter. J. Mol. Biol. *172*:489-506.

Shinagawa, H., K. Makino, and A. Nakata. 1983. Regulation of the *pho* regulon in *Escherichia coli* K-12: genetic and physiological regulation of the positive regulatory gene *phoB*. J. Mol.Biol. *168*:477-488.

Shoeman, R., B. Redfield, T. Coleman, R.C. Greene, A.A. Smith, N. Brot, and H. Weissbach. 1985. Regulation of methionine synthesis in *Escherichia coli*: effect of *metJ* gene productand S-adenosylmethionine on the expresstion of the *metF*gene. Proc. Natl. Acad. Sci. USA *82*:3601-3605.

Silhavy, T.J., M.L. Berman, and L.W. Enquist. 1984. Experiments with gene fusions. Cold Spring Harbor laboratory. Cold Spring Harbor. N.Y.

Silverman, P.M. 1985. Host cell-plasmid interactions in the expression of DNA donor activity by F^+ strains of *Escherichia coli* K-12. Bioessays *2*:254-259.

Silverman, P.M., K. Nat, J. McEwen, and R. Birchman. 1980. Selection of *Escherichia coli* K-12 chromosomal mutants that prevent expression of F-plasmid functions. J. Bacteriol. *143*:1519-1523.

Silverman, P.M., S. Rother, and H. Gaudin. 1991. Arc and Sfr functions of the *Escherichia coli* K-12 *arcA* gene product are geneticlly and physiologically separable. J. Bacteriol. *173*:5648-5652.

Singer, M., T.A. Baker, J.Schnitzler, S.M. Deichel, M. Goel, W. Dove, K.J. Jaack, A.D. Grossman, J.W. Frickson, and C.A. Gross. 1989. A collection of strains containing genetically linked alternating antibiotic resistance for genetic mapping of *Escherichia coli*. Microbiol. Rev. *53*:1-24.

Smith, M.W., and F.C. Neidhardt. 1983a. Proteins induced by anaerobiosis in *Escherichia coli*. J. Bacteriol. *154*:336-343.

Smith, M.W., and F.C. Neidhardt. 1983b. Proteins induced by aerobiosis in *Escherichia coli*. J. Bacteriol. *154*:344-350.

Sϕgaard-Andersen, L. J. Martinussen, N.K. Mϕllegaard, S.R. Douthwaite, and P. Valentin-Hansen. 1990. The CytR repressor antagonizes cyclic AMP-cyclic AMP receptor protein activation of the $deoCp_2$ promoter of *Escherichia coli* K-12. J. Bacteriol. *172*:5706-5713.

Sϕgaard-Andersen, H. Pedersen, B. Holst, and P. Valentin-Hansen. 1991. A novel function of the cAMP-CRP complex in *Escherichia coli* K-12: cAMP-CRP functions as an adaptor for the CytR repressor in the *deo* operon. Mol. Microbiol. *5*:969-975.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J.Mol. Biol. 98:503-517.

Spiro, S., R.E. Roberts, and J.R. Guest. 1989. FNR-dependent repression of *ndh* gene of *Escherichia coli* and metal ion requirement for FNR-regulated gene expression. Mol. Microbiol. 3:601-608.

Squires, C.H., M. DeFelice, S.R. Wessler, and J.M. Calvo. 1981. Physical characterization of the *IlvIH* operon of *Escherichia coli* K-12. J. Bacteriol. *147*:797-804.

Stauffer, G.V. 1987. Biosynthesis of serine and glycine. PP. 412-418. *In Escherichia coli* and *Salmonella typhimurium* Cellular and Molecular Biology. Ed. Frederick C. Neidhardt. American Society for Microbiology, Washington.

Stauffer, L.T., M.D. plamann, and G.V. Stauffer. 1986. Cloning and characterization of the glycine-cleavage enzyme system of *Escherichia coli*. Gene *44*:219-226.

Stenzel, T.T., P, Patel, and D. Bastia. 1987. The integration host factor of *Escherichia coli* binds to bent DNA at the origin of replication of the plasmid pSC101. Cell *49*:709-717.

Stewart, V. 1982. Requirement of Fnr and NarL functions for nitrate reductase expression in *Escherichia coli* K-12. J. Bacteriol. *151*:1320-1325.

Stewart, V., and B. L. Berg. 1988. Influence of *nar* (nitrate reductase) genes on nitrate inhibition of formate hydrogenlyase and fumarate reductase gene expression in *Escherichia coli* K-12. J. Bacteriol. *170*:4437-4444.

Stewart, V., and J. Parales,Jr. 1988. Identification and expression of gene *narL* and *narX* of the *nar* (nitrate reductase) locus in *Escherichia coli* K-12. J. Bacteriol. *170*:1589-1597.

Stock, J.B., A.J. Ninfa and A.M. Stock. 1989. Protein phosphorylation and regulation of adaptive responses in bacteria. Microbiol. Rev. *53*:450-490.

Storz, G., L.A. Tartaglia, and B.N. Ames. 1990. Transcriptional regulator of oxidative stress-inducible genes: direct activation by oxidation. Science *248*:189-194.

Straney, D.C., S.B. Straney, and D.M. Crothers. 1989. Synergy between *Escherichia coli* RNA polymerase in the *lac* promoter open complex. J. Mol.Biol. *206*:41-57.

Stratagene Inc. Bluescript DNA sequencing system. Instruction Manual.

Su, H., B. F. Lang, and E.B. Newman. 1989. L-Serine degradation in *Escherichia coli* K-12. Cloning and sequencing of the *sdaA* gene. J. Bacteriol. *171*:5095-5102.

Su, H., and E.B. Newman. 1991. A novel L-serine deaminase activity in *Escherichia coli* K-12. J. Bacteriol. *173*:2473-2480.

Su, H. 1991. Molecular studies of the *sdaA* and *sdaB* genes and their gene products in *Escherichia coli* K-12. PhD thesis at Concordia University.

Su, W., S. Porter, S. Kustu, and H. Echols. 1990. DNA-looping and enhancer activity: association between DNA-bound NtrC activator and RNA polymerase at the bacterial *glnA*promoter. Proc. Natl. Acad. Sci. USA *87*:5504-5508.

Tabor, H., S.M. Rosenthal, and C.W. Tabor. 1958. Biosynthesis of spermidine and spermine from putrescine and methionine. J. Biol. Chem. *233*:907-914.

Tam, A., M.B. Herrington, and E.B. Newman. 1978. A single mutation affects L-serine deaminase, L-leucyl-, L-phenylalanyl-tRNA protein transferase, and proline oxidase activity in *Escherichia coli* K-12. J. Bacteriol. *135*:1154-1155.

Tartaglia, L.A., G. Storz, and B.N. Ames. 1989. Identification and molecular analysis of *oxyR*-regulated promoters important for the bacterial adaptation to oxidative stress. J. Mol. Biol. *210*:709-719.

Taylor, W.E., D.E. Straus, A.D. Grossman, Z.F. Burton, C.A. Gross, and R. Burgess. 1984. Transcription from a heat-inducible promoter causes heat shock regulationof the sigma subunit of *E. coli* RNA polymerase. Cell *38*:371-381.

Thorbjarnardottir, S., R. Magnusdottir, G. Eggertsson, S. Kagan, and D. Andresson. 1978. Mutations determining generalized resistance to aminoglycoside antibiotics in *Escherichia coli*. Mol. Gen. Genet. *161*:89-98.

Tobey, K.L., and G.A. Grant. 1986. The nucleotide sequence of *serA* gene of *Escherichia coli* and the amino acid sequence of the encoded protein, D-3-phosphoglycerate dehydrogenase. J. Biol. Chem. *261*:12179-12183.

Tommassen, J., P.Degeus, and B. Lugtenberg. 1982. Regulation of the pho regulon of *E. coli* K12. Cloning of the regulatory genes *phoB* and *phoR* and identification of their gene products. J. Mol. Biol. *157*:265-274.

Tran, V.S., E. Schaeffer, O. Bertrand, R. Marluzza, and P. Ferrara. 1983. Purification, molecular weight and N-terminal sequence of cystathionine-r-synthase of *Escherichia coli*. J. Biol. Chem. *258*:14872-14873.

Tsaneva, I.R., and B. weiss. 1990. *soxR*, a locus governing a superoxide response regulon in *Escherichia coli* K-12. J. Bacteriol. *172*:4197-4205.

Tsui, P., L. Huang, and M. Freunglich. 1991. Integration host factor binds specifically to multiple sites in the *ompB* promoter of *Escherichia coli* and inhibits transcription. J. Bacteriol. *173*:5800-5807.

Tsung, K., R.E. Brissette, and M. Inouye. 1990. Enhancement of RNA polymerase binding to promoters by a transcriptional activator, OmpR, in *Escherichia coli*: Its positive and negative effects on transcription. Proc. Natl. Acad. Sci. USA *87*:5940-5944.

Tullius, T.D. 1989. Physical studies of protein-DNA complexes by footprinting. Ann. Rev. Biophysics and Biophysical Chemistry *18*:213-237.

Ullmann, A., and A. Danchin. 1983. Role of cyclic AMP in bacteria. Adv. Cyclic Nucleotide Res. *15*:1-53.

Umbarger, H.E., M.A. Umbarger, and P.L.M. Siu. 1963. Biosynthesis of serine in *Escherichia coli* and *Salmonella typhimurium*. J. Bacteriol. *85*:1431-1439.

Umbarger, H.E. 1987. Biosynthesis of the branched-chain amino acids. In *Escherichia coli* and *Salmonella typhimurium* Cellular and Molecular Biology. Ed. Frederick C. Neidhardt. American Society for Microbiology, Washington.

Urata, G., and S. Granick. 1963. Biosynthesis of aminoketones and the metabolism of aminoacetone. J. Biol. Chem. *238*:811-820.

Uzan, M., and A. Danchin. 1978. Correlation between the serine sensitivity and the derepressibility of the *ilv* genes in *Escherichia coli relA*⁻ mutants. Mol. Gen. Gent. *165*:21-30.

Valentin-Hansen, P. 1985. DNA sequences involved in expression and regulation of *deoR*, *cytR* and cAMP/CRP-controlled genes in *IN scherichia coli*. p. 273-288. In Gene manipulation and expression. R.E. Glass, J. Spizek (ed.). London: Croom Helm.

Valentin-Hansen, P., J.E. Lφve Larsen, P. Hojrup, S.A. Short, and C.S. Barbier. 1986. Nucleotide sequence of the *cytR* regulatory gene of *E. coli* K-12. Nucleic Acids Res. *14*:2215-2228.

Van Bogelen, R.A., P.M. Kelley, and F.C. Neidhardt. 1987. Differential induction of heat-shock, SOS, and oxidation-stress regulons and accumulation of nucleotides in *Escherichia coli*. J. Bacteriol. *169*:26-32.

Van Bogelen, R.A., V.Vaughn, and F.C. Neidhardt. 1983. Gene for heat-inducible lysyl-tRNA synthetase (*lysU*) maps near *cadA* in *Escherichia coli*. J. Bacteriol. *153*:1066-1068.

van Delft, J.H.M., B. Marlon, D.S. Schmidt, and L. Bosch. 1987. Transcription of the tRNA-*tufB* operon of *Escherichia coli*: activation, termination and antitermination. Nucleic Acids Res. *15*:9515-9530.

Van Vliet, F., A. Jacobs, J. Piette, D. Gigot, M. Lauwreys, A. Plerard, and N. Glansdorff. 1984. Nucleic Acids Res. *12*:6277-6289.

Vershon, A.K., S.M. Liao, W.R. McClure, and R.T. Sauer. 1987. Interaction of the bacteriophage P22 Arc repressor with operator DNA. J. Mol. Biol. *195*:323-331.

Vidal-Ingigliardi, D., E. Richet, and O. Raibaud. 1991. Two MalT binding sites in direct repeat: A structural motif involved in the action of all the promoters of the maltose regulons in *Escherichia coli* and *Klebsiella pneumoniae*. J. Mol. Biol. *218*:323-334.

Vieira, J., and J. Messing. 1987. Production of single-stranded plasmid DNA. Methods In Enzymology. *153*:3-11.

von Hippel, P.H., and O.G. Berg. 1989. Facilitated target location in biological systems. J. Biol. Chem. *264*:675-678.

von Ossowski, I., M.R. Mulvey, P.A. Leco, A. borys, and P.C. Loewen. 1991. Nucleotide sequence of *Escherichia coli katE*, which encloses catalase HPII. J. bacteriol. *173*:514-520.

Walker, G.C. 1984. Mutagenesis and inducible responses to DNA damage in *Escherichia coli*. Microbiol. Rev. *48*:60-93.

Walker, G.C. 1985. Inducible DNA repair systems. Ann. Rev. Biochem. *54*:425-457.

Walker, G.C. 1987. The SOS response of *Escherichia coli*. p.1346-1357. *In Escherichia coli* and *Salmonella typhimurium* Cellular and Molecular Biology. F.C. Neidhardt (ed.). American Society for Microbiology, Washington.

Walkup, L.K.B., and T. Kogoma. 1989. *Escherichia coli* proteins inducible by oxidative stress mediated by the superoxide radical. J. Bacteriol. *171*: 1476-1484.

Wanner, B.L. 1983. Overlapping and separate controls on the phosphate regulon in *Escherichia coli* K12. J. Mol. Biol. *166*:283-308.

Wanner, B.L. 1987. Phosphate regulation of gene expression in *Escherichia coli*. pp.1326-1333. *In Escherichia coli* and *Salmonella typhimurium* Cellular and Molecular Biology. F.C. Neidhardt (ed.). American Society for Microbiology, Washington.

Wanner B.L., and R. McSharry. 1982. Phosphate-controlled gene expression in *Escherichia coli* K-12 using *Mud1*-directed *lacZ* fusion. J. Mol. Biol. *158*:347-363.

Weber, R.F., and P.M. Silverman. 1988. The Cpx proteins of *Escherichia coli* K-12: structure of the CpxA polypeptide as an inner membrane component. J. Mol. Biol. *203*:467-478.

Wedel, A., D.S. Weiss, D. Popham, P. Droge, and S. Kustu. 1990. A bacterial enhancer functions to tether a transcriptional activator near a promoter. Science *248*:486-490.

Wek, R.C., and G.W. Hatfield. 1988. Transcriptional activation at adjacent operators in the divergent-overlapping *IlvY* and *IlvC* promoters of *Escherichia coli*. J. Mol.Biol. *203*:643-663.

Wessler, S.R., and J.M. Calvo. 1981. Control of *leu* operon expression in *Escherichia coli* by a transcription attenuation methanism. J. Mol. Biol. *149*:579-597.

Willets, A.J. 1972. Metabolism of threonine by fusaria: growth of threonine as the sole carbon and nitrogen source. Antonie van Leeuwenhoek J. Microbiol. Serol. *38*:591-603.

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

Willsky, G.R., and M.H. Malamy. 1980. Characterization of two genetically separable inorganic phosphate transport systems in *Escherichia coli*. J. Bacteriol. *144*:356-365.

Wood, J.M. 1981. Genetics of L-proline utilization in *Escherichia coli*. J. Bacteriol. *146*:895-901.

Wu, H.M. and D.M. Crothers. 1984. The locus of sequence-directed and protein-induced DNA bending. Nature *308*:509-513.

Wu, H.Y., and L.F. Liu. 1991. DNA looping alters local DNA conformation during transcription. J. Mol. Biol. *219*: 615-622.

Wu, J., and B. Weiss. 1991. Two divergently transcribed genes, *soxR* and *soxS*, control a superoxide response regulon of *Escherichia coli*. J. Bacteriol. *173*:2864-2871.

Wylie, D., A. Stock, C.Y. Wong, and J. Stock. 1988. Sensory transduction in bacterial chemotaxis involves phosphotransfer between Che proteins. Biochem. Biophys. Res. Commun. *151*:891-896.

Xiong, X.F., N. de la Cruz, and W.S. Reznikoff. 1991. Downstream deletion analyss of the *lac* promoter. J. Bacteriol. *173*:4570-4577.

Yamamori, T., T. Osawa, T. Tobe, K. Ito, and T. Yura. 1982. *Escherichia coli* gene (*hin*) controls transcription of heat-shock operons and cell growth at high temperatures. pp.131-137. In M. Schlessinger, M. Ashburner, and A. Tissieres (ed.), Heat shock from bacteria to man. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Yamamori, T., and T. Yura. 1980. Temperature-induced synthesis of specific proteins in *Escherichia coli*: evidence for transcription control. J. Bacteriol. *142*:843-851.

Yamamori, T., and T. Yura. 1982. Genetic control of heat-shock protein systhesis and its bearing on growth and thermal resistance in *Escherichia coli* K-12. Proc. Natl. Acad. Sci. USA *79*:860-864.

Yamato, I., and Y. Anraku. 1980. Genetic and biochemcal studies of transport systems for branched-chain amino acids in *Escerichia coli* K-12: Isolation and properties of mutants defective in leucine-repressible transport activities. J. Bacteriol. *144*:36-44.

Yanofsky, C. 1960. The tryptophan synthetase system. Bacteriol. Rev. *24*:221-245.

Yokota, T., H. Sugisaki, M. Takanami, and Y. Kaziro. 1980. The nucleotide sequence of the cloned *tufA* gene of *Escherichia coli*. Gene *12*:25-31.

Zhou, Y.N., N. Kusukawa, J.W. Erickson, C.A. Gross, and T. Yura. 1988. Isolation and characterization of *Escherichia coli* mutants that lack the heat shock sigma factor 32. J. Bacteriol. *170*:3640-3649.

Zinkel, S.S., and D.M. crothers. 1991. Catabolite activator protein-induced DNA bending in transcription initiation. J. Mol. Biol. *219*:201-215.