INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Order Number 9304675

Regulation of gene expression in *Escherichia coli ilvGMEDA* cluster

Huang, Fei, Ph.D.

City University of New York, 1992

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

A

Regulation of Gene Expression in *Escherichia coli* *ilvGMEDA* Cluster

BY

FEI HUANG

A dissertation submitted to the Graduate Faculty in Biochemistry
in partial fulfillment of the requirements for the degree of Doctor of Philosophy,
The City University of New York

1992

This manuscript has been read and accepted for the Graduate Faculty in Biochemistry in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

10 Aug 92
Date

David H. Cal
Dr. David H. Calhoun, Ph.D.
Chair of Examining Committee

12 Aug 92
Date

Horst Schulz
Dr. Horst Schulz, Ph.D.
Executive Officer

Sharon Cosloy
Dr. Sharon Cosloy, Ph.D.

Joseph S Krakow
Dr. Joseph Krakow, Ph.D.

Catherine L. Squires
Dr. Catherine Squires, Ph.D.

Mark Steinberg
Dr. Mark Steinberg, Ph.D.

Supervisory Committee

The City University of New York

Abstract

REGULATION OF GENE EXPRESSION IN *ESCHERICHIA COLI* *ilvGMEDA* CLUSTER

by

FEI HUANG

Adviser: Professor David H. Calhoun

The *ilvGMEDA* cluster encodes five gene products needed for the biosynthesis of leucine, isoleucine, and valine in *Escherichia coli* K-12. While some of the sites of transcription initiation and termination have been identified for the *ilvGMEDA* cluster, the sizes and distribution of mRNAs present in vivo have not previously been described. In this study, four transcripts have been identified. Two relatively stable transcripts of 4.6 and 1.1 kb correspond to the products of the *ilvEDA* and *ilvE* genes, and two relatively unstable transcripts of 6.7 and 3.6 kb correspond to the products of the *ilvGMEDA* and *ilvDA* genes. The identity of the transcripts was based on the use of eight probes derived from segments of the *ilvGMEDA* cluster. In addition, two strains with deletions of *ilvG* or *ilvDA* were used, and the expected decrease in transcript size was observed in Northern (RNA) blots. Primer extension using a synthetic oligonucleotide derived from the 5'-end of the *ilvD* gene and reverse transcriptase generated a 169 nt product corresponding to a 5'-end of transcript within the *ilvED* intercistronic region, 37 nt from the AUG codon of the *ilvD* gene. This primer extension product presumably indicates the 5'-end of the *ilvDA* transcript detected in Northern blots. The stability of the transcripts was monitored following rifampicin addition, and RNase E, RNase P, RNase II and PNPase, but not RNase III, were found to participate in degradation of the *ilv* transcripts. The *ilv* transcript levels varied as predicted in response to growth in the presence and absence of the end product amino acids, and in response to the presence of the polar frameshift site in *ilvG* and regulatory allele *ilvA538*. In a strain with a *rho* mutation, the level of 4.6 kb *ilvEDA* transcript increased, but the 1.1 kb *ilvE* transcript decreased. This

indicated that a Rho dependent termination site may be present in the *ilvE-ilvD* intercistronic region that causes the release of the 1.1 kb *ilvE* transcript rather than readthrough to generate the 4.6 kb *ilvEDA* transcript in wild-type (Rho$^+$) strain.

It was previously observed that an *ilv*-related protein, (ORFI), was produced at higher levels in UV irradiated cells infected with *λdilvGMEDA* phage with specific *ilvG* (Val^R) mutations, compared to phage carrying the wild type *ilvG*$^+$ (Val^S) allele. The gene coding for this protein was further localized by analyzing restriction fragment subsets in maxicells to a region between *rrnC* and *ilvGMEDA*, approximately 2.1 kb from the *ilvGMEDA* cluster. In this study the DNA sequence of the 3.5 kb segment between *rrnC* and *ilvGMEDA* was determined. Two open reading frames (ORFs) predict proteins of 18,751 (ORFI) and 20,085 (ORFII) daltons, and both ORFs have a strong probability to code for proteins based on codon frequency analysis. Maxicell analysis revealed that a *Hind*III to *Sma*I restriction fragment of 1319 bp contains ORFI and encodes the *ilv*-related peptide. A *Cla*I fragment deletion removed a portion of ORFI encoding the carboxyl-terminal region of the peptide, and revealed a decrease in the size of the protein produced in maxicell analysis in accord with the prediction. Northern blots were used to characterize transcripts encoding ORFI. A transcript initiated 112 nt from the *ilvG*p2 promoter, but proceeding in the opposite direction, may encode the ORFI peptide.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Dr. David H. Calhoun for his advice, guidance, encouragement, and continuous support throughout the course of this work. I also greatly acknowledge Dr. George Coppola for his guidance and technical advice. I would like to thank Drs. Sharon Cosloy, Joseph Krakow, Catherine Squires, and Mark Sternberg for serving on my thesis committee.

I would like to express my appreciation to my friends and colleagues, Yan Yan, Li-Bo Zhou, Janice Riley and Noren Dove for the shared experience, friendship and kind help, which have made my years at the City College enjoyable. Thanks also to Chance Neale who helped me to prepare some of figures.

Mostly, I thank my parents and relatives in my home country. It was their continuous moral support and encouragement that inspired me to achieve this goal. Finally, special thanks goes to my husband for his love, care, support, and carefully typing and editing this thesis, which made it all possible.

This thesis is dedicated to my dearest mother for everything she has done for me.

TABLE OF CONTENTS

PART TWO

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

bp	base pair
DEPC	diethylpyrocarbonate
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
IHF	integration host factor
kb	kilobase(s)
MOPS	3-(N-morpholino)-propanesulfonic acid
NaOAc	sodium acetate
nt	nucleotide
oli	oligonucleotide(s)
PCR	polymerase chain reaction
pfu	plaque forming units
PNPase	polynucleotide phosphorylase (*pnp*)
Rho	transcriptional termination protein
RNase II	exoribonuclease II (*rnb*)
RNase III	endoribonuclease III (*rnc*)
RNase E	endoribonuclease E (*rne*)
RNase P	endoribonuclease P (*rnpA*)
rRNA	ribosomal RNA
Tris-HCl	tris(hydroxymethyl)aminomethane hydrochloride
ts	temperature sensitive
Val^S	valine sensitive
Val^R	valine resistant

PART I

Multiple Transcripts and Differential Expression of the *ilvGMEDA* Cluster in *Escherichia coli* K-12

INTRODUCTION

The biosynthesis of leucine, isoleucine, and valine in *Escherichia coli* (*E. coli*) K-12 occurs in parallel pathways and involves the participation of common precursors and shared bifunctional enzymes (Fig. 1, Baez *et al.*, 1979). The genes coding for these enzymes are located at different positions on chromosome (Fig. 1). The *leu* genes and *ilvHI* are closely located at 2 min (Bachman *et al.*, 1976; Squires *et al.*, 1981); the *ilvBN* at 82 min (Friden *et al.*, 1985); and the *ilvGMEDA* and *ilvYC* at 85 min (Lawther *et al.*, 1981; Kuramitsu *et al.*, 1985; Wek and Hatfield, 1986). The biosynthesis of these amino acids is regulated by complex allosteric control of the activity of key branch point enzymes and by controlling the level of synthesis of most or all of these enzymes (Umbarger, 1971). One of the gene clusters, *ilvGMEDA,* codes for four of the enzymes required in the biosynthesis of the branched-chain amino acids (Lawther *et al.*, 1987; Cox *et al.*, 1987; Umbarger, 1987). Several mechanisms are involved in control of the expression of the *ilvGMEDA* gene cluster. The mechanisms include attenuation, internal promoters and terminators, translational coupling, binding of regulatory proteins, and possibly autoregulation by the *ilvA* gene product, threonine deaminase. Some other mechanisms, such as posttranscriptional and translation efficiency, may also be involved in controlling the expression of the *ilvGMEDA* genes.

I. Gene-Enzyme Relationship

The *ilv* genes control the synthesis of the branched and parallel pathways leading to leucine, isoleucine and valine (Fig. 1), a group of amino acids that constitute a significant fraction of the total amino acid biosynthetic effort of the cell. The pathways leading to isoleucine and valine involve the condensation of an α-keto acid with an active acetaldehyde group derived from pyruvate. In the pathway to valine, the α-keto acid acceptor is pyruvate itself, and acetolactate is formed. For isoleucine, the acceptor is α-ketobutyrate. Thus, for the valine pathway, the condensation step is the one that serves to draw off a central metabolite specifically into its pathway, and it is the first unique step in the pathway. For the isoleucine pathway, a different specific step is required since α-ketobutyrate, required for acetohydroxybutyrate formation, is not an intermediate in the central metabolic routes. Threonine is converted to α-ketobutyrate by threonine deaminase, the product of the *ilvA* gene, and this is the first step unique to the isoleucine pathway.

The enzymes forming acetolactate and acetohydroxybutyrate present a complex picture. There are three different isozymes of acetohydroxy acid synthase. The activities of two, acetohydroxy acid synthase I and III, are inhibited by valine (Val^S). The activity of the third isozyme, acetohydroxy acid synthase II, is not inhibited by valine (Val^R). The *ilvG* and *ilvM* genes encode the large subunit and small subunit of acetohydroxy acid synthase II, respectively (Favre *et al.*, 1976). A single gene cluster located at 82 min consists of the *ilvB* and *ilvN* genes encoding the large and small subunits of acetohydroxy acid synthase I (Newman and Levinthal, 1980). The *ilvI* and *ilvH* structural genes for acetohydroxy acid synthase III lie on the *ara* side and adjacent to the *leu* genes (DeFelice *et al.*, 1974).

The next step in each pathway is the isomerization and NADPH-dependent reduction of the acetohydroxy acids by the isomeroreductase, specified by the *ilvC* gene, to yield the α,β-dihydroxy acids. These intermediates, in turn, undergo dehydration reactions to yield the α-keto acid precursors of isoleucine and valine. These reactions are catalyzed by dihydroxy acid dehydratase, the product of the *ilvD* gene.

The final step in the two pathways is a transamination reaction between the two α-keto acids and glutamate. Transaminase B, the product of *ilvE* gene, is absolutely required for isoleucine biosynthesis, and it also participates in the final transamination reaction in the leucine and valine pathways.

The *ilvY* gene, located next to the *ilvGMEDA* genes, encodes a regulatory protein which activates transcription of the *ilvC* gene (Watson *et al.*, 1979). Transcription of the *ilvY* is in a direction opposite to the *ilvGMEDA* genes (Wek and Hatfield, 1986). The converging transcripts from the *ilvA* and *ilvY* genes are overlapping and terminate in the coding region of the opposing gene (Samishima *et al.*, 1989).

II. Regulatory Sites

(i) Promoters. Three promoters, *ilvGp2, ilvEp* and *ilvAp* have been well characterized, and these initiate transcription just upstream of the *ilvG, ilvE* and *ilvA* genes, respectively (Adams *et al.*, 1985; Lopes and Lawther, 1989; Pereira *et al.*, 1988; Wek and Hatfield, 1986). The presence of the two internal promoters, *ilvEp* and *ilvAp*, as well as a third internal promoter *ilvDp*, was first inferred using strains with polar insertions or mutations located within the *ilvGMEDA* cluster (Berg *et al.*, 1981; 1979; Blazey and Burns, 1979), in *λilv* phage carrying only *ilvEDA*, *ilvDA*, or *ilvA* (Gray *et al.*, 1982), and by fusion to reporter genes (Calhoun *et al.*, 1985; Lopes and Lawther, 1989; Wek and Hatfield, 1986).

(ii) Leader region and attenuator. The *ilvGMEDA* genes are multivalently regulated by an attenuator mechanism (Lawther and Hatfield, 1980; Nargang *et al.*, 1980). A leader region upstream of the *ilvG* gene yields a short transcript about 185 bases in length which could specify a 32-amino acid peptide. This putative leader peptide contains four leucine, five isoleucine and six valine residues. The 185-base transcript is potentially capable of forming several secondary structures through base pairing, including a Rho independent transcriptional terminating stem and loop. The rate of movement of a ribosome behind the RNA polymerase transcribing this leader region could determine which of the secondary structures is formed and which is excluded. The rate of ribosomal movement should, in turn, be controlled by the charging levels of the various isoleucine, leucine and valine tRNAs. If a ribosome stalls at an isoleucine, leucine or valine codon in the leader transcript, the terminator stem-loop structure could not be formed, and the RNA polymerase should read through the attenuator to the structural genes. If the ribosome does not stall at any of these codons, the terminator structure should form, and the transcript would be attenuated before reaching the structural genes. Loss of the attenuator by deletion results in loss of response of the *ilvGMEDA* coded enzymes to either excess or limiting levels of the branched-chain amino acids (Bennett and Umbarger, 1984).

(iii) Terminators. In wild type (Val^S) *E. coli* K-12 there are three tandem *rho* dependent termination sites downstream of the frameshift site (Wek and Hatfield, 1987) which reduce readthrough of *ilvGp2* initiated transcripts to *ilvEDA*. Strains with mutations in the *rho* gene produce elevated levels of the *ilvEDA* gene products and derepress to a greater extent than isogenic rho^+ strains (Baez *et al.*, 1979; Wasmuth and Umbarger, 1973). The effect of the *rho* mutation is similar to isoleucine limitation itself in that the relative changes in gene expression as determined by direct enzyme assay are greater downstream. That is, the ratio of derepressed to repressed enzyme levels is greatest for *ilvA*, intermediate for *ilvD*, and least for *ilvE*. These results indicate that Rho dependent sites of natural polarity may be present in the *ilvEDA* region (Smith *et al.*, 1976). There are two transcriptional termination sites at the distal end of the *ilvGMEDA* cluster: a Rho independent termination site, *ilvA t*, 116 base pair (bp) downstream of the *ilvA* transcriptional stop codon, and a Rho dependent termination site, *ilvA t'*, beginning 70 bp beyond the *ilvA t* site (Sameshima *et al.*, 1989). This *tt'* termination pattern is similar to the *trp* operon (Wu *et al.*, 1981), and the *rrnG* cluster (Albrechtsen *et al.*, 1991) of *E. coli* and the *leu* cluster of *Salmonella typhimurium* (Friedberg *et al.*, 1985).

III. The Cryptic *ilvG* Gene and the Val^R Growth Phenotype

A cryptic gene is defined as phenotypically silent DNA sequences not normally expressed during the life cycle of an individual but capable of activation as a rare event in a few members of a large population by mutation, recombination, insertion elements, or other genetic mechanisms (Hall *et al.*, 1983). In wild type *E. coli* K-12 the *ilvGMEDA* cluster contains a naturally occurring frameshift site near the middle of the *ilvG* gene (Lawther *et al.*, 1981; 1982). This causes the formation of an inactive *ilvG* gene product that lacks acetohydroxy acid synthase II activity. The premature translation termination within the *ilvG* gene of the wild type strain generates transcriptional polarity which decreases the expression of the distal *ilvEDA* genes. When the *ilvG* pseudogene carried by the wild type *E. coli* K-12 cells is activated by spontaneous mutations (one-base-pair deletions or two-base-pair insertions) which correct the translational alignment of the *ilvG* gene and remove the frameshift site, the *ilvG* gene is expressed, isozyme II (Val^R) is produced and the cell is phenotypically valine resistant.

IV. Multivalent Control of *ilv* Regulon

The branched biosynthetic pathway contains common precursors and shared bifunctional enzymes. The branched-chain amino acids and some of their precursors are structurally related and growth inhibition is seen with aberrant ratios of these amino acids or precursors. The levels of all three end products and some key precursors, such as α-ketobutyrate (Primerano and Burns, 1982; Daniel *et al.*, 1983), must be closely coordinated, and this is achieved by allosteric control of the activity of key branch point enzymes and by controlling the level of synthesis of most or all of these enzymes. Sites of allosteric control are present at the beginning of the isoleucine, valine, and leucine branchlets, and both primary (i.e. the end product) and remote (other than the end product) effectors (Nargang *et al.*, 1980) control metabolite flow (Fig. 2, Hahn and Calhoun, 1978).

Repression of the *ilvGMEDA* genes (Fig. 3, Umbarger, 1987) requires all three branched-chain amino acids, and the limitation of any one of them leads to derepression (Freundlich *et al.*, 1962). The acetohydroxy acid synthase I is derepressed when either valine or leucine is limiting; it is not derepressed if isoleucine is limiting. Acetohydroxy acid synthase III formation is controlled by the level of leucine alone. Thus when valine is added at a high concentration (e.g., 1 mM) to a minimal medium culture of the Val^S K-12 strain of *E.coli* (which contains only the two valine-sensitive isozymes of acetohydroxy acid synthase, AHSI and III), isoleucine is no longer formed and becomes increasingly more limiting. Since valine can be converted through α-ketoisovalerate to leucine, both valine and leucine become relatively more in excess. This process leads to a progressively tighter repression of both acetohydroxy acid synthase I and III. Acetohydroxy acid isomeroreductase, the product of the *ilvC* gene, is indirectly regulated by valine, because it is induced by either of its substrates, α-acetolactate or α-acetohydroxybutyrate in a process mediated by an activator protein encoded by the *ilvY* gene.

The substrates for the isomeroreductase are formed by their acetohydroxy acid syntheases (Ratzkin *et al.*, 1972). Since the wild type (Val^S) *E. coli* K-12 has only the valine-sensitive isozymes, the inducer will be formed only when valine is limiting. In the Val^R strain, the isomeroreductase is induced when either leucine, valine, or isoleucine are limiting. This results from increased inducer levels as metabolite flow to the end product amino acids increases.

V. Differential Expression of the *ilvGMEDA* Genes

The expression of the *ilvGMEDA* cluster is not completely coordinate, and the differential expression is particularly noted with threonine deaminase, the product of the *ilvA* gene, which is derepressed more than the other products of the cluster upon limitation by the branched-chain amino acids (Smith *et al.*, 1976), as indicated in Fig. 3. The differential derepression is the greatest with isoleucine limitation and least with leucine limitation. Thus, in effect, there is a mechanism that allows threonine deaminase, which is needed only for isoleucine biosynthesis, to be formed at the highest rate when isoleucine is limiting. The basis for the differential expression has not been established. Three regulatory features could result in this discoordinate expression of the cluster.

(i) Expression from internal promoters. The direct examination of the DNA sequence (Lawther *et al.*, 1981) and RNA polymerase binding to restriction fragments (Subramanyam *et al.*, 1980) indicate the presence of a promoter, *ilvEp,* in the *ilvM* coding region upstream of *ilvE* gene. The *ilvE-lacZ* fusions produce β-galactosidase at about 60% the levels of *ilvG-lacZ* fusions in minimal medium under repressing conditions (Calhoun *et al.*, 1985). Thus, in the wild-type *E. coli*, this internal promoter *ilvEp* accounts for more than half of *ilvEDA* transcription, at least under repressing conditions. Some studies indicate that *ilvEp* might be regulated (Lopes and Lawther, 1989; Subrahmanyam *et al.*, 1980), and others indicate that it is constitutively expressed (Berg *et al.*, 1979; Wek and Hatfield, 1986). A more recent and very thorough analysis indicates that *ilvEp* responds neither to the availability of the end product amino acids nor to a wild range of growth rates (Harms and Umbarger, 1991).

Another internal promoter, designated *ilvAp,* was identified which allows independent expression of *ilvA* gene under anaerobic conditions in a cAMP dependent fashion (Lopes and Lawther, 1989). This *ilvAp* promoter has been analyzed both in vivo and in vitro by fusions to *galK* gene, and it is a weaker promoter than *ilvGp2*.

It was observed that insertion of Tn5 into *ilvE* gene was only partially polar on expression of the downstream genes, implying the existence of a promoter designated *ilvDp,* which expresses *ilvDA* genes (Berg *et al.*, 1979). However, no further evidence for the presence of *ilvDp* has been reported. Thus, the significance of the earlier genetic observations for *ilvDp* is currently unclear.

These three internal promoters, *ilvEp, ilvDp* and *ilvAp* could permit expression of three genetic subsets of the *ilvGMEDA* gene cluster: *ilvEDA, ilvDA* and *ilvA,* respectively.

(ii) Existence of transcriptional and/or translational polarity sites. The *ilvG* gene in the wild-type *E. coli* K-12 contains a frameshift site which results in termination of translation near the middle of the gene. Three tandem Rho dependent sites have been located in the *ilvGM* region downstream of the frameshift site in the *ilvG* gene (Wek *et al.*, 1987). These transcriptional polarity sites are responsible for terminating transcription of the genes upstream of *ilvEp*.

(iii) Differential mRNA stability. In addition to the internal promoters and polarity sites, the DNA sequence of the *ilvGMEDA* gene cluster indicates the presence of regions that may form RNA secondary structure (Lawther *et al.*, 1987; Cox *et al.*, 1987). An inverted repeat presented in the intercistronic between *ilvE* and *ilvD* and another one near the *ilvD-ilvA* boundary potentially could form particularly stable stem and loop structures. However, the role of such structures on the *ilvGMEDA* gene expression remains unestablished. The stem and loop structure could possibly act as a barrier to 3' to 5' specific exonucleases to affect mRNA stability.

VI. Role for Threonine Deaminase in Regulation of *ilv* Gene Expression

Evidence was presented that the first enzyme in a pathway can play a role in the repression of enzymes in that pathway (Kovach *et al.*, 1969). Direct in vivo evidence that the *ilvA* gene product, threonine deaminase, affects *ilvGMEDA* and *ilvC* expression comes from an analysis which demonstrated that mutated strains unable to produce the holoenzyme form of threonine deaminase (containing the pyridoxal phosphate cofactor) are unable to repress the *ilvEDA* genes products and are defective in the *ilvC* induction (Hatfield and Burns, 1970; Calhoun and Hatfield, 1973; Levinthal *et al.*, 1973). The most thoroughly studied mutation is *ilvA538*, in which a C-to-T point mutation changes leucine-481 to phenylanine (Taillon *et al.*, 1988). The *ilvA538* mutation also alters the sensitivity of the enzyme to allosteric inhibition (Calhoun, 1976) and the regulation of the *ilvGMEDA* expression (Levinthal *et al.*, 1973). It is also accompanied by greatly reduced expression of the genes for isoleucyl and valyl tRNA synthesis (Singer *et al.*, 1984). The C-terminal segment of threonine deaminase presumably is involved in either (i) the binding or interaction of the allosteric effector isoleucine with the enzyme, or (ii) subunit interactions (Taillon *et al.*, 1988). From considerable evidence it was suggested that threonine deaminase might act as a negative control element for the *ilvGMEDA* (Hatfield and Burns, 1970) and a positive control element for the *ilvC* (Calhoun and Hatfield, 1973). Although there is little experimental evidence, a model was proposed (Singer *et al.*, 1984) in which isoleucyl and valyl tRNA synthetases are inhibited by virtue of a complex between synthetase, threonine deaminase, and α-ketobutyrate.

VII. Mutations Affecting *ilvGMEDA* Expression

Most mutations causing increased expression of the *ilvGMEDA* gene cluster in *E. coli* might be explained on the basis of interference with attenuation control. These include mutations affecting the amount of branched-chain aminoacyl-tRNA synthetase activity (Eidlic and Neidhardt, 1965; Blatt and Umbarger, 1972) and tRNA modification enzymes (Cortese *et al.*, 1974; Harris *et al.*, 1983).

However , other mutations leading to enhanced expression of the *ilvGMEDA* gene cluster appear to be independent of attenuation. Mutations in the *rho* gene produce elevated levels of the *ilvEDA* gene products (Baez *et al.*, 1979) and generate downstream amplification, indicating that Rho dependent sites of natural polarity may be present in the *ilvGMEDA* region (Smith *et al.*, 1976). Mutation in *relA* strongly impairs the derepression of the *ilvA* gene (Freundlich, 1977). In vitro studies indicate that ppGpp stimulates the *ilvGMEDA* specific mRNA expression (Smolin and Umbarger, 1975).

It was reported that two mutations, *avr-16* and *flr-9*, lead to constitutive expression of the *thr*, *leu* and *ilv* genes, and this constitutivity can be significantly reversed by increasing the intracellular levels of the *trp* aporepressor protein (Johnson and Somerville, 1983). Sequences with homology to the *trp* operator have been found upstream from the transcription start points of the *thr*, *leu* and *ilv* genes. These observations were interpreted to indicate the existence of specific repressor proteins for the *thr, leu* and *ilv* genes which are functionally analogous to the repression known for the *trp* system (Bogosian and Somerville, 1983; Weiss *et al.*, 1986). Deletion of a gene designated *ilvR* results in an Ilv$^-$ growth phenotype and causes an undetectable level of β-galactosidase expression from a λ prophage with a fusion of the upstream *ilvGMEDA* promoter, *ilvGp2*, to *lacZ*. These investigators speculated that *ilvR* coded an activator which could bind upstream of *ilvGp* (Johnson and Somerville, 1984).

Integration host factor (IHF) is a sequence specific DNA-binding protein (Craig and Nash, 1984). IHF mutants (*himA* and *himD*) have been shown to affect several functions of controlling expression in *ilv* genes, but were originally identified as mutations preventing λ-mediated, site-specific recombination (Friedman *et al.*, 1984; Kikuchi and Nash, 1978; Miller *et al.*, 1979). IHF specifically binds to a site just upstream of the *ilvBN* promoter, and alters the conformation of *ilvBN* promoter-leader DNA, and strongly decreases transcriptional pausing and termination in the *ilvBN* leader (Tsui and Freundlich, 1990). In the presence of *himA* or *himD* mutations, transcription of the *ilvBN* cluster is strongly inhibited, and the expression of the *ilvGMEDA* cluster is also reduced (Frieden *et al.*, 1984). Direct footprint analysis reveals two binding sites for IHF in the region of *ilvGMEDA* promoters (Tsui and Freundlich, 1988; Winkelman and Hatfield, 1990). IHF binding to this region inhibits transcription from the *ilvGp*1 promoter and activates expression from the *ilvGp*2 promoter (Pagel and Hatfield, 1991; Pereira *et al.*, 1988). IHF also interacts with the attenuator site in *ilvGMEDA*, and increases the termination of the transcription at the downstream of the attenuator (Pagel and Hatfield, 1991).

VIII. Relationship of ORFI to *ilvGMEDA*

A protein designated ORFI is produced at remarkably elevated levels in Val^R strains (that have mutations in *ilvG*) compared to Val^S strains (the wild type *E. coli* K-12 with the *ilvG* pseudogene) (Gray *et al.*, 1981 and 1982). The location of the coding region for ORFI has been identified by systematic analysis to a region about 2.1 kilobases (kb) counterclockwise to the *ilvGMEDA*, close to *rrnC*. The DNA sequence of this region has been determined, and the promoter ORF*p* may direct the expression of ORFI protein (Coppola *et al.*, 1991). ORF*p* and *ilvGp1* promoters are overlapping and divergent, and one IHF DNA binding site overlaps both of these promoters. This may provide a clue that a potential regulatory relationship exists between ORFI and the *ilvGMEDA*. These studies are described in more detail in Part II (Page 38).

MATERIALS AND METHODS

Materials

Restriction endonucleases, T_4 DNA polymerase, T_4 DNA ligase, T_4 polynucleotide kinase were purchased from New England Biolabs. The Klenow fragment of DNA polymerase was purchased from International Biotechnologies Inc. RNasin and RQ1 RNase-free DNaseI were purchased from Promega Corp. AMV reverse transcriptase was purchased from Gibco BRL. Calf intestine phosphatase, S1 nuclease and RNA molecular weight markers were purchased from Boehringer Mannheim. Taq DNA Polymerase was purchased from Perkin-Elmer Cetus Corp. Sequenase Version 2.0 kit was from United States Biochemical Corp. (α^{35}S)dATP, (α^{32}P)dCTP and (γ^{32}P)dATP were obtained from New England Nuclear. Lysozyme, diethylpyrocarbonate (DEPC), formamide, formaldehyde, 3-(N-morpholino)-propanesulfonic acid (MOPS) and rifampicin were from Sigma.

Bacterial Strains and Plasmids

The bacterial strains used in this study are all derivatives of *E. coli* K-12. The relevant genotypes of these bacterial strains are described in Table I. The characteristics of plasmids constructed and utilized in this study are listed in Table II. The sequence and the position of synthetic oligonucleotides are described in Table III and in Fig. 4.

Construction of Plasmids

Standard methodologies (Maniatis *et al.*, 1989) were used to construct following plasmids:

(i) pCC17. Plasmid pJG51, which has unique sites for *Nsi*I and *Kpn*I separated by 1,110 bp in the *ilvG* gene, was treated with both enzymes, followed by T_4 DNA polymerase to remove the protruding 3' ends, and T_4 DNA ligase to join the blunt ends. The deletion of 1,110 bp in the resulting plasmid, pCC17, generated a protein fusion within the *ilvG* gene (Fig. 5), so that translation proceeds without the formation of a polar frameshift site.

(ii) pCC20. Plasmid pJG51 (Fig. 6) was treated with *Xho*I, which cleaves in the *ilvD* gene, and *Sac*II, which cleaves in the *ilvA* gene. Phage T_4 DNA polymerase was used to remove two overhanging bases at 3' end of the *Sac*II site, Klenow fragment was used to fill in four bases at the *Xho*I site, and then T_4 DNA ligase was used to join the blunt ends. The deletion of 1290 bp generates a protein fusion with the amino terminal region coded by the *ilvD* and the carboxyl terminal region coded by the *ilvA*.

(iii) pCC40. Plasmid pCC40 was constructed by exchanging the 1,100 bp of *Nsi*I-*Kpn*I fragment in the *ilvG* gene of plasmid pJG51 (Val^S) with the same fragment from plasmid pJG16 (Val^R). Plasmid pJG16 was treated with both *Nsi*I and *Kpn*I, 1,100 bp fragment was isolated by ultracentrifugation using the procedure of Ericson (1990). Plasmid pJG51 was digested with *Nsi*I and *Kpn*I, and treated with calf intestine phosphatase to remove 3'-phosphate group, then ligated by T_4 DNA ligase with purified *Nsi*I-*Kpn*I DNA fragment from plasmid pJG16. The ligation mixture was used to transform FD1054 (Δ*ilvGMEDA*::Tn5-131). The transformants were selected on M9 (Miller, 1972) minimal plates containing valine (0.3 mM) and ampicillin (50 µg/ml) to select the Val^R phenotype. A plasmid designated pCC40 was obtained from a Val^R transformant.

(iv) pCC36. Phage P1 grown on strain CU2501 [*rbs-200, ilvG468* (Val^R)] was used to transduce strain CC84 [strain FD1054 (Δ*ilvGMEDA*::Tn5-131) containing plasmid pJG51 (Val^S)] to Val^R. A plasmid designated pCC36 obtained from a Val^R transductant was used to transform strain FD1054 to Val^R, generating strain CC221. Strains CC84 (Val^S) and CC221 (Val^R) are Ilv prototrophs since the plasmid contains the *ilvGMEDA* genes and the chromosome contains the *ilvYC* genes.

Preparation of Competent Cells and Transformation

Competent cells were made using the one-step protocol as described by Chung *et al.* (1989). A fresh overnight culture of bacteria was diluted 1:100 into LB medium (10 g tryptone, 5 g yeast extract, 5 g NaCl per liter pH 7.4) and incubated at 37°C until the cells reach an

OD_{600} of 0.3 to 0.4. The cells were pelleted by centrifugation 10 minutes at 1000 X g at 4°C and resuspended at one tenth of original volume in 1 X TSS [10% (wt/vol) polyethylene glycol 4000, 1 X LB, 50 mM $MgCl_2$, 5% (vol/vol) dimethyl sulfoxide, pH6.5]. Small aliquots (1 ml) of the suspension were then frozen in a dry ice/ethanol bath and stored at -70°C.

Transformation (Chung *et al.*, 1989) was carried out by adding 100 µl competent cells and 1 to 5 µl DNA (intact plasmid DNA or ligation mixture) to an ice cold polypropylene tube, swirling to mix, followed by incubation on ice for 30 minutes. 900 µl LB medium containing 20 mM glucose was added followed by incubation at 37°C with mild shaking for 60 minutes to allow expression of the antibiotic resistance gene. An aliquot of transformed cells was spread on YT agar plates containing appropriate antibiotics (ampicillin at 50 µg/ml or tetracycline at 15 µg/ml) to select transformants.

Transduction

Phage P1 lysates were prepared by confluent lysis method (Rosner, 1972). To 3 ml of a standing overnight, room-temperature culture of strain CU2501 (Val^R), 1.5 to 3.0 X 10^6 phage previously grown on strain GT103 were added and allowed to adsorb for 30 minutes at 37°C in the presence of 5 mM $CaCl_2$. Then 10 ml of LB Mg (10 mM $MgCl_2$ with LB) and 10 ml of LB top agar were added, equal volumes of this mixture were poured onto each of three LBCa plates (LB with 1.5% agar and 5 mM $CaCl_2$). The plates were incubated at 37°C until confluent lysis was observed (6—8 hours). The top agar layer was removed, triturated with a glass rod, and treated with 4 drops of $CHCl_3$ for 20 minutes at 32°C. The treated lysate was then centrifuged for 5 minutes at 9,000 rpm in IEC centrifuge with #870 rotor. The supernatant fluid was collected and assayed for phage P1 activity by determining the titer of plaque forming units (pfu). Three successive lysates were prepared to minimize carryover of phage P1 grown on strain GT103.

The recipient strain CC84 [strain FD1054 (*ΔilvGMEDA*::Tn5-131) containing plasmid pJG51 (Val^S)] was grown in LBMg at 32°C to a concentration of 2—3 X 10^8 pfu/ml. The culture was iced and $CaCl_2$ was added to a final concentration of 10 mM. Aliquots of 2.5 ml were added to centrifuge tubes containing 2.5 ml LBMg (control) or P1·CU2501 lysate diluted in LBMg to a concentration of 4 X 10^8 pfu/ml. After 30 minutes at 32°C the tubes were iced and centrifuged at 9,000 rpm for 5 minutes in IEC centrifuge with a #870 rotor. The pellet was rinsed with 1 ml of iced phosphate-buffered saline (PBS; 0.4% NaCl, 0.3% KH_2PO_4 and 0.7% Na_2HPO_4, pH 6.8) then resuspended in 1 ml iced PBS. The aliquots were plated on M9 minimal plates containing valine (0.3 mM) and ampicillin (50 µg/ml). The plates were kept at room temperature for 2 hours to minimize P1 growth and then incubated for 3 days at 32°C to select valine resistance transductants.

Bacterial Growth Conditions

Bacterial strains with or without plasmids were grown on a rotary shaker at 200 rpm in M9 medium supplemented with required nutrient to certain stage as indicated and used for RNA isolation (Figs. 7-17).

For identification of the *ilvGMEDA* specific transcripts (Figs. 7-9), strain C600 derivatives containing plasmids pJG51, pCC17 and pCC20 (Strains CC85, CC86 and CC89, respectively) were grown at 37°C in M9 medium supplemented with 0.3 mM threonine, leucine, isoleucine and valine, and 0.05 mM thiamine to mid-log stage, and used for RNA isolation.

For the amino acid limitation experiments (Figs. 15, 16), cells were grown on a rotary shaker at 200 rpm and 37°C to early exponential phase (A_{550}=0.3) in the presence of excess leucine, isoleucine and valine, pelleted by centrifugation, washed with M9 medium, resuspended in M9 medium containing limiting concentrations of one amino acid (0.015 mM for leucine and isoleucine, 0.03 mM for valine) as described (Driver and Lawther, 1985), and incubated on a rotary shaker at 200 rpm and 37°C for 3 hours prior to RNA isolation.

In order to determine if the endonucleases RNase III, RNaseE, RNaseP (encoded by the *rnc*, *rne*, and *rnpA* genes, respectively) play a role in the *ilvGMEDA* transcripts processing or stability, isogenic wild type and mutant pairs were used. Strain N3431 (*rne-3701*) and the isogenic *rne*$^+$ strain N3433 carrying plasmid pJG51 (strains CC147 and CC148, respectively) were grown in M9 medium with 0.05 M thiamine. Strain NHY322 (*rnpA49*) and isogenic *rnpA*$^+$ strain NHY312 harboring plasmid pJG51 (strains CC287 and CC283, respectively) were grown in M9 medium with thiamine and proline. Strains CC147, CC148, CC283 and CC287 were grown at 30°C to ΔKlett reading of 70, and then shifted to 43°C for 10 minutes and rifampicin was added as described (Mudd *et al.*, 1990; Nilsson and Uhlin, 1991). RNase III deficient strain BL321 (*rnc-105*) and the isogenic *rnc*$^+$ strain BL322 (*rnc*$^+$) containing plasmid pJG51 (strains CC325 and CC326, respectively) were grown at 37°C to ΔKlett 70 prior to rifampicin addition. Rifampicin was added at final concentration of 200 μg/ml and RNA samples were prepared at 0, 5, 10, and 15 min, as indicate (Figs. 11-13).

For determining the effect of exoribonucleases RNase II and polynucleotide phosphorylase (PNPase), the isogenic strains SK5003 (*pnp-7, rnb-500*^{ts}), SK5004 (*pnp-7, rnb*$^+$) and SK5005 (*pnp*$^+$, *rnb-500*^{ts}), and wild type strain SK5006 (*pnp*$^+$, *rnb*$^+$) were transformed with plasmid pJG51. The resulting transformants (strains CC327, CC328, CC329, and CC330, respectively) were grown at 30°C in M9 medium with leucine and threonine to ΔKlett 70, then shifted to nonpermissive temperature (44°C) for 10 minutes, rifampicin was added, and samples were taken at the indicated times (Fig. 14).

Synthesis and Purification of Oligonucleotides

Oligonucleotides (Table III) were synthesized on a Cyclone™ plus DNA synthesizer (Milligen/Biosearch) using phosphoramidite chemistry as described by the manufacturer. When synthesis was completed, the oligonucleotides were separated from the protecting group and purified using either Olig-Pak columns (Milligen/Biosearch) according to the manufacturer's procedure, or by electrophoresis on a 20% polyacrylamide preparative gel containing 8M urea, visuablization by UV shadowing, and elution from the gel using the crush and soak procedure of Maxam and Gilbert (1980).

Labeling of Oligonucleotides

To label the oligonucleotide probes used for Northern blots with (^{32}P) using T_4 polynucleotide kinase, 30 pmoles of oligonucleotide was added to 30 µl of DEPC-treated H_2O, 5 µl of 10X kinase buffer [0.5 M tris(hydroxymethyl) aminomethane hydrochloride (Tris-HCl), pH 7.6, 100 mM $MgCl_2$, 50 mM dithiothreitol, 1 mM spermidine)], 10 µl of (γ-^{32}P) dATP (New England Nuclear, specific activity 3000 Ci/mmol) and 1 µl of T_4 polynucleotide kinase (10 units/ µl) and incubated at 37°C for 30 minutes. The reaction was stopped with the addition of 50 µl of STE [0.1 M NaCl, 10 mM Tris-HCl pH7.6, 1 mM ethylenediaminetetraacetic acid (EDTA)]. The labeled oligonucleotides were separated from unreacted ATP by spin-column chromatography using Sephadex G-50 as described by Maniatis *et al.* (1989).

Polymerase Chain Reaction

The asymmetric PCR procedure was performed as described by McCabe (1990) to produce single-stranded DNA products either as templates for directly sequencing the constructed plasmids to confirm the expected deletions, or as probes used in the Northern blots. The reaction contains 50 µM dNTP, 50 pmol excess primer (antisense primer, i.e., primers 20, 6, 7, 18, D5, 47, and 38, which anneal to mRNA strand and generate the probe used for hybridization), 1 pmol limiting primer (sense primer, i.e., primers 19, 40, 41, and 17), 2.5 units Taq DNA polymerase (Perkin-Elmer Cetus), 25 mM Tris-HCl (pH 8.5), 1.5 mM $MgCl_2$, 50 mM KCl, 0.1% Tween 20, and 0.1 ng plasmid DNA template. An automated thermocycler (Perkin-Elmer Cetus) was programmed for 40 cycles with the following cycle profile: 95°C for 30 seconds, 60°C for 30 seconds, 72°C for 2 minutes. After the cycling finished, the reaction was incubated at 72°C for an addition 15 minutes to ensure completion of final extensions. For instance, to generate pcrDA probe, 0.1 ng plasmid pJG51 DNA, 50 pmol (^{32}P)-labeled primer 18, and 1 pmol primer 17 were used for PCR. The reaction was monitored for production of single-stranded DNA by agarose gel electrophoresis. Special tips (capillaries and pistons) and pipettors from Gilson were used for PCR work to minimize contamination.

Preparation of RNA

RNA was isolated as described by Ausubel *et al.* (1991) with one modification (DNase I treatment). Cells (50 ml) were collected by centrifuging 10 min at 4°C at 10,000 rpm in a IEC centrifuge with #870 rotor in two of 30 ml glass tubes (Coret, cat. #156; DEPC treated and autoclaved). The cell pellets were resuspended in 10 ml protoplasting buffer (15 mM Tris-HCl pH 8.0, 0.45 M sucrose, 8 mM EDTA). Lysozyme (Sigma; 80 µl at 50 mg/ml) was added. The mixture was incubated for 15 min on ice, and centrifuged at 4°C at 7,000 rpm for 5 min. The protoplasts were resuspended in 0.5 ml lysing buffer (10 mM Tris-HCl pH 8.0, 10 mM NaCl, 1 mM Na-citrate, 1.5% SDS), and 15 µl DEPC was added. The mixture was transferred to a microcentrifuge tube (autoclaved, RNase free, Jersey Lab & Glove Supply, cat. #T-15-NGS-16), incubated 5 min at 37°C and chilled on ice, and 250 µl saturated NaCl (40%) was added. The mixture was inverted and incubated for 10 min on ice, centrifuged at 4°C for 15 min, extracted twice at room temperature with a phenol-chloroform (25:24) solution purchased from Jersey Lab & Glove Supply (cat. #0883, ultra pure grade). After ethanol precipitation the RNA was resuspended in 100 µl of DNase digestion buffer (80 mM Tris-HCl pH 7.9, 10 mM EDTA, 10 mM $MgCl_2$, 10 mM 2-mercaptoethanol), 10 µl of RQ1 RNase-free DNase(1 unit/µl; Promega) was added, and incubated 60 min at 37°C. An equal volume of phenol-chloroform (25:24) was added and extracted twice. The RNA was precipitated by adding 1/10 vol of 3 M NaOAc, pH 5.2, and 2 vol of ethanol at -70°C for 30 min, and collected by centrifugation at 4°C for 10 min. The pellet was washed with 70% ethanol and resuspended in 30 µl of DEPC-treated H_2O, and stored at -70°C. The concentration of RNA was determined by adding 4 µl to 996 µl of water and measuring absorbance at 260 nm using a LKB spectrophotometer, assuming 1 OD = 40 µg/ml.

The difficulty in RNA isolation is that most ribonucleases are very stable and active enzymes that require no cofactors to function. To maintain RNase free conditions, the following precautions were taken. (i) Separate chemicals and electrophoresis apparatus were used only for

RNA work. (ii) Water and solutions were treated with DEPC which inactivates ribonucleases by covalent modification. DEPC was added to solutions to be treated at 0.1%, and solutions were shaken vigorously to get the DEPC into solutions, put in a 37°C incubator overnight, and autoclaved to inactivate the remaining DEPC. Solutions containing Tris-HCl cannot be effectively treated with DEPC because Tris reacts with DEPC to inactivate it. Therefore DEPC-treated water was used to prepare Tris containing solutions. (iii) Glassware was soaked in 0.1% DEPC overnight and rinsed several times with sterile water, and then autoclaved. (iv) Disposable gloves were always worn during RNA isolation and analysis. (iv) Reagents were tested for RNase contamination by adding 5 µl of tRNA (5 µg) to 5 µl of reagent to be tested, and analyzed on a 6.6% formaldehyde denaturing 1.0% agarose gel. The tRNA samples were visualized by ethidium bromide staining.

Northern Hybridization and Probes

In a typical experiment RNAs (15 µg) were denatured at 55°C for 15 min in 1 X MOPS, 50% formamide, 6.6% formaldehyde and analyzed on 1.0% agarose gels containing 6.6% formaldehyde, and transferred overnight by capillary action in 10 X SSC to nylon filters (Nytran, 0.45 mm, Schleicher & Schuell, Inc.). The filters were placed in sealable plastic bags (Kapak/Scotchpak #402) with aqueous hybridization solution (25 mM KPO_4 pH7.4, 5 X SSC, 5 X Denhardt's solution, 10 mM EDTA, 0.5% SDS, 100 µg/ml salmon testes DNA) and prehybridized for at least 4 hours at the desired temperature. The prehybridization solution was then removed, the hybridization solution with the appropriate probe labeled with (^{32}P) radioactivity was added to the bag. The filters were incubated overnight at the same temperature as prehybridization, and washed according to the procedure of Gilman (1987).

The probes (Fig. 4B) used for the Northern blots were either synthetic oligonucleotides derived from sequences encoding the *ilvE* gene product (7: nt 3113 to 3091; oliE), the *ilvD* gene product (47: nt 4293 to 4244; oliD) or the *ilvA* gene product (38: nt 5842 to 5793; oliA), or PCR products derived from primers in *ilvG*, (19: nt 283 to 302, and 20: nt 1584 to 1565; pcrG), in *ilvE* (41: nt 2766 to 2785, and 7: nt 3113 to 3091; pcrE), from the *ilvGM* junction (40: nt 1596 to 1615, and 6: nt 2153 to 2131; pcrGM), the *ilvED* junction (41: nt 2766 to 2785, and D5: nt 3323 to 3301; pcrED), and the *ilvDA* junction (17: nt 4856 to 4876, and 18: nt 6363 to 6341; pcrDA). Oligonucleotides 19, 40, 41, and 17 anneal to the coding DNA strand and primer from left to right in Fig. 4B, and the oligonucleotides 20, 6, 7, 18, D5, 47, and 38 anneal to the opposite strand (mRNA strand). The asymmetric PCR protocol of McCabe (1990) was used to preferentially generate a product complementary to the mRNA, and the anti-sense primer was labeled with (^{32}P) using T$_4$ polynucleotide kinase prior to amplification.

For the Northern blots (Fig. 7-9, 11-17), ethidium bromide stained gels confirmed that comparable quantities of RNA were loaded in each lane, and differences in plasmid copy number were not detected in the plasmids used in those experiments. Each experiment was repeated at least twice, and consistent results were obtained for each replicate.

Primer Extension

Primer D5 (5'gggtgaacgagttcacaaccgcg3'), which was derived from sequence near the 5'-end of the *ilvD* gene, was labeled at 5'-end with (γ-^{32}P) using T4 polynucleotide kinase, and incubated with 50 µg of RNA isolated from FD1054 with or without plasmid pJG51 in a reaction volume of 20 µl of hybridization buffer (250 mM KCl, 5 mM Tris-HCl, pH 8.0, and 0.5 mM EDTA) at 80°C for 10 min, then at 42°C for 2 hr. RNasin (0.5 µl containing 20 units), reverse transcriptase (2 µl containing 400 units), and 130 µl of reaction buffer provided by the supplier of reverse transcriptase (250 mM Tris-HCl, pH 8.3, 375 mM KCl, 50 mM dithiothreitol, and 15 mM $MgCl_2$, 1 mM dNTPs) were then added and incubated at 37°C for 60 min. The reaction was terminated by the addition of 4 µl of 0.5 M EDTA, 15 µl of 3 M NaOAc pH 5.2, and 400 µl of ethanol, precipitated at -70°C for 30 min. The pellet was collected by centrifugation at 4°C for 10 min, washed with 70% ethanol, and resuspended in 10 µl of TE solution, 10 µl of formamide loading buffer. Half volume of samples were loaded onto an 8% polyacrylamide sequencing gel.

Nucleotide Sequence Analysis

In order to confirm the nucleotide sequence flanking the deletions introduced in the generation of plasmids pCC17 (Fig. 5) and pCC20 (Fig. 6), the enzymatic sequencing protocol (Sanger *et al.*, 1977) was used. The single-stranded DNA templates were generated by asymmetric PCR procedure (McCabe, 1990). The sense primer was used for sequencing reaction. The junction regions of both deletions in plasmids pCC17 and pCC20 were sequenced and confirmed the presence of the expected deletions.

RESULTS

Characterization of the *ilvGMEDA* Transcripts

Northern blot hybridization was used to determine the identities of the mRNA species coded in vivo by the *ilvGMEDA* cluster. As probes (Fig. 4B) synthetic oligonucleotides derived from the *ilvE, ilvD* and *ilvA* genes, and asymmetric PCR products derived from the *ilvG* and *ilvE* genes and from the *ilvGM, ilvED* and *ilvDA* junctions were used. It is noted, as has been previously reported by others (Murakawa et al., 1991), that the rRNAs (1,541 and 2,904 bases) in the samples saturate the binding sites on the membrane at their positions and create "shadows" in the *ilv* mRNA pattern. The *ilv* mRNAs accumulated just ahead of the rRNAs may represent mRNA "trapped" by the large amount of rRNA, thus providing internal controls for the size estimation, which was based on RNA size markers in parallel lanes.

In Northern blots, four transcripts of 6.7, 4.6, 3.6 and 1.1 kb are identified, which correspond respectively to the products of the *ilvGMEDA*, *ilvEDA*, *ilvDA* and *ilvE* genes. The 4.6 and 1.1 kb transcripts are relatively stable; the 6.7 and 3.6 kb transcripts are much less stable, and are not always detectable at lower levels of expression or on shorter exposures of the blots to film.

The 6.7 kb transcript hybridizes to all eight probes (Figs. 7-9, 11-17, and data not shown), the asymmetric PCR products (pcrG, pcrGM, pcrE, pcrED and pcrDA) and the *ilvE, ilvD* and *ilvA* specific oligonucleotides (oliE, oliD, and oliA). The 6.7 kb transcript is the size predicted for a full length *ilvGMEDA* transcript initiated at the *ilvGp* promoter at nt 1 and terminating at nt 6703 at the end of the cluster.

The 4.6 kb transcript hybridizes to all probes except pcrG (Figs. 7-9, 11-17, and data not shown) and is the size predicted if initiation occurred at the *ilvEp* promoter at nt 2101 and continued to the terminator at bp 6703. It should be noted that the probe pcrGM extends from nt 1596 to 2131, so that it overlaps the site of transcription initiation of the *ilvEp* at nt 2101 by 30 nt, which is of sufficient size to account for the observed hybridization to the 4.6 kb transcript.

The 3.6 kb transcript hybridizes to probes pcrED, oliD, pcrDA, and oliA (Figs. 8, 9, 15 and data not shown), but not pcrG, pcrGM, pcrE, or oliE (Figs. 7, 11-17, and data not shown), and is the size predicted for a transcript encoding the *ilvD* and *ilvA* genes. The primer extension experiment described below (Fig. 10) is consistent with a transcript extending from nt 3153, in the intercistronic region between the *ilvE* and *ilvD*, to the termination site at nt 6703.

The 1.1 kb transcript hybridizes to probes pcrGM, pcrE, oliE and pcrED (Figs. 7, 11-17, and data not shown), but not pcrG, pcrDA, oliD, or oliA (Figs. 8, 9, 15 and data not shown). Thus, this transcript contains internal coding regions, but lacks the 5'- and 3'-ends of the *ilvGMEDA* cluster. This transcript could be derived from longer precursor, such as the 6.7 kb or 4.6 kb transcripts, or it could be a primary transcription product. In the latter case, it could be derived from initiation at the *ilvEp*, which predicts a transcript extending from nt 2101 to approximately nt 3201 encoding the *ilvE* gene, but little or none of the *ilvD* gene.

Thus, four major transcripts are detected , and the 4.6, and 1.1 kb messages form stronger, more distinct bands than the 6.7 and 3.6 kb messages. Similar variations in intensity and sharpness are also seen with transcripts derived from the *lac* operon (Murakawa *et al.*, 1991).

Effects of Specific Deletions in the *ilvGMEDA* cluster

In order to confirm the initial assignments of transcripts based on hybridization to specific probes, two derivatives of plasmid pJG51 were constructed with specific deletions (Figs. 5 and 6). Plasmid pCC17 (Fig. 5) has a 1,110 bp deletion extending from the *Nsi*I site to the *Kpn*I site in the *ilvG* gene, and plasmid pCC20 has a 1,290 bp deletion extending from the *Xho*I site in the *ilvD* gene to the *Sac*II site in the *ilvA* gene. Both deletions, which were verified by nucleotide sequence analysis (see Materials and Methods) result in protein fusions so that polar effects due to termination of translation would not be expected (Figs. 5 and 6). Predictions can be made as to the effects of these deletions upon specific transcripts if the initial assignments were correct. The predicted changes were detected as a loss of the expected transcript and, in some instances, the appearance of shortened transcripts. In general, the shortened transcripts appear to be more diffuse and are presumably less stable than full length transcripts produced by the intact cluster.

As predicted, the 1,110 nt *ilvG* deletion in plasmid pCC17 results in the loss of the 6.7 kb transcript, the appearance of a new 5.6 kb transcript, but no effect upon the 4.6, 3.6, or 1.1 kb transcripts (Figs. 7 and 9, compare lane wt with lane G').

Also as predicted, the 1,290 bp deletion in the *ilvDA* junction in plasmid pCC20 results in the loss of the 6.7 kb transcript and the appearance of a new 5.4 kb transcript, and this result was seen with all probes, (Figs. 7 and 8, compare lane wt with lane A') except for the oliA (Fig. 9, lane A'). As expected, the probe oliA, which is based on sequences from nt 5842 to 5793, failed to hybridize to any transcripts derived from plasmid pCC20, which has a deletion extending from the *Xho*I site at nt 5,000 to the *Sac*II site at nt 6290.

The absence of any hybridization in Fig. 9, lane A', also indicates that the diffuse hybridization seen with this probe (Fig. 9, lanes wt and G'), and presumably for the other *ilv* probes as well, is specific for the *ilv* mRNAs rather than cross hybridization to other mRNAs.

The 4.6 kb transcript is not affected by the *ilvG* deletion (Figs. 7-9, lane G'), but the *ilvDA* deletion results in the loss of the 4.6 kb transcript and the appearance of a 3.3 kb transcript (Figs. 7 and 8, lane A'). This result is consistent with the interpretation that the 4.6 kb transcript does not include *ilvG* coding sequences, but does include the *ilvDA* regions.

The 3.6 kb transcript is not affected by the *ilvG* deletion (Figs. 8 and 9, lane G', probe pcrDA, oliA), but the *ilvDA* deletion results in the loss of the 3.6 kb transcript and the appearance of a 2.3 kb transcript that migrates just ahead of the 23S RNA (Fig. 8, lane A'). This result is consistent with the interpretation that the 3.6 kb transcript does not include the *ilvG* coding sequences, but does include the *ilvDA* regions.

The 1.1 kb transcript is not affected by the deletions in plasmids pCC17 and pCC20 (Fig. 7, lanes wt, G', and A'), confirming that this mRNA is derived from sequences outside of the regions deleted.

It should be noted that the *ilvG* and *ilvDA* deletions generate in frame protein fusions (Figs. 5,6) so that polar effects due to termination of translation would not be expected. The deletion in plasmid pCC17 also removes the polar frameshift site present in the wild type *ilvG* (Val^S) allele present in plasmid pJG51, and this presumably accounts for the increased transcript levels derived from plasmid pCC17 (Figs. 7 and 8, compare lane G' to lanes wt and A').

Primer Extension Detects a 5'-end of *ilvDA* transcript

The 5'-ends of the 6.7 kb *ilvGMEDA* and 4.6 kb *ilvEDA* transcripts can be attributed to the well characterized *ilvGp2* and *ilvEp* promoters, respectively. The existence of an *ilvDp* promoter has been inferred (Berg *et al.*, 1979), but no direct in vivo or in vitro experiments have been reported to characterize the 5'-end of such a transcript. Accordingly, oligonucleotide D5, with a sequence derived from nt 3323 to 3301 near the 5'-end of the *ilvD* gene, was used as a primer for reverse transcriptase using mRNA as a template. An extension product of approximately 169 nt was detected (Fig. 10) which corresponds to a 5'-end at nt 3154, located approximately 37 nt upstream of the start codon of the *ilvD*, within the intercistronic region between the *ilvE* and *ilvD*. This result, in conjunction with previous reports cited above, is consistent with the existence of an *ilvDp* promoter, but at present we cannot exclude the possibility of a processing event that generates this 5'-end.

Effect of RNase E upon Stability of the *ilvGMEDA* Transcripts

RNase E, encoded by the *rne* (also designated *ams* gene), has been postulated to have a general role in *E. coli* mRNA turnover (Mudd *et al.*, 1990). Since the *rne* mutation appeared to stabilize bulk *E. coli* mRNA, Northern blot analysis was used to examine the effect of this mutation on the decay of the *ilvGMEDA* transcripts. Isogenic *rne-3701* and *rne*$^+$ strains containing plasmid pJG51 were grown in unsupplemented minimal medium. RNA was isolated following rifampicin addition. In Fig. 11A, Northern RNA blot showed that the 6.7 and 4.6 kb transcripts were clearly more stable in the *rne-3701* strain than in the *rne*$^+$ strain. The amount of 1.1 kb transcript was similar in both strains at time zero, but the rate of decay was significantly reduced in the *rne-3701* strain relative to its *rne*$^+$ parent (Fig. 11B). It was apparent that in the *rne-3701* strain , the decay of 6.7 and 4.6 kb transcripts was rapid, and the decay of the 1.1 kb transcript was biphasic. While the transcript was initially stabilized, it decayed more rapidly between the 10- and 15-minute time points (Fig. 11B). The half life of the 1.1 kb transcript was about 9 min in the *rne-3701* strain, and about 2.5 min in *rne*$^+$ strain. The mRNA was degraded in the *rne-3701* mutant, albeit more slowly. This could be due to the activities of other endonucleases or exonucleases involved in degradation of the *ilvGMEDA* transcripts.

Effect of RNase P upon Stability of the *ilvGMEDA* Transcripts

RNase P is one of the enzymes that participates in the processing of precursor tRNAs. This enzyme generates the 5' termini of mature tRNAs by cleavage of extra nucleotides from the 5' termini of precursor tRNAs. The role of RNase P on processing of mRNA is not as well defined. To determine if RNase P might affect the stability of the *ilv* transcripts, the RNase P deficient strain NHY322 (*rnpA49*^{ts}) and the isogenic *rnpA*$^+$ strain NHY321 were transformed with plasmid pJG51 to generate strains CC287 and CC283, respectively, to monitor mRNA stability. The cultures were shifted to the non-permissive temperature (43°C) for 10 min, and rifampicin was added (T_0 in Fig. 12). The *ilv* mRNA was stabilized in *rnpA49* strain as shown in Northern blot analysis (Fig. 12A). The half life of the *ilv* mRNA is greater in the mutant than in the wild type, especially the 1.1 kb *ilvE* transcript (about 6 min in *rnpA49* strain, and 2 min in *rnpA*$^+$ strain).

Effect of RNase III upon Stability of the *ilvGMEDA* Transcripts

The endoribonuclease III, the *rnc* gene product of *E. coli*, seems to act as the rate-limiting step in the degradation of a few specific mRNAs (Bardwell *et al.*, 1989; Portier *et al.*, 1987; Schmeissner *et al.*, 1984). The RNase III mutant strain BL321 (*rnc105*) does not show any severe growth deficiency, although the enzyme is normally involved in the maturation of ribosomal RNAs (Nikolaev *et al.*, 1985). To determine if RNaseIII might be involved in the *ilvGMEDA* transcripts degradation, the RNase III mutant strain BL321 (*rnc-105*) and the wild type isogenic strain BL322, harboring plasmid pJG51 were grown at 37°C to mid-log phase. Total RNA was isolated after rifampicin treatment, and analyzed by Northern blotting (Fig. 13A). As shown in Fig. 13B, the turnover of the *ilvGMEDA* transcripts occurred to the same extent in the RNase III mutant and wild type, indicating the degradation of the *ilvGMEDA* transcripts was independent of RNase III.

Effect of RNase II and PNPase upon Stability of the *ilvGMEDA* Transcripts

The turnover of mRNA in *E. coli* is mediated by a combination of endonucleolytic and exonucleolytic enzyme activities (Belasco and Higgins, 1988). RNase II and polynucleotide phosphorylase (PNPase) are major exonucleases that participate in the chemical decay of mRNA (Donovan and Kushner, 1986). Cells lacking both RNase II and PNPase are inviable. The loss of cell viability in double mutant (*rnb-500*^{ts}, *pnp-7*) strain at non-permissive temperature is related to an in vivo accumulation of RNA species 100-1500 nucleotides long and to an increase in the chemical half life of total mRNA. To determine the role of RNase II and PNPase in degradation of the *ilvGMEDA* transcripts, strain SK5003 (*pnp-7*, *rnb-500*^{ts}), SK5004 (*pnp-7*, *rnb*$^{+}$), SK5005 (*pnp*$^{+}$, *rnb-500*^{ts}), and SK5006 (*pnp*$^{+}$, *rnb*$^{+}$) were transformed with plasmid pJG51 and used to examine the *ilv* transcripts stability. Cells were grown in minimal medium supplemented with leucine and threonine at 30°C to mid-log phase, then shifted to 44°C for 10 minutes, rifampicin was added (T_0, Fig. 14), and cells were harvested 0, 5, 10, 15 and 30 min later. RNA was isolated for Northern blot analysis using oliE probe (Fig. 14A). After incubation at 44°C, the *ilv* mRNA levels were higher at time zero (corresponding to rifampicin addition) in both single mutants and in the double mutant, compared to the wild type (Fig. 14B). These results indicated that RNase II and PNPase participate in *ilv* mRNA turnover.

The mRNA levels in the four wild type strains (Figs. 11-14) vary due to differences in exposure times to X-ray film, and possibly due to differences in genetic background and growth supplements.

Effect of End Product Amino Acids upon the *ilvGMEDA* Transcripts

Expression of the *ilvGMEDA* cluster is regulated by the levels of branched chain amino acids, and is affected by various regulatory loci, both linked and unlinked (Umbarger, 1987). RNA was isolated from strain FD1054 derivatives (Fig. 15A) containing plasmids pJG51 (Val^S) or pCC36 (Val^R) grown in medium containing no branched chain amino acids (lane M), repressing concentrations of all three end products (lane X), or in the presence of limiting concentrations of one of the three end products (lanes L, I, and V). Based on direct enzyme assay (Fig. 3), these prototrophic strains would be predicted to have the lowest mRNA levels in the presence of an excess of all three end product amino acids, intermediate levels in the absence of the three end products, and the highest levels when isoleucine, leucine, or valine is limiting. Isoleucine limitation (the presence of excess valine and leucine) inhibits growth of the Val^S but not the Val^R strain, and therefore causes a more severe limitation in the Val^S strain. Also, the Val^R strain does not have the polar frameshift site in *ilvG*, so that higher levels of the downstream *ilvEDA* transcripts would be predicted.

As expected for the Val^S strain (Fig. 15B), the lowest, intermediate, and highest mRNA levels correspond to growth with an excess of all three end products, no end products added, and isoleucine starvation, respectively. Leucine or valine limitation resulted in mRNA levels intermediate between those detected with an excess of all three end products and unsupplemented minimal medium.

As expected for the Val^R strain (Fig. 15B), higher levels of transcripts were present compared to the Val^S strain in the presence of all three end products, with no end products added, and with limiting leucine or valine. The Val^R strain had the lowest mRNA levels in the presence of all three end products, intermediate levels with limiting leucine, isoleucine or valine, and the highest levels in unsupplemented minimal medium.

The 6.7, 4.6, 3.6, and 1.1 kb transcripts appear to change in parallel in response to amino acid levels. These results do not provide support for the downstream amplification model (Smith *et al.*, 1976), which predicts that, under derepressing conditions, there is an increase in full length *ilvGMEDA* transcripts at the expense of shorter transcripts postulated to arise from internal Rho dependent transcription termination sites. However, these Northern blots represent steady state mRNA levels, and we have not yet measured the separate rates of synthesis and degradation of each transcript under these growth conditions.

Effects of the *ilvA538* Regulatory Allele and the *ilvG671* (Val^R) Allele upon the *ilvGMEDA* Transcripts

A variety of experimental results implicate the *ilvA* gene product, threonine deaminase, as an autoregulatory protein that affects the expression of its own gene and those coding for some related proteins. The most direct evidence that threonine deaminase affects the *ilv* gene expression comes from mutations in *ilvA* gene that affect repression control such as the *ilvA538* allele, which increases the sensitivity of threonine deaminase to the allosteric inhibitor, isoleucine, and to the remote effector, leucine (Calhoun, 1976). Direct enzyme assay of the plasmid and phage coded *ilvA538* gene product in Δ*ilv* hosts confirmed the feedback hypersensitivity of the enzyme product (Calhoun and Gray, 1982).

To test the postulated regulatory models (Calhoun and Hatfield, 1975) proposing that the *ilvA* gene product affects the expression of the *ilvGMEDA* cluster, strain CU505 (Δ *ilvGMEDAYC2014*) was transformed with plasmids pJG51(Val^S, *ilvA*$^+$), pCC40 (Val^R, *ilvA*$^+$), and pJG45 (Val^R, *ilvA538*). As indicated in Fig. 16A, RNA was isolated from cells grown in minimal medium (lane M), in the presence of excess (lane X) leucine, isoleucine, and valine, or in the presence of limiting concentrations of one of the three end products leucine, isoleucine, or valine (lane L, I, and V respectively). Comparing Val^R, A$^+$ with Val^S, A$^+$, higher levels of *ilv* mRNA can be seen in presence of *ilvG671* (Val^R) allele, than in *ilvG*$^+$ (Val^S) allele (Fig. 16A). It is possible to test for an effect of the *ilvA538* allele by comparison of the Val^R, *ilvA*$^+$ and the Val^R, *ilvA538*. The mRNA levels (Fig. 16A) in the presence of the *ilvA538* allele are much lower than in the isogenic *ilvA*$^+$ comparing pCC40 (Val^R, *ilvA*$^+$) with pJG45 (Val^R, *ilvA538*) particularly for cells grown in unsupplemented minimal medium (Fig. 16A compare lanes M). For example, at the highest levels of expression (Fig. 16B, column M, minimal medium) the mutation reduces transcript levels 12-fold (10.0/0.83) for the *ilvGMEDA*, 6.9-fold for *ilvEDA* (10.9/1.59), and 2.9-fold for *ilvE* (24.2/8.28). At lower levels of expression (X, L, I, and V) the *ilvA538* mutation results in a 2- to 3-fold effect, or less. This result indicates that the down regulatory effect of the *ilvA538* allele on the *ilvGMEDA* genes, which were previously detected only by enzyme assay for the *ilvGMEDA* products, is actually mediated, at least in part, by decreased transcripts levels.

Effect of the *rho115* Allele upon the *ilvGMEDA* Transcripts

Premature termination of translation of a polycistronic mRNA can reduce transcription of operon distal genes in prokaryotes. Transcriptional polarity is the result of a mutation that causes termination of translation within a proximal gene of a polycistronic gene cluster. Three Rho dependent termination sites within *ilvG* and *ilvM* genes are responsible for transcriptional polarity in the *ilvGMEDA* cluster in wild type *E. coli* (Wek *et al.*, 1987). The *ilvG* gene in wild type (Val^S) *E. coli* contains a frameshift site which results in termination of translation near the middle of the gene, so that the Rho dependent termination sites are recognized by Rho factor and transcription stops at these sites. Mutations either in the *ilvG* gene (which restore a full length protein) or in the *rho* gene relieve this polarity (Smith *et al.*, 1976; Lawther *et al.*, 1981 and 1982). Strains TN1276 (*rho*$^+$) and TN1271 (*rho-115*) were transformed with plasmids pJG51 (Val^S) and pCC40 (Val^R). RNA samples were prepared from cells grown in minimal medium supplemented with leucine (a necessary growth factor for this isogenic pair), and analyzed by Northern blot using probe oliE (Fig. 17A).

A very revealing result was that in strains with an altered Rho factor, the 4.6 kb *ilvEDA* transcript increased (Fig. 17A and B, compare lane 1 with 2; lane 3 with 4), but the 1.1 kb *ilvE* transcript decreased (Fig. 17A and B, compare lane 3 with 4). The most straightforward interpretation of this result is that the presence of a Rho dependent termination site causes the release of the 1.1 kb *ilvE* transcript, rather than readthrough to generate the 4.6 kb *ilvEDA* transcript in wild type Rho^{+} strain.

The effect of the altered Rho factor differs, as expected, in the $\mathrm{Val}^{\mathrm{R}}$ and $\mathrm{Val}^{\mathrm{S}}$ strains. The $\mathrm{Val}^{\mathrm{S}}$, but not the $\mathrm{Val}^{\mathrm{R}}$, strain has the frameshift site in *ilvG* followed by three Rho dependent transcriptional termination sites. As indicated in Fig. 17 B, in the $\mathrm{Val}^{\mathrm{R}}$ strain, the Rho mutation affects only the postulated Rho dependent termination site between *ilvE* and *ilvD*, and causes the *ilvE* transcript to decrease 3.2-fold (19.9/6.25) and the *ilvEDA* transcript to increase 2.3-fold (7.46/3.26). In the $\mathrm{Val}^{\mathrm{S}}$ strain, the Rho mutation affects both transcriptional termination sites downstream of the frameshift site in *ilvG* and the postulated Rho dependent termination site between *ilvE* and *ilvD*, and causes the *ilvE* transcript to decrease only 1.2-fold (11.1/9.09) and the *ilvEDA* transcript to increase 3.4-fold (6.38/1.90). We interpret this to indicate that in the $\mathrm{Val}^{\mathrm{S}}$ strain the Rho mutation results in more readthrough of the polar sites in *ilvG*, resulting in an increase of the total *ilvE* and *ilvEDA* transcripts. This would account for the greater increase of the *ilvEDA* transcript in the $\mathrm{Val}^{\mathrm{S}}$ (3.4-fold) than in the $\mathrm{Val}^{\mathrm{R}}$ (2.3-fold) strain, and for the greater decrease of the *ilvE* transcript in the $\mathrm{Val}^{\mathrm{R}}$ (3.2-fold) than in the $\mathrm{Val}^{\mathrm{S}}$ (1.2-fold).

DISCUSSION

Four major transcripts derived from the *ilvGMEDA* cluster have been detected, and these transcripts encode the products of the *ilvGMEDA*, *ilvEDA*, *ilvDA*, and *ilvE* genes. The *ilvGMEDA* and *ilvEDA* transcripts can be interpreted in terms of the previously characterized transcription start sites, *ilvGp2* and *ilvEp*, (Adams *et al.*, 1985; Lopes and Lawther, 1989; Pereira *et al.*, 1988; Wek and Hatfield, 1986), with 5'-ends at nt 1 and 2101, respectively, (Fig. 4C) and the single termination site at the end of the cluster at nt 6703 (Sameshima *et al.*, 1989). A transcript initiated at the weak internal promoter, *ilvAp* (Lopes and Lawther, 1989) is not detected, but this can be attributed to the low levels of activity of this promoter under the aerobic conditions used in this study. In addition, truncated transcripts derived from the *ilvGM* region corresponding to those generated in vitro from the wild type (Val^S) *ilvG* gene (Lopes *et al.*, 1989; Wek *et al.*, 1987) are not detected, suggesting that these are less stable in vivo compared to the four major transcripts.

Two transcripts are detected that were not predicted by the previously established promoters and terminators, and these encode the *ilvE* and *ilvDA* gene products. It remains to be definitively established whether these two mRNAs are primary transcripts, or the result of posttranscriptional events. Since the presence of the *ilvEp* promoter has been documented, it is possible that the *ilvE* transcript results from initiation at *ilvEp* and termination or processing near the *ilvED* intercistronic region. Primer extension analysis identified the 5'-end of the *ilvDA* transcript to be in the *ilvED* intercistronic region, and this transcript presumably extends to the previously described termination site following the *ilvA* gene.

It is apparent that whereas the *ilvEDA* and *ilvE* transcripts often are present as discrete and relatively intense transcripts, the *ilvGMEDA* transcript is accompanied by a broad distribution of degradation form. The *lacZYA* operon of *E. coli* has been reported (Murakawa *et al.*, 1991) to produce multiple transcripts, including distinct *lacZYA*, *-ZY*, *-Z*, and *-A* messages, and more diffuse *lacY*, and *-YA* transcripts. Similarly, the *ilv* transcripts vary in that the *ilvEDA* and *ilvE* messages are usually sharper and more distinct than the *ilvGMEDA* and *ilvDA* messages. Most of the Northern blots have a background of the *ilv* specific hybridization that presumably represents breakdown products of the major transcripts. We assume this represents the distribution of species present in vivo, but we cannot exclude the possibility that some degradation occurs after extraction.

Identification of some of the potential sites of transcription initiation and termination is not sufficient to predict the actual populations of the *ilvGMEDA* coded transcripts present in the cell, particularly during the response to regulatory signals. Specifically, several questions remain about factors affecting the expression of these genes, including possible sites of internal transcription termination, differential mRNA stability of specific segments, the presence of specific mRNA segments generated during the early stages of mRNA turnover, and changes in transcription patterns during repression or derepression in response to changes in the end product amino acids, leucine, isoleucine, and valine.

Many factors were postulated to specifically affect expression of the *ilvGMEDA* cluster as measured by direct enzyme assay (Umbarger, 1987). In this study, we used plasmids having mutations in *ilvA538* and *ilvG671* for Northern blot analysis. The *ilvG* mutation [compare plasmid pCC40 (*ilvG671*, *ilvA*$^+$) to plasmid pJG51 (*ilvG*$^+$, *ilvA*$^+$) in Fig. 16], increases the expression of the *ilvGMEDA* transcripts as predicted for the loss of the frameshift site. The *ilvA538* allele decreases the *ilv* transcript levels [compare plasmid pJG45 (*ilvG671*, *ilvA538*) to plasmid pCC40 (*ilvG671*, *ilvA*$^+$) in Fig. 16] indicating that the *ilvA* gene product, threonine deaminase has autoregulatory effects on expression of the *ilvGMEDA* transcripts.

It was reported that in the wild type (Val^S) cell growing under repressing condition, *ilvEp* accounts for 41% (Wek *et al.*, 1987), 62% (Calhoun *et al.*, 1985), and 90% (Lopes *et al.*, 1989) of downstream gene expression. The results present in this work confirm these analysis that under repressing conditions, the *ilvEDA* transcript is much more abundant than the *ilvGMEDA* transcript in Val^S strains. Unexpectedly the same is true for Val^R strains (Figs. 15 and 16).

As shown in Fig. 15 and 16, the four *ilv* transcripts change in parallel in response to the end product amino acids. The changes of the 6.7 kb *ilvGMEDA* transcript from repression to derepression could be explained by the deattenuation mechanism in upstream of *ilvG* gene. Since *ilvEp* is probably not regulated (Harms and Umbarger, 1991), the changes of the 4.6 kb *ilvEDA* and 1.1 kb *ilvE* transcripts could be due to processing of *ilvGMEDA* transcripts to *ilvEDA* and/or *ilvE* transcripts, so that the attenuator indirectly controls the levels of *ilvEDA* and *ilvE* transcripts. Alternatively, some other possible mechanisms, such as an effect of end product levels upon mRNA stability or Rho dependent termination within the *ilvGMEDA* genes cluster cannot be excluded.

It is obvious that the *ilvEDA* and *ilvE* transcripts are relatively more stable than the *ilvGMEDA* and *ilvDA* transcripts in Northern blots. Some specific *ilv* sequences may account for these differences in stability of the *ilv* transcripts. It has become clear that untranslated segments of procaryotic messages can influence gene expression through their impact on mRNA stability. There is considerable evidence that a short sequence having the potential to fold into a stem-loop structure is found at 3'-end of many bacterial mRNA species and plays an important role in protecting mRNA from digestion by 3' exoribonucleases (Higgins *et al.*, 1988; Newbury *et al.*, 1987). When located internally in polycistronic mRNAs, such sequences can lead to segmental stability, with the downstream cistrons being more vulnerable to 3' exonucleolytic attack (Chen *et al.*, 1988; Mott *et al.*, 1985). It is apparent that the *ilvE* transcript often constitutes a major fraction of the *ilvGMEDA* encoded transcripts present in Northern blots, while the *ilvDA* transcript is present in lower quantities and is not as discrete in size. These two transcripts are separated by a potential stem and loop structure in the intercistronic region between *ilvE* and *ilvD* that may affect mRNA stability or processing, possibly by acting as a barrier to 3'-to-5' specific exonuclease action. It has been reported that the 5' untranslated region of mRNA also can control mRNA stability (Belasco *et al.*, 1986; Cho and Yanofsky, 1988; Gorski *et al.*, 1985). For example, the 5' untranslated RNA segment of the *ompA* functions as a growth-rate-regulated mRNA stabilizer whose activity is unrelated to translation efficiency.

Substitution of the *ompA* mRNA segment for the corresponding segment of the labile *bla* gene transcripts prolongs their lifetime by a factor of 6 (Emory and Belasco, 1990). It was observed that the *ilvGMEDA* transcript is accompanied by a broad distribution of degradation forms, while the *ilvEDA* transcript forms a more discrete and relatively intense band in Northern blots. Since these two transcripts apparently differ only in the 5'-proximal *ilvG* sequences, it is possible that the 5'-end of the *ilvEDA* transcript confers greater stability relative to the 5'-end of the *ilvGMEDA* transcript, or that sequences at the 5'-end of the *ilvGMEDA* transcript promote processing to form the *ilvEDA*, *ilvE,* or *ilvDA* transcripts.

The *ilvEDA, ilvE,* and *ilvDA* transcripts could be primary transcripts, or the products of processing of larger transcripts, or some combination of both. When RNA sample are prepared immediately after rifampicin addition (Figs. 11-14), there was no obvious evidence for precursor to product relationships in either the wild type or in ribonuclease deficient strains. It remains to be established what fraction of the *ilvEDA, ilvE* and *ilvDA* mRNA derive from primary transcription events, and what fraction is the result of posttranscriptional events.

If the *ilvEDA* and *ilvE* transcripts are primary products, they could be initiated at the internal promoter *ilvEp.* The *ilvEDA* transcript could terminate at the end of the *ilvGMEDA* cluster at nt 6703. The *ilvE* transcript might end within the intercistronic region between *ilvE* and *ilvD* genes, or in *ilvD* gene. It is possible that there is a Rho dependent termination site that generates 3' end of the 1.1 kb *ilvE* transcript. In a *rho-115* host strain the amount of 4.6 kb *ilvEDA* transcript increases, but the amount of 1.1 kb *ilvE* transcript decreases compared with the mRNA levels in the wild type (rho^+) host strain (Fig. 17). This putative Rho dependent termination site would cause the release of the 1.1 kb *ilvE* transcript rather than readthrough to generate the 4.6 kb *ilvEDA* transcript in wild type (rho^+) host strain. The presence of such an internal Rho dapendent terminator is consistent with other reported observations that mutations in the *rho* gene produce elevated levels of the *ilvEDA* gene products determined by direct enzyme assay. The effect is greater downstream (Baez *et al.*, 1979; Smith *et al.*, 1976), indicating that Rho dependent site may be present in *ilvGMEDA* cluster. Taking these together, it is logical to propose the presence of a Rho dependent termination site, probably in the intercistronic region between *ilvE* and *ilvD* genes.

Detection of a 3.6 kb *ilvDA* transcript in Northern blots, along with other genetic data (Berg and Shaw, 1981; Berg *et al.*, 1979; Blazey and Burns, 1979; Gray *et al.*, 1982) is consistent with the possible presence of an *ilvDp* promoter. However, the activity of such an *ilvDp* promoter is not detected on the 1,659 bp *Hind*III fragment containing *ilv'ED'* segment in vitro transcription (C. Squires, unpublished) or using fusions to β-galactosidase (F. Huang, unpublished). Two possibilities were considered: (i) the potential *ilvDp* promoter exists, but the activity is very low and expressed only at certain conditions, or (ii) there is no *ilvDp* promoter, so that the 3.6 kb *ilvDA* transcript detected in the Northern blots (Figs. 8, 9 and 15) would derive from the processing of the *ilvEDA* or *ilvGMEDA* precursors.

The level of gene expression is determined primarily by three factors: the rate of transcription, the efficiency of translation, and the stability of the mRNA. In *E. coli* some mono- and polycistronic transcripts decay at different rates for different segments (von Gabain *et al.*, 1983; Belasco *et al.*, 1985; Newbury *et al.*, 1987). Thus, it has been suggested that the decay of a transcript, or a segment thereof, is controlled by a rate-limiting step followed by rapid degradation mediated by a combination of endonucleolytic and exonucleolytic enzyme activities (Blundell and Kennell, 1974; Schmeissner *et al.*, 1984; Belasco and Higgins, 1988). The major exonucleases that participate in the chemical decay of mRNA, RNase II and PNPase, degrade single-strained RNA exonucleolytically and processively in the 3'-to-5' direction in vitro (Donovan and Kushner, 1986). Several endonucleases have been identified so far in *E. coli* to be involved in mRNA degradation. The functional inactivation of polynucleotide phosphorylase mRNA is initiated by RNase III processing at the 5'-end (Portier *et al.*, 1987). RNase E is also involved in processing 9S RNA (Ghora and Apiron, 1978), RNA I (Tomcsanyi and Apiron, 1985) and T_4 mRNA (Mudd *et al.*, 1988), and it was the first endoribonuclease identified as having a general role in mRNA degradation in *E. coli* (Mudd *et al.*, 1990). A comparison of the sequences flanking the RNase E processing sites of these three transcripts revealed some striking similarity, which is a G-A/U-cut-A-U-U motif followed by a stem-loop structure (Mudd *et al.*, 1990). In a RNase E mutant the rate of chemical decay of bulk *E. coli* mRNA has been shown to reduced, and some specific mRNAs, including *oriC* and *ompA* mRNA, have been shown to be more stable (Mudd *et al.*, 1990). In this study, strains with mutations in exonucleolytic and endonucleolytic enzymes were used to test their role in the *ilv* transcripts stability. These results indicate that (i) exonucleases, RNase II and PNPase play a role in *ilv* mRNA turnover (Fig. 14), (ii) endoribonucleases RNase E and RNase P are both involved in degradation of *ilv* mRNA (Figs. 11 and 12), and (iii) endoribonuclease RNase III seems to have no effect on *ilv* mRNA stability (Fig. 13). Presumably, different transcripts in *E. coli* are subject to decay by what may be loosely defined as different "degradation pathways" depending on which endonucleolytic activity is responsible for the rate-limiting step. Subsequent to initial attack(s) by endonucleolytic activities, the mRNA molecules are supposedly degraded in a processive manner by the concerted action of exonucleases in 3'→5' direction (Belasco and Higgins, 1988; Brawerman, 1987).

The results in this study indicate that the expression of the *ilvGMEDA* genes involves four major transcripts that differ in their stabilities and levels of expression. The role of these transcripts in response to changes in end product amino acid levels remains to be established.

PART II

Nucleotide Sequence and Transcriptional Activity of the *Escherichia coli* K-12 Chromosome Region Between *rrnC* and *ilvGMEDA*

INTRODUCTION

The focus of the this study is an *ilv*-related protein of 18,751 daltons, the product of a gene designated ORFI, that is expressed at higher levels when the polar frameshift site in *ilvG* is removed (Gray *et al.*, 1981, 1982). The product of the ORFI locus was first detected (Gray *et al.*, 1981) by monitoring the synthesis of (^{35}S) methionine labeled proteins produced following infection of UV irradiated cells with λd*ilv* phage containing the wild type (Val^{S}) or several mutant (Val^{R}) alleles of *ilvG*. This *ilv*-related protein was only detected when the λd*ilv* phage contained a mutated *ilvG* allele which had lost the polar frameshift site. This pattern of expression would be consistent with a location for ORFI downstream of the polar frameshift site in *ilvG*. However, when the approximate location of the sequences coding for the ORFI product was determined in maxicells by testing a variety of restriction fragments cloned in plasmid pBR322 (Gray *et al.*, 1981,1982), the gene was found to be outside of the *ilvGMEDA* cluster, and close to a *Hind*III site approximately 2.1 kb from the transcription start sites for *ilvGMEDA*. Thus, there was no immediately obvious explanation for the increased expression seen in the presence of the Val^{R} alleles. In this study we report the entire 3.5 kb sequence between *ilvGMEDA* and *rrnC*, and we have identified ORFs for the *ilv*-related protein (ORFI) and for another ORF (ORFII). We have detected specific transcripts corresponding to ORFI, which is transcribed in a direction opposite to that of *rrnC* and *ilvGMEDA*, by slot blots and Northern blots. The ORFII peptide is not produced at detectable levels in maxicells, even when the gene is transcribed by an exogenous promoter.

MATERIALS AND METHODS

Bacterial Strains and Plasmids

As indicated in Table I, the maxicell analyses used strain CSR603 (Sancar *et al.*, 1979). Restriction fragments were cloned to plasmids pBR322 or pUC9 (Fig. 18) as previously described (Gray *et al.*, 1981). Nucleotide sequence analysis (Fig. 19) by the chain termination method (Sanger *et al.*, 1977) used bacteriophage mp18 and mp19 for the preparation of template DNA using strain JM103 as host (Messing, 1983). Plasmid pMC81 (Casadaban and Cohen, 1980) was used to clone *Hind*III fragments in the orientations indicated in Fig. 18.

Nucleotide Sequence Analysis

Both the chemical (Maxam and Gilbert, 1977) and enzymatic (Sanger *et al.*, 1977) sequencing protocols were used. The method of Dale *et al.*, (1985) was used to construct deletion derivatives of mp18 and mp19 derivatives containing several restriction fragments spanning the region show in Fig. 18. Other sequencing methods were as described (Cox *et al.*, 1987). Computer analyses of DNA sequences and database searches were conducted on an IBM PS2 Model 80 personal computer using MicroGenie and PC/GENE software, and on a MicroVAXII using the Wisconsin Genetics Computer Group software.

Maxicell Experiments

After transformation of the *E. coli* strain CSR603 with appropriate plasmid, a single ampicillin resistant transformant was isolated and purified. The protocol was previously described (Sancar *et al.*, 1979; Calhoun and Gray, 1981; Gray *et al.*, 1982), except that after irradiation, methicillin (10 μg/ml) and carbenicillin (10 μg/ml) were included along with cycloserine (20 μg/ml) during overnight incubation. Next day, 1.5 ml of the irradiated cells were used for detection of the *ilv*-related peptide coded by restriction enzyme fragments in plasmids (Fig. 20), and the rest of cells were used for RNA preparation.

Isolation and Analysis of RNA

RNA was isolated as described in Part I. The ORFI specific probe used for hybridization (Fig. 21) to mRNA (5'-tatggtttccctgataatgtcc ggtagggctgggattcat-3') was a synthetic oligonucleotide prepared on a Biosearch Cyclone DNA Synthesizer that was designed to be complementary to nt 2825 to 2864 (V, in Table III). We considered the possibility that the hybridization we observed (Fig. 21) could be due to contamination of the RNA preparations with trace amounts of plasmid DNA. This possibility seems unlikely since (i) the RNA preparations were treated with DNase during the isolation procedure, (ii) treatment of the RNA samples with additional DNase did not affect the hybridization observed, (iii) treatment of the RNA preparations with RNase eliminated hybridization, and (iv) another synthetic oligonucleotide corresponding to nt 2531 to 2553 (5'gctgcgatattaagtctggtaaa-3', D, in Table III) in the *Hind*III to *Sma*I fragment (Figs. 18 and 19), with the same sequence as the ORFI mRNA rather than complementary as for the hybridization probe described above, did not hybridize to the mRNA preparations.

RESULTS AND DISCUSSION

Nucleotide Sequence of 3.5 kb Segment Between *rrnC* and *ilvG*

The nucleotide sequence of 4890 bp extending from the *EcoR*I site in *rrnC* to the *Kpn*I site in *ilvG* was determined from both strands using a combination of the chemical (Maxam and Gilbert, 1980) and enzymatic (Sanger *et al.*, 1977) methods (Fig. 19). Computer analyses were carried out with software from MicroGenie, PC/GENE, and the University of Wisconsin Genetics Computer Group. We confirmed the published sequence for the 3'-end of the *rrnC* cluster (Young, 1979) and the 5'-end of *ilvG* (Lawther *et al.*, 1981), and we determined approximately 3.5 kb of new sequence that linked the known sequences for *rrnC* and *ilvG.* Two ORFs (Figs. 18 and 19), that would be transcribed counterclockwise, opposite in direction to the flanking *rrnC* and *ilvGMEDA* clusters, were analyzed by the codon frequency method of Fickett (1982) and the predicted coding probability values were 77% and 98% for ORFI and ORFII, respectively. We found no relevant homologies when we analyzed ORFI and ORFII using the PROSITE routine (Bairoch, 1989) of PC/GENE which uses pattern matching to detect 159 amino acid sequence profiles specific to a family or group of proteins, such as signal peptides, DNA binding motifs, enzyme active sites, and others. Also, when the entire region between *rrnC* and *ilvGMEDA*, including the segments including ORFI and ORFII, were searched for relatedness to the GenBank database (release 54), no significant homologies were detected. Several smaller ORFs encoding less than 100 amino acids are present in the *EcoR*I to *Kpn*I region. These predicted proteins, however, were not seen in maxicell analyses (Gray *et al.*, 1981 and 1982), no homologies were seen when these ORFs were compared to the DNA and protein databases, and their significance is unknown.

As shown in Fig. 22, from the pairwise alignment, ORFII is detected to have homology with *ilvY*, one of the LysR family of transcriptional regulatory proteins (Henikoff, 1991, personal communication). LysR is a very large family with more than 20 diverse members known, characterized by a helix-turn-helix DNA binding motif within the N-terminal region (Fig. 23).

The *ilvY* coded regulatory protein which activates transcription of the *ilvC* gene (Watson *et al.*, 1979). It is interesting that *ilvY* and ORFII form a direct repeat, with divergently transcribed *ilvGMEDA* and ORFIII located in between.

The ORFI begins at nt 2826 and extends to a stop codon at nt 3314 (Fig. 19). We noted, however, that another ATG codon in the same translation frame is located upstream at nt 1767. This is hypothetical peptide beginning at nt 1767, designated ORFIII in Fig. 19, would extend from nt 1767 to 3314, and would contain ORFI as a C-terminal segment. The hypothetical ORFIII peptide would potentially encode a peptide of 516 amino acids, and would have a coding probability of 92% based on codon utilization. We have never detected a protein corresponding to plasmids containing this coding region and flanking regions (Gray *et al.*, 1981; 1982). This potential ORFIII coding region has no significant homologies to *ilvGM, ilvBN* or *ilvIH*. However, Dr. Henikoff has some interesting findings (personal communication). In his search of the SWISS-PROT 17 database, it shows that the ORFIII very likely matches to two proteins, Rhodobacter CrtA, encoding spheriodene monooxygenase (Armstrong, *et al.*, 1989), and a "novel" *Arabadopsis* chloroplast protein (Koncz *et al.*, 1990). Similar results are obtained using another program, FastA. The partial multiple alignment of ORFIII with these two proteins (Fig. 24) shows the ORFIII is likely to be a very distant homology. The most similar regions of ORFIII and other two include the most highly conserved regions seen from the comparison of *CrtA* and the "novel" protein. It is notable that one of the highly similar regions looks like an ATP-binding site motif (the consensus for which is A/G-x-x-x-x-G-K-S/T), and these proteins are essentially the same lengths, with similar regions at the same positions in the sequence, suggesting global similarity. Now there is no logical interpretation for this finding, and the significance is unknown.

Several possibilities to account for the absence of an ORFIII peptide in maxicells include: (i) irradiated maxicells may lack the proper signals to allow for expression of ORFIII; (ii) ORFs that lack biological significance are present in DNA sequences and ORFIII could simply represent an example of a large, randomly occurring ORF; and (iii) the ORFIII peptide could be subject to proteolytic processing or degradation in irradiated maxicells. ORFI could arise as a partial degradation product of ORFIII or it could arise by initiation of translation at a suitably placed ribosome binding site at position 2812. The production of the ORFI peptide does not depend upon translation of the upstream region in ORFIII since plasmids such as pMS147 that contains the *Hind*III to *Sma*I fragment (nt 2344 to 3666) lead to the production of ORFI protein.

Maxicell Analysis for Detection of the *ilv*-related Peptide Encoded by ORFI

Strain CSR603 (*uvrA6, recA*) is very sensitive to UV light, and appears to extensively degrade chromosomal DNA, but not plasmid DNA following mild doses of irradiation. Accordingly, plasmid coded proteins can be detected by radioactive labeling of proteins synthesized after UV irradiation of plasmid containing derivatives of CSR603 (Sancer *et al.*, 1979). As indicated in Fig. 18, the restriction fragments were cloned to pBR322 and/or to pUC9 and used to transform strain CRS603 for maxicell analysis.

Plasmid pMS147, which contains a *Hind*III to *Sma*I restriction fragment of 1319 bp cloned to pUC9, produces the *ilv*-related peptide (Fig. 20, lane 4). An examination of ORFI, which is located on this restriction fragment, revealed a *Cla*I site near the region coding for the carboxyl-terminal region of the peptide, and extending 504 bp to another *Cla*I site downstream of the ORFI stop codon. In order to confirm that ORFI encodes the *ilv*-related peptide, this *Cla*I fragment was removed from plasmid pMS147, generating plasmid pCC2. The predicted size of the shortened protein, which extends past the distal *Cla*I site until a stop codon is encountered, is 16,930 daltons. As seen in Fig. 20, lane 5, the size of the *ilv*-related protein seen in maxicells decrease as expected. This result indicates that ORFI encodes the *ilv*-related peptide. We note that much less ORFI protein is produced in maxicells containing pMS147 (Fig. 20, lane 4) compared to those containing plasmid pMS143 (Fig. 20, lane 3). This could be due to specific regulatory effects or due to read-through from vector promoters.

Detection of ORFI Specific Transcripts

RNA was isolated from cells with and without specific plasmids in order to test for the presence of transcripts specific for ORFI. Northern blots (Fig. 21) using an ORFI specific probe and RNA from cells containing plasmid pMS147 revealed the presence a major transcript of approximately 3.4 kb (Fig. 21, lane 3), while a major transcript of approximately 2.9 kb is detected in cells containing plasmid pCC2 (Fig. 21, lane 4). We assume that the lower M_r material seen in Fig. 21, lane 3 and 4, represents specific hybridization to lower M_r forms of ORFI mRNA. It is evident that the transcripts of 3.4 kb and 2.9 kb are initiated within and or extend into vector sequences since plasmid pMS147 contains only 1.3 kb of *E. coli* DNA. The main purpose of the present analysis is to identify the location of the DNA segment encoding the *ilv*-related peptide, ORFI, and to confirm that it is expressed in vivo. Preliminary analysis of these ORF-*lacZ* fusions are consistent with expression of ORFI from the promoter designated ORF*p* in Fig. 18, that is initiated at a position 112 bp upstream of the transcript initiated at *ilvGp*2, in a direction opposite to the *ilvGMEDA* transcript (Tsui and Freundlich, 1988).

Analysis of the ORFII Region

The maxicell analysis of a variety of plasmids containing sequences encoding ORFII did not reveal a peptide of the expected size. The very high score (98%) using Fickett's method of coding probability indicated that there may be some biological significance to ORFII. If ORFII is expressed in vivo, it may require some specific cellular control signal. In order to test for the potential expression of ORFII, we cloned ORFII on a *Hind*III fragment to plasmid pMC81, so that the ORF would be transcribed from the exogenous *ara* promoter. Plasmid pMC81 also has the *lacZ* gene distal to the *ara* promoter, which make it possible to monitor transcripts that traverse the cloned restriction fragment (Fig. 18). Expression of the *ara* promoter is dependent upon the addition of arabinose to the culture medium. In the absence of arabinose, expression of *lacZ* will occur only if the cloned restriction fragment contains a promoter. Plasmids pMS135 and pMS136 are pMC81 derivatives with the *Hind*III fragment cloned in the two possible orientations (Fig. 18). Production of the ORFII peptide was not seen in maxicells containing plasmid pMS136 grown with or without arabinose or in maxicells harboring other plasmids (Gray *et al.* 1981, 1982). Cells containing plasmid pMS135 produced significant levels of β-galactosidase indicating the presence of a promoter of unknown location and significance located on the *Hind*III fragment containing *rrnC* and ORFII.

In summary, we have determined the 3.5 kb sequence between *rrnC* and *ilvGMEDA*, identified the gene encoding the *ilv*-related peptide, ORFI, demonstrated expression of mRNA transcripts specific for ORFI, and identified other ORFs (designated ORFII and ORFIII) that are located between ORFI and *rrnC*. It will now be possible to analyze the specific expression of ORFI and to determine its relationship to the *ilvGMEDA* genes.

TABLE I

Bacterial Strains Used in This Study

Strain	Description	Reference
BL321	F$^-$, *thi-1, argH1, gal-6, lacY1, mtl-2, xyl-7, malA1, ara-13, str-9, tonA2,* λ^-, *supE44, rnc-105*	Studier, 1975
BL322	F$^-$, *thi-1, argH1, gal-6, lacY1, mtl-2, xyl-7, malA1, ara-13, str-9, tonA2,* λ^-, *supE44, rnc*$^+$	Studier, 1975
C600	F$^-$, *thi-1, thr-1, leuB6, lacY1, tonA21, supE44,* λ^-	Appleyard, 1954
CC84	Strain FD1054 (Δ*ilvGMEDA732)* with plasmid pJG51 (Val^S)	this work
CC85	Strain C600 with plasmid pJG51 (Val^S)	this work
CC86	Strain C600 with plasmid pCC17 (Δ*G*, Val^S)	this work
CC89	Strain C600 with plasmid pCC20 (Δ*DA*, Val^S)	this work
CC147	Strain N3431 (*rne-3071*^{ts}) with plasmid pJG51 (Val^S)	this work
CC148	Strain N3434 (*rne*$^+$) with plasmid pJG51 (Val^S)	this work
CC221	Strain FD1054 (Δ*ilvGMEDA732)* with plasmid pCC36 (Val^R)	this work
CC245	Strain CU505 (Δ*ilvGMEDAYC2014, leu455)* with plasmid pJG51 (Val^S)	this work
CC250	Strain CU505 (Δ*ilvGMEDAYC2014, leu455)* with plasmid pCC40 (Val^R)	this work
CC253	Strain TN1271 (*rho115*) with plasmid pJG51 (Val^S)	this work
CC254	Strain TN1271 (*rho115*) with plasmid pCC40 (Val^R)	this work
CC256	Strain TN1276 (*rho*$^+$) with plasmid pJG51 (Val^S)	this work
CC257	Strain TN1276 (*rho*$^+$) with plasmid pCC40 (Val^R)	this work
CC274	Strain CU505 (Δ*ilvGMEDAYC2014, leu455)* with plasmid pJG45 [*ilv538, ilvD670, ilvG671*(Val^R)]	this work

Strain	Description	Reference
CC283	Strain NHY312 (*rnpA*$^+$) with plasmid pJG51 (Val^S)	this work
CC287	Strain NHY322 (*rnpA49*^{ts}) with plasmid pJG51 (Val^S)	this work
CC325	Strain BL321 (*rnc-105*) with plasmid pJG51 (Val^S)	this work
CC326	Strain BL322 (*rnc*$^+$) with plasmid pJG51 (Val^S)	this work
CC327	Strain SK5003 (*pnp-7, rnb-500*^{ts}) with plasmid pJG51 (Val^S)	this work
CC328	Strain SK5004 (*pnp-7, rnb*$^+$) with plasmid pJG51 (Val^S)	this work
CC329	Strain SK5005 (*pnp*$^+$, *rnb-500*^{ts}) with plasmid pJG51 (Val^S)	this work
CC330	Strain SK5006 (*pnp*$^+$, *rnb*$^+$) with plasmid pJG51 (Val^S)	this work
CSR603	*thi-1, leuB6, proA2, phr-1, recA1, argE3, thi-1, uvrA6, ara-14, lacY1, galK2, xyl-5, mtl-1, gyrA98(nalA98), rpsL31, tsx-33,* λ^-, *supE44*	Sancar *et al.*, 1979
CU2501	*rbs-200 ilvG468* (Val^R), λ^-	Umbarger, 1972
CU505	F$^-$, *galT12*, Δ*ilvGMEDAYC2014*, *leu455*, λ^-	Watson *et al.*, 1979
FD1054	*rbs-301*::Tn*5*, Δ*ilvGMEDA732*::Tn*5-131*	Gayda *et al.*, 1980
JM103	Δ(*lac-pro*), *thi, strA, supE, endA, sbcB, hsdR*$^-$, F'*traD36, proAB, lacI*^{q}*,Z*Δ*M15*	Messing *et al.*, 1981
N3431	*lacZ43, relA1, spotl, thi-1, rne-3071*^{ts}	Mudd *et al.*, 1990
N3433	*lacZ43, relA1, spotl, thi-1, rne*$^+$	Mudd *et al.*, 1990
NHY322	Δ(*proB-lac*), *ara, gyrA, thi, zic-501*::Tn10, *rnpA49*^{ts}	Kirsebon, 1988
NHY312	Δ(*proB-lac*), *ara, gyrA, thi, zic-501*::Tn10, *rnpA*$^+$	Kirsebon, 1988
SK5003	*leu*$^-$, *thr*$^-$, Cam^r, *pnp-7, rnb-500*^{ts}	Kushner, 1986
SK5004	*leu*$^-$, *thr*$^-$, Cam^r, *pnp-7*	Kushner, 1986

Strain	Description	Reference
SK5005	*leu*$^-$, *thr*$^-$, Cam^r, *rnb-500*^{ts}	Kushner, 1986
SK5006	*leu*$^-$, *thr*$^-$, Cam^r, *pnp*$^+$, *rnb*$^+$	Kushner, 1986
TN1271	*araD139, Δ(ara leu)7697, ΔlacX74, galU, galK, rpsL, rho-115*	Laboratory collection
TN1276	*araD139, Δ(ara leu)7697, ΔlacX74, galU, galK, rpsL, rho*$^+$	Laboratory collection

TABLE II

Plasmids Used and Constructed in This Study

Plasmid	Description	Reference
pBR322	Cloning vector	Bolivar *et al.*, 1977
pJG16	*EcoR*I segment from λ d*ilv581* (Val^{R}) cloned to pBR322, *ilvGMEDA'* (Val^{R})	Gray *et al.*, 1981
pJG51	tandem *EcoR*I segments from λ *c112* cloned to pBR322, *ilvGMEDAYC* (Val^{S})	Gray *et al.*, 1982
pJG45	tandem *EcoR*I segments from λ d*ilv582* to pBR322, *ilvGMEDAYC ilvA538 ilvD670 ilvG671* (Val^{R})	Calhoun and Gray, 1982
pCC17	pJG51 with *Nsi*I-*Kpn*I deletion in *ilvG* (Val^{S})	this work
pCC20	pJG51 with *Xho*I-*Sac*II deletion in *ilvDA* (Val^{S})	this work
pCC36	*ilvGMEDAYC* (Val^{R}) from a Val^{R} transductant of strain CC84 [pJG51 (Val^{S}) in strain FD1054] with P1.CU2501 (*ilvG468*, Val^{R})	this work
pCC40	replace *Nsi*I-*Kpn*I fragment in pJG51 (Val^{S}) with same fragment from pJG16, *ilvGMEDAYC* (Val^{R})	this work
pUC 9	cloning vector	Vierra and Messing, 1982
pMS147	1.4 kb *Hind*III-*Sma*I fragment including ORFI coding region cloned into pUC 9	this work
pCC2	pMS147 with *Cla*I deletion in ORFI	this work

TABLE III

Oligonucleotides Used in This Study

oligonucleotide[1]	sequence[2]	sequence coordinates[3]
19	5'-CTGcagtgggtggtacatgcgtt-3'	283-302 in *ilvG*
20	5'-CTGcagctcttgcacattcatca-3'	complementary to 1565-1584 in *ilvG*
40	5'-acgcaagcagttaccgttga-3'	1596-1615 in *ilvG*
6	5'-ggttgtgctctggcagatggcaa-3'	complementary to 2131-2153 in *ilvM*
41	5'-ttatatctctgaaggcgcag-3'	2766-2758 in *ilvE*
7 (oli E)	5'-tgatctaaccagccccatttatc-3'	complementary to 3091-3113 in *ilvE*
D5	5'-gggtgaacgagttcacaaccgcg-3'	complementary to 3301-3323 in *ilvD*
47 (oli D)	5'-agcataacgtcgtattgttccagcg tttgcggcaacgtcaggccaagtac-3'	complementary to 4244-4293 in *ilvD*
17	5'-GAattcagttacaggtaagcgat-3'	4856-4876 in *ilvD*
18	5'-gaattcgaagctgtagaggcgtt-3'	complementary to 6341-6363 in *ilvA*
38 (oli A)	5'-tatcgacggtgatgatgtcgtcgag atactcctggcataaacggaaggtt-3'	complementary to 5793-5842 in *ilvA*
V	5'-tatggtttccctgataatgtccggtagggctgggattcat-3'	complementary to 2825-2864 in ORFI
D	5'-gctgcgatattaagtctggtaaa-3'	2531-2553 in ORFIII

---

1. Oligonucleotides were synthesized and purified as described in Materials and Methods.
2. Uppercase letters indicate residues not in the *ilvGMEDA* sequence which have been included to facilitate molecular cloning.
3. Sequence coordinates are numbered as the +1 residue of the *ilvGp2* promoter is residue 1 in *ilvGMEDA* sequence and as the same as in Coppola *et al.*, 1991, for the gene encoded the *ilv*-related protein.

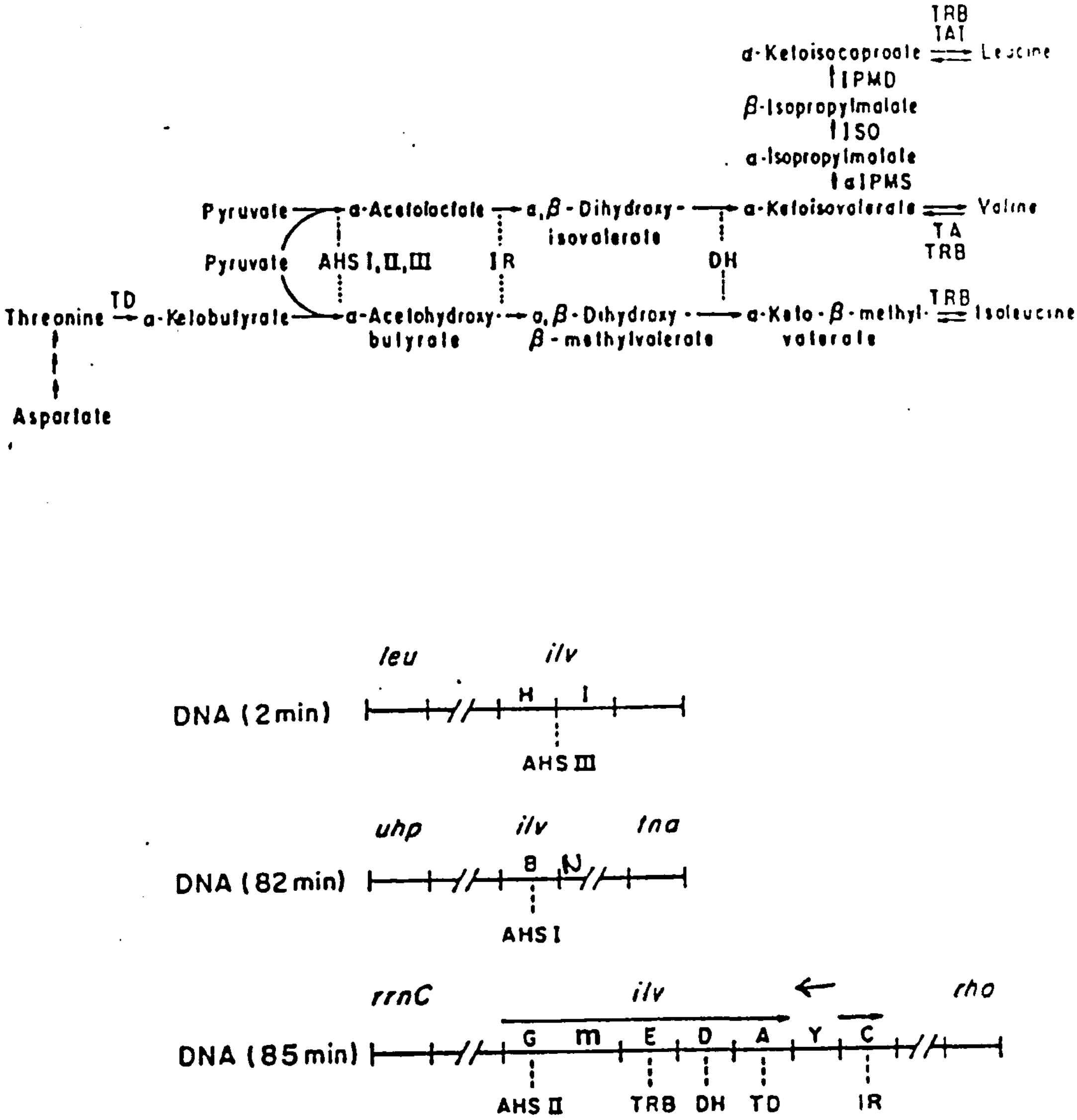


Fig. 1. Biosynthetic pathways leading to the branched chain amino acids and the locations of the various *ilv* genes. The corresponding structural genes are indicated as follows: TD(*ilvA*), threonine deaminase; AHSI (*ilvBN*) and AHSIII (*ilvIH*), valine sensitive acetohydroxy acid synthases; AHSII (*ilvGM*), valine resistant acetohydroxy acid synthase; IR (*ilvC*), acetohydroxy acid isomero-reductase; DH (*ilvD*), dihydroxy acid dehydrase; TrB (*ilvE*), transaminase B; v (*ilvY*) is a positive control element required for substrate induction of the *ilvC* gene (Baez *et al.*, 1979).

BRANCH POINT ENZYME		BRANCHLET	PRIMARY EFFECTORS	REMOTE EFFECTORS
L-Threonine Deaminase (TD)		ILE	ILE	LEU, VAL
	I	VAL	(I) VAL	
α-Acetohydroxyacid Synthase (AHS)	II	VAL	(II) ?(VAL^R)	LEU, ILE
	III	VAL	(III) VAL	
α-Isopropylmalate Synthase (α-IPMS)		LEU	LEU	ILE

Fig. 2. The primary and remote effectors for the feedback regulated enzymes in each branchlet of the pathways (Hahn and Calhoun, 1978).

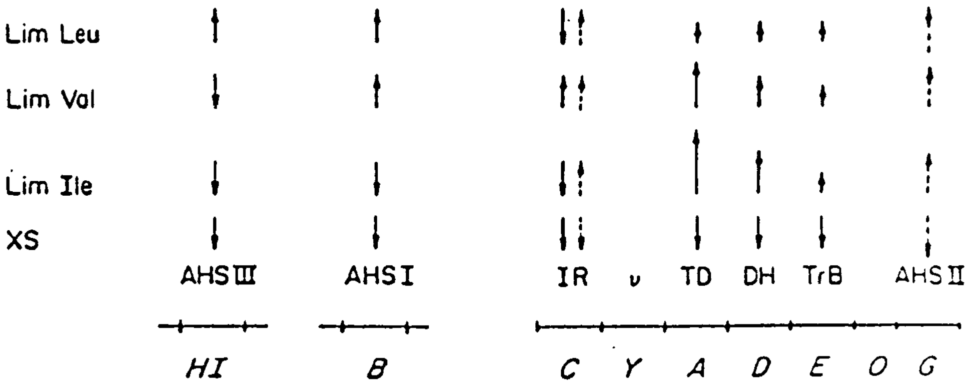
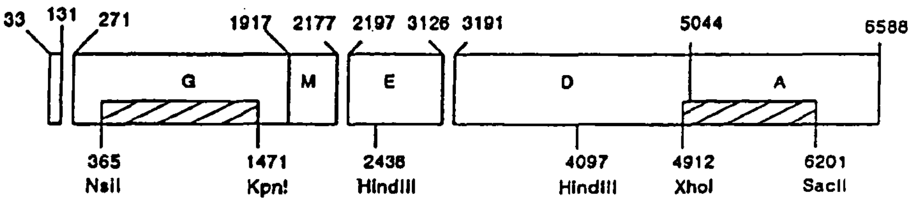


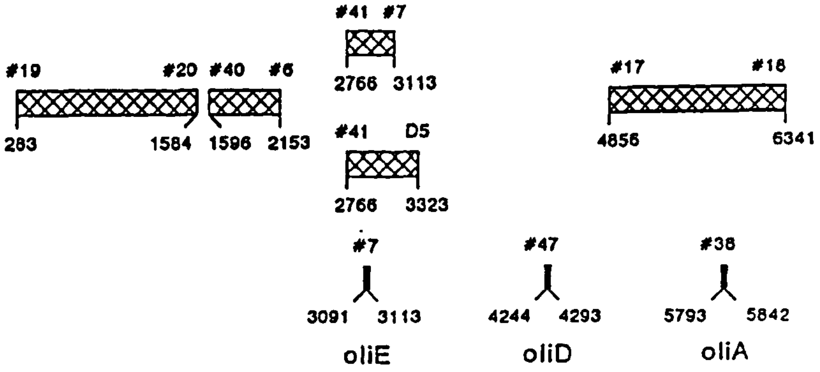
Fig. 3. Pattern of multivalent repression in isoleucine and valine biosynthesis. The arrows pointing upward indicate derepression under the indicated condition (limiting leucine, limiting valine, or limiting isoleucine). The arrows pointing downward indicate repression under the indicated condition. XS, excess branched chain amino acids in medium. Broken lines represent behavior in Val^R derivatives of *E. coli* K-12. Size of the upward-pointing arrow is a qualitative indicator of the derepression, relative to that of transaminase B, obtained upon growing cells under various derepression conditions . The *ilvO* site is now recognized to be the naturally occurring frameshift in the wild type *ilvG* gene. Also, we now know that AHSI has a small subunit encoded by *ilv B* and a large subunit encoded by *ilvN*, and AHSII has a small subunit encoded by *ilvM* and a large subunit encoded by *ilvG* (Umbarger, 1987).

Fig. 4. The *ilvGMEDA* cluster, probes used in Northern blots, and major transcripts detected in this study. Panel A indicates the genes, locations of the coding regions for each polypeptide, locations of selected restriction enzyme sites, and deletions (hatched) introduced in *ilvG* to construct plasmid pCC17 and in *ilvDA* to construct plasmid pCC20. Panel B indicates the PCR products and oligonucleotides used as probes in Northern blots. The polymerase chain reaction products (hatched) were generated using oligonucleotides #19 and #20 (pcrG), #40 and #6 (pcrGM), #41 and #7 (pcrE), #41 and D5 (pcrED), and #17 and #18 (pcrDA). Synthetic oligonucleotides used as probes were #7 (oliE), #47 (oliD), and #38 (oliA). Oligonucleotide D5 was also used as a primer for reverse transcriptase (Fig. 10). Panel C indicates the relatively unstable (dotted lines) 6.7 and 3.6 kb transcripts, and the relatively stable (bold line) 4.6 and 1.1 kb transcripts.

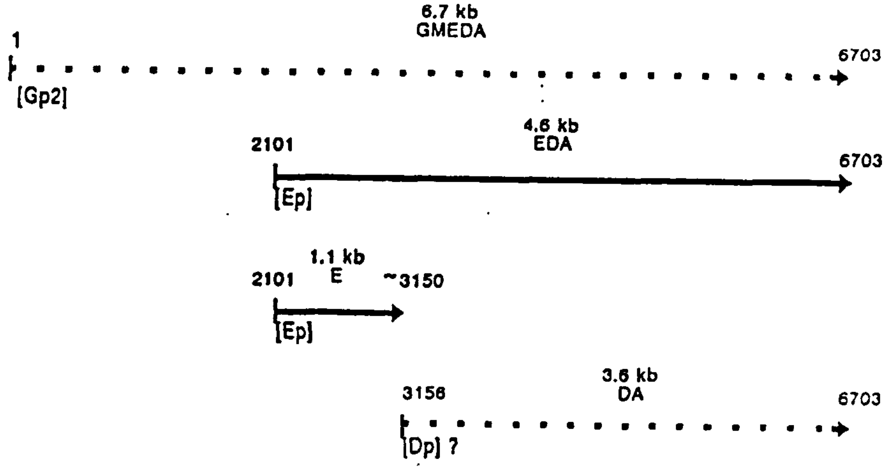
A. Genes, Deletions, and Restriction Enzyme Sites



B. PCR Products and Oligonucleotides



C. Transcripts



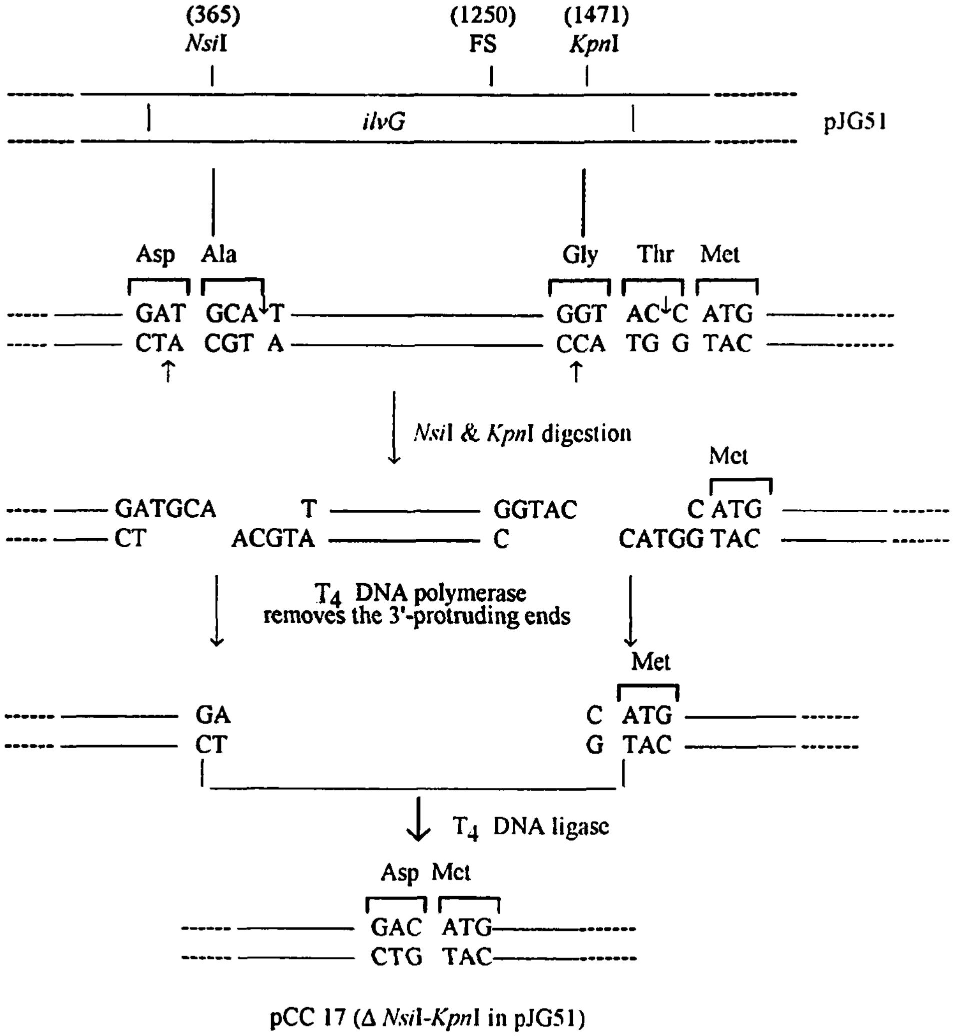


Fig. 5. Construction of plasmid pCC17 containing a deletion in the *ilvG* gene. Plasmid pJG51 (Val^S, has a frameshift site in *ilvG* gene), which has unique sites for *Nsi*I and *Kpn*I separated by 1.110 bp in the *ilvG* gene, was treated with both enzymes, followed by T_4 DNA polymerase to remove the protruding 3' ends, and T_4 DNA ligase to join the blunt ends. The deletion of 1110 bp in the resulting plasmid, pCC17, generated a protein fusion within the *ilvG* gene, so that translation proceeds without the formation of a polar frameshift site.

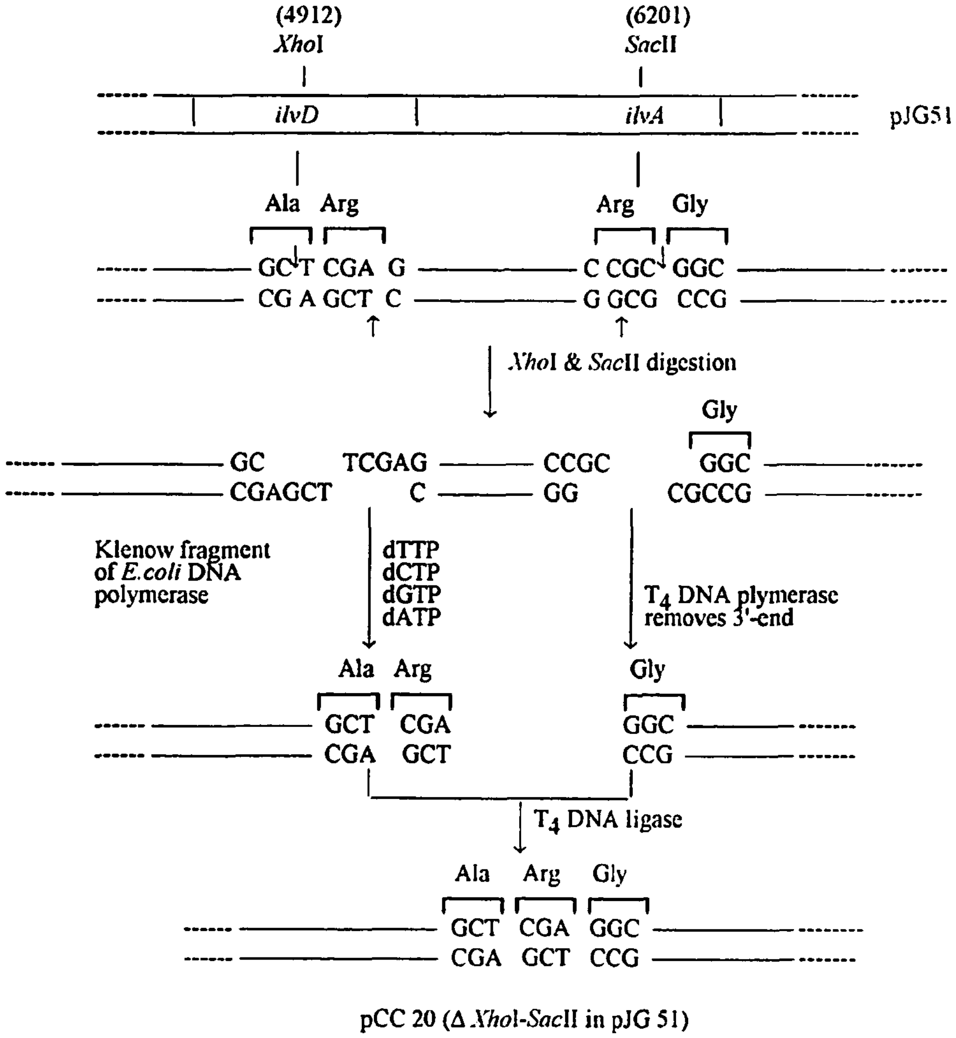


Fig. 6. Construction of plasmid pCC20 containing a deletion in *ilvDA* genes. Plasmid pJG51 was treated with *Xho*I, which cleaves in the *ilvD* gene, and *Sac*II, which cleaves in the *ilvA* gene. Phage T_4 DNA polymerase was used to remove two overhanging bases at 3' end of the *Sac*II site. Klenow fragment was used to fill in four bases at the *Xho*I site, and then T_4 DNA ligase was used to join the blunt ends. The deletion of 1290 bp generates a protein fusion with the amino terminal region coded by *ilvD* and the carboxyl terminal region coded by *ilvA*.

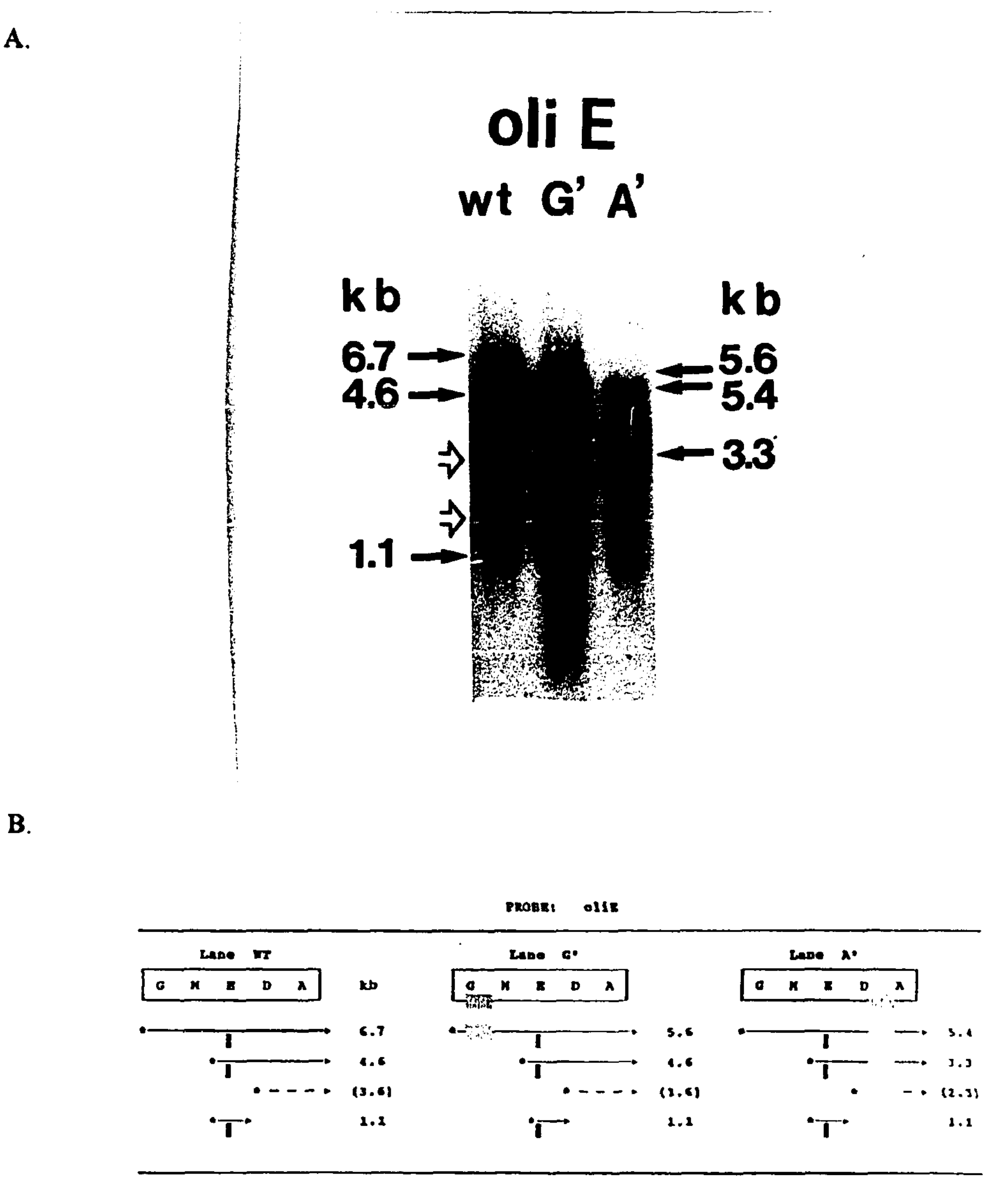


Fig. 7. (A) Detection of *ilv* transcripts using probe oli E. RNA was extracted from strain C600 derivatives containing plasmid pJG51 which contains the full length *ilvGMEDA* cluster (lane wt), plasmid pCC17, which contains a deletion in the *ilvG* gene (lane G'), or plasmid pCC20, which contains a deletion in the *ilvDA* region (lane A'), and hybridized to probe oliE. Solid arrows indicate the sizes of the full length transcripts (left) or shortened transcripts (right) generated as a result of the deletions of plasmid pCC17 and pCC20. Open arrows (left) indicate the positions of the large (2,904 bp) and small (1,541 bp) rRNA bands. (B) Approximate positions of the deletions (▨), PCR (■) and synthetic (▌)hybridization probe, and transcript (•——→), indicating approximate sizes in kilobases. The dotted arrow with parentheses around the size in kb indicates a transcript not detected in this blot.

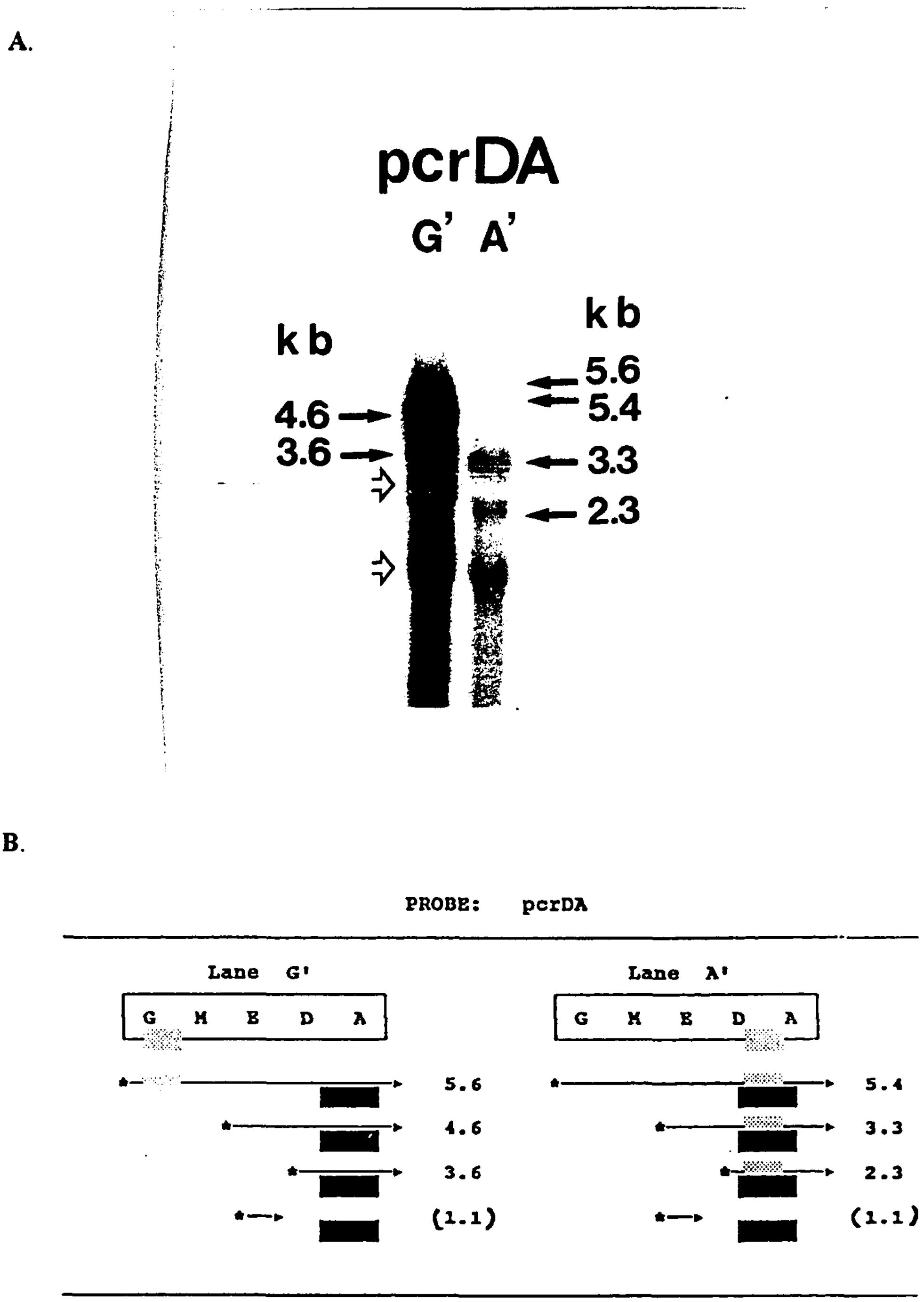


Fig. 8. Detection of *ilv* transcripts using probe pcrDA. RNA was isolated from strain C600 derivatives containing plasmids pCC17 (lane G') or pCC20 (lane A'), and hybridized to an asymmetric PCR probe derived from the *ilvDA* junction (pcrDA). (A) Northern blot. (B) Approximate positions of deletions, probe, and transcripts. Symbols are described as in Fig. 7.

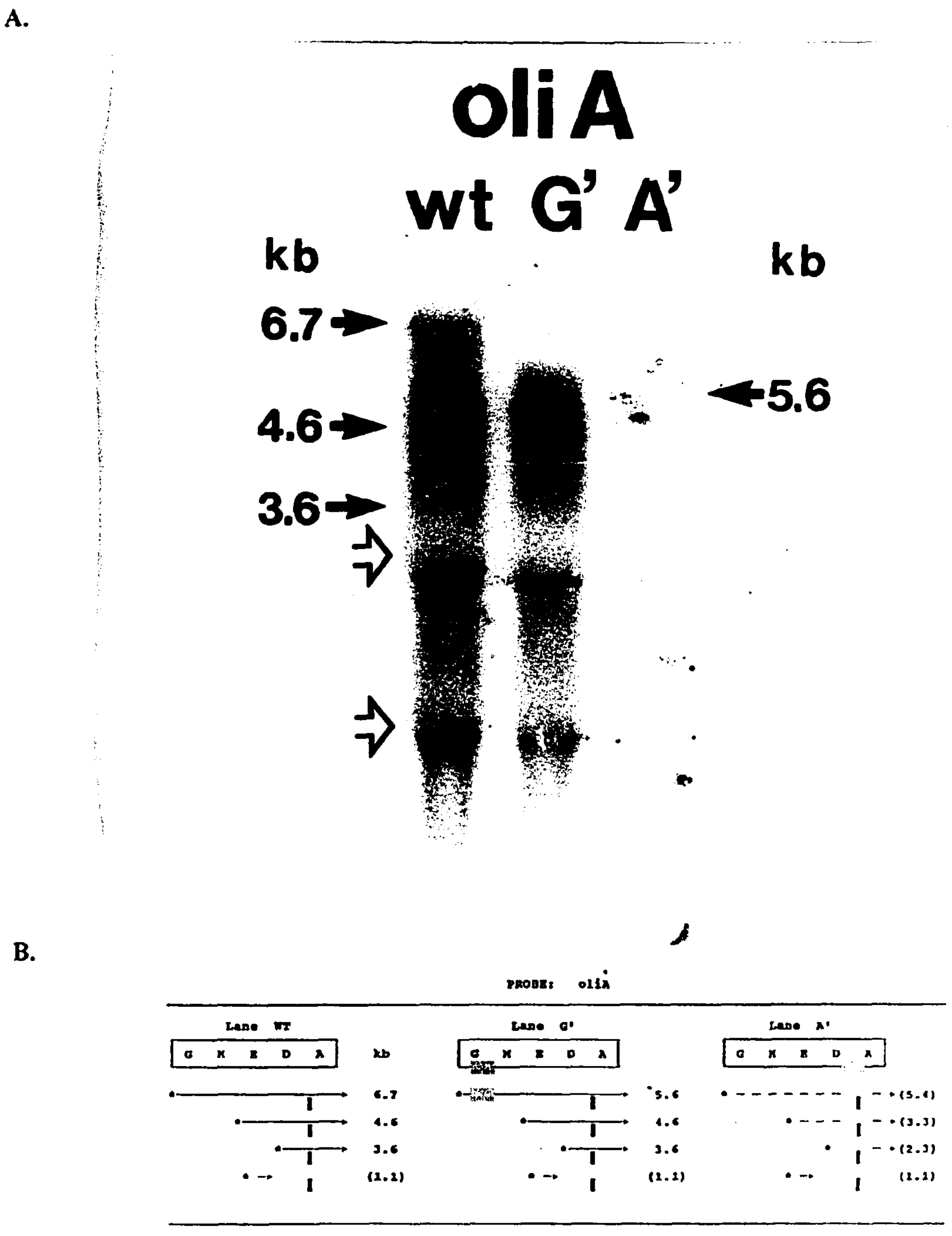


Fig. 9. Detection of *ilv* transcripts using probe oliA. RNA was isolated from strain C600 derivatives containing plasmids pJG51 (lane wt), pCC17 (lane G'), or pCC20 (lane A'), and hybridized to a synthetic oligonucleotide derived from the *ilvA* gene (oliA). (A) Northern blot. (B) Approximate positions of deletions, probe, and transcripts. Symbols are described as in Fig. 7.

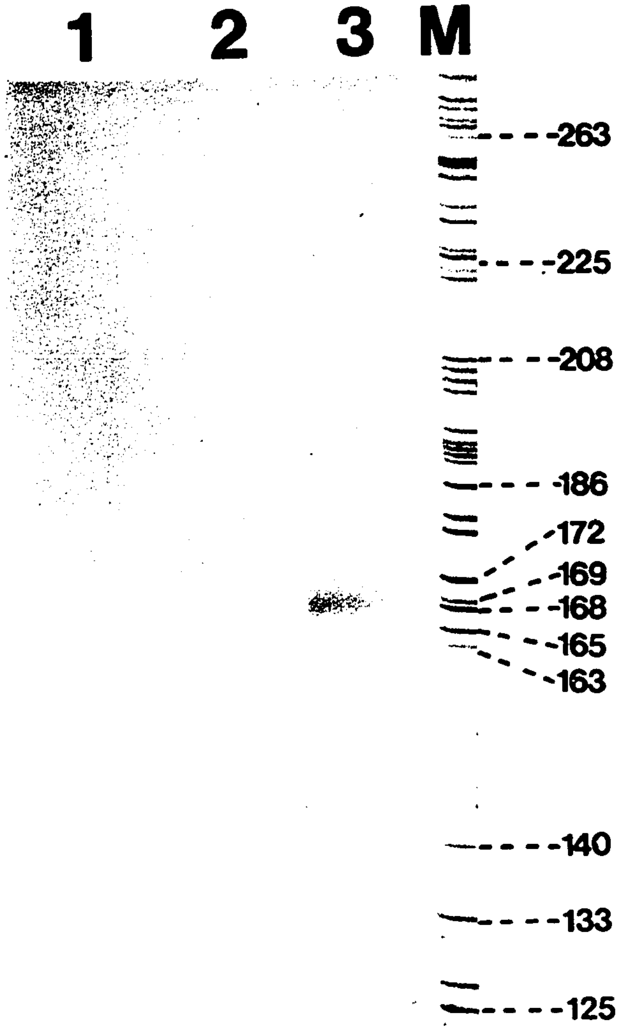
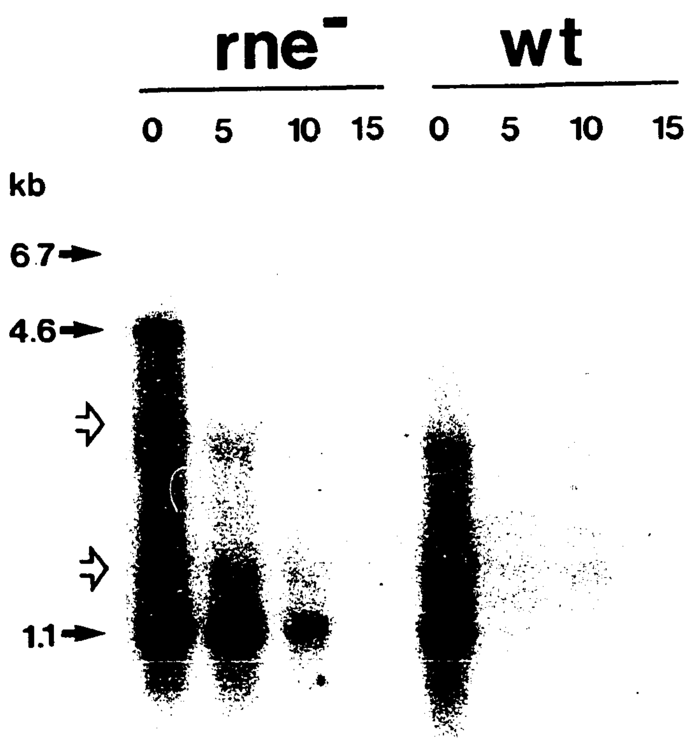


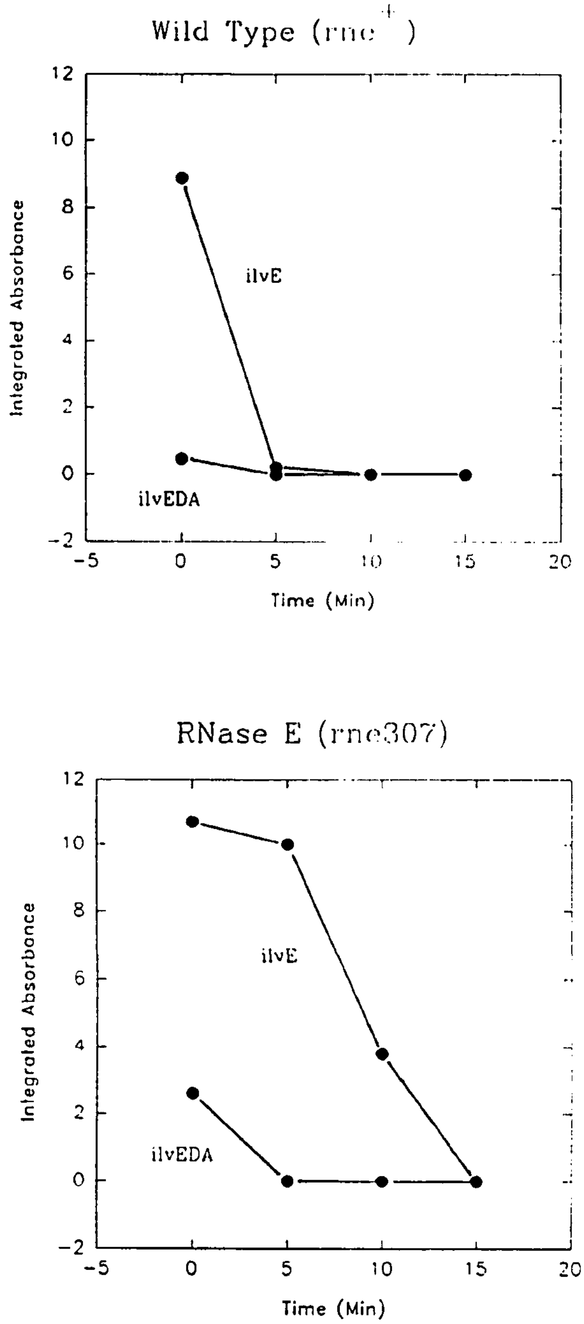
Fig. 10. Identification of 5'-end of *ilvDA* transcript by primer extension. Total *E. coli* mRNA was used as template for reverse transcriptase and synthetic oligonucleotide D5 derived from sequences near the 5'-end of the *ilvD* gene as primer. The RNA was from strain FD1054 that contains an *ilvGMEDA* deletion (lane 1), or from a strain C600 derivative containing plasmid pJG51 (lanes 2 and 3). Lane 2 is a control in which reverse transcriptase was omitted. Lane M contains molecular size markers as indicated. As a size standard, wild type M13mp18 was used as template with the universal primer for sequencing reaction. Only G lane is shown.

Fig. 11. Effect of RNase E upon *ilvGMEDA* transcripts. (A) Strains N3431 (*rne-3071*) and N3433 (*rne*$^+$) were transformed with plasmid pJG51 and used to examine transcript stability. Cells were grown at 30°C in minimal medium with thiamine to mid-log phase, then shifted to 43°C for 10 min. Rifampicin was added and RNA samples were prepared 0, 5, 10, and 15 min later, as indicated, and analyzed using probe oliE. (B) Relative amounts of *ilv* mRNA in the samples of panel A at different times after rifampicin addition were quantited using a Millipore BioImage system. The values obtained were plotted in the integrated absorbance versus time of sampling.

A.



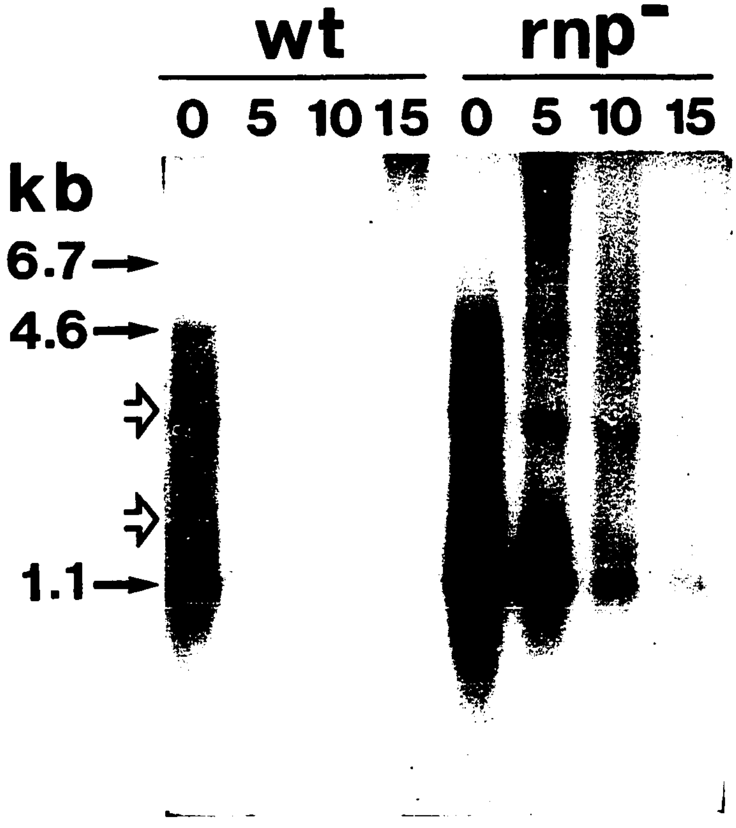
B.


Wild Type (rne^{+})

RNase E (rne307)

Fig 12. Effect of RNase P upon the *ilvGMEDA* transcripts. (A) Strains NHY312 (*rnpA*$^+$) and NHY322 (*rnpA49*) were transformed with plasmid pJG51 and used to examine transcript stability. Cells were grown at 30°C in minimal medium supplemented with thiamine and proline to mid-log phase, then shifted to 43°C for 10 min. Rifampicin was added and RNA samples were prepared 0, 5, 10, and 15 min later, as indicated, and analyzed using probe oliE. (B) Autoradiogram in panel A was analyzed as in Fig. 11B.

A.



B.

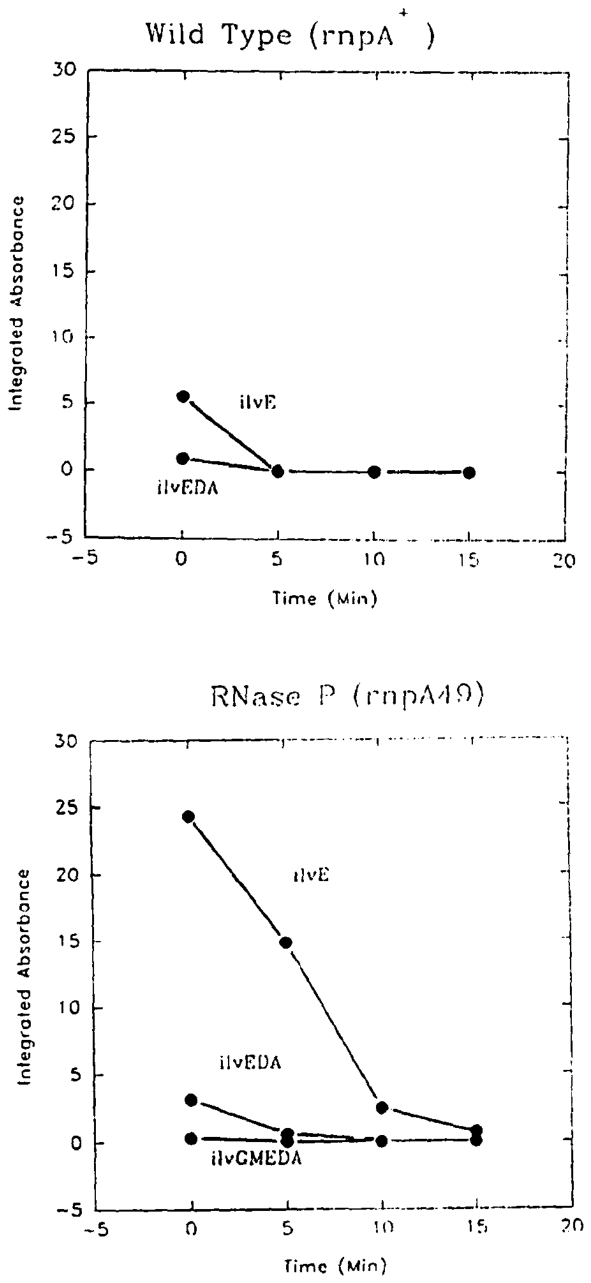
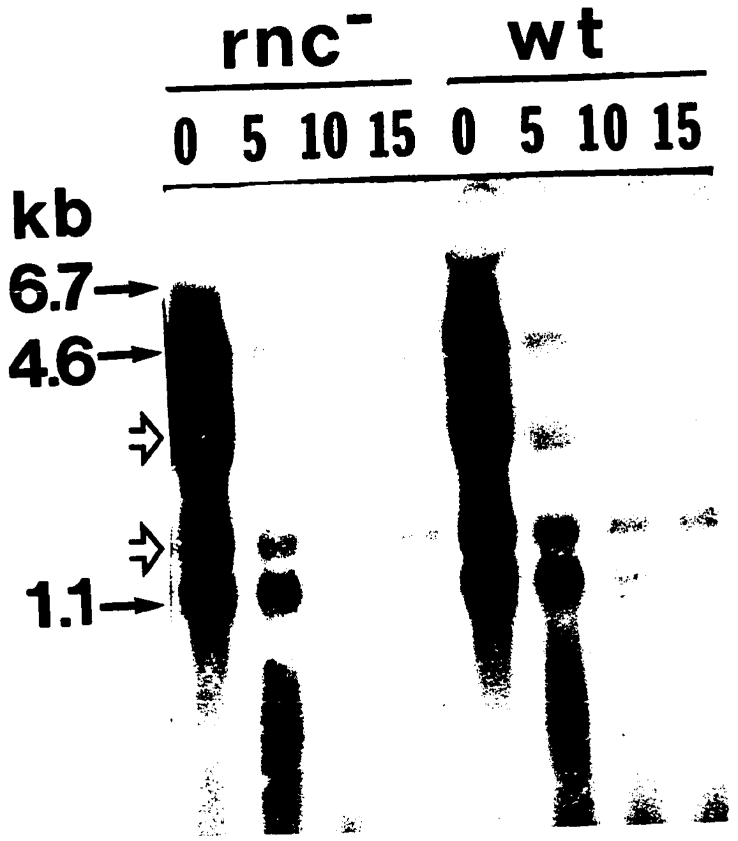
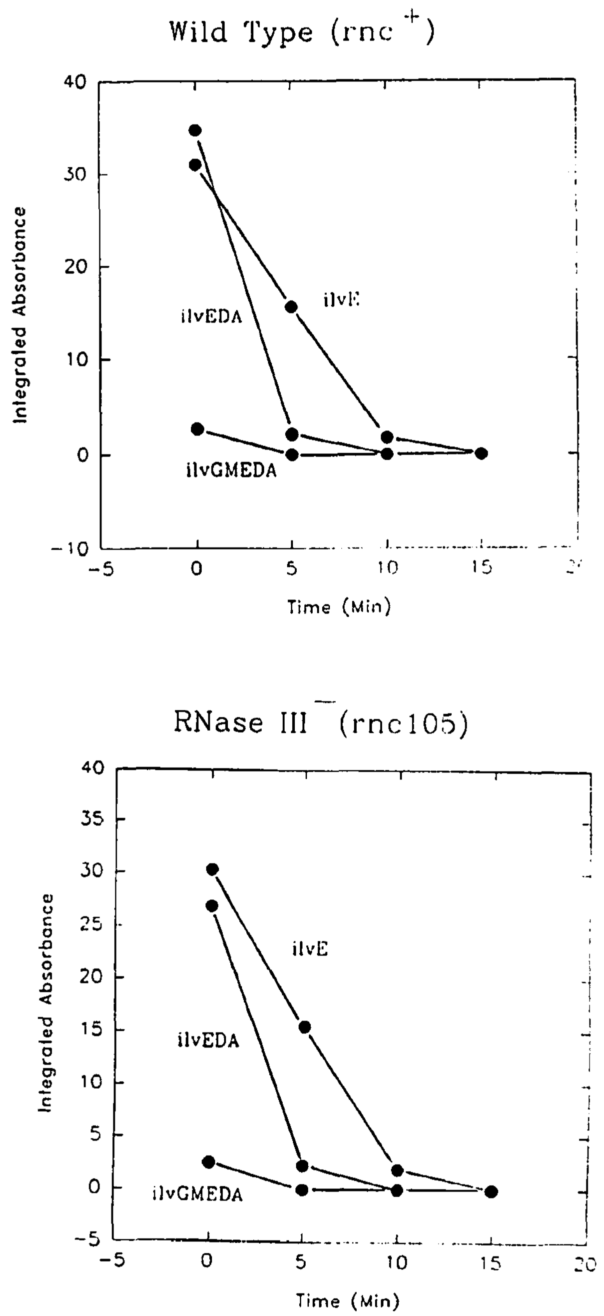

Wild Type (rnpA+)
Integrated Absorbance
30
25
20
15
10
5
0
-5
ilvE
ilvEDA
-5
0
5
10
15
20
Time (Min)
RNase P (rnpA49)
Integrated Absorbance
30
25
20
15
10
5
0
-5
ilvE
ilvEDA
ilvGMEDA
-5
0
5
10
15
20
Time (Min)

Fig. 13. Effect of RNase III upon the *ilvGMEDA* transcripts. (A) Strain BL322 (*rnc*$^+$) and BL321 (*rnc-105*) were transformed with plasmid pJG51 and used to examine transcript stability, cells were grown at 37°C in minimal medium supplemented with thiamine and arginine to mid-log phase. Rifampicin was added and RNA sample were prepared 0, 5, 10, and 15 min later, as indicated, and analyzed using probe oliE. (B) Autoradiogram in panel A was analyzed as in Fig. 11B.

A.



B.

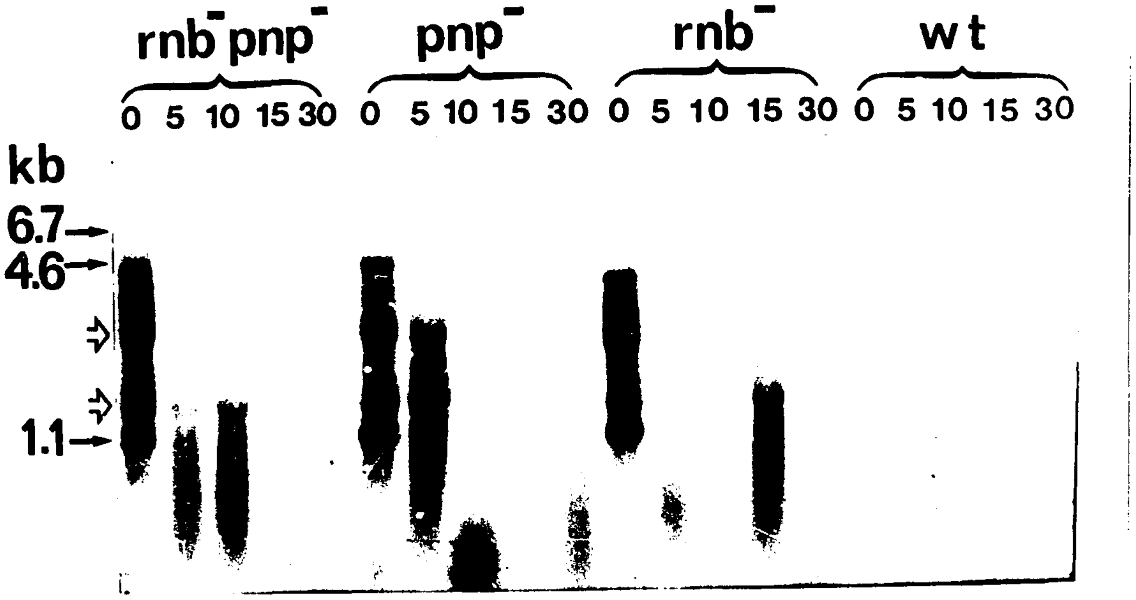


Wild Type (rnc$^+$)

RNase III$^-$ (rnc105)

Fig. 14. Effect of RNase II and PNPase upon the *ilvGMEDA* transcripts. (A) Strains SK5003 (*pnp-7, rnb-500*^{ts}), SK5004 (*pnp-7, rnb*$^{+}$), SK5005 (*pnp*$^{+}$, *rnb-500*^{ts}), and SK5006 (*pnp*$^{+}$, *rnb*$^{+}$) were transformed with plasmid pJG51 and used to examine transcript stability. Cells were grown at 30°C in minimal medium supplemented with leucine and threonine to mid-log phase, then shifted to 44°C for 10 min. Rifampicin was added and RNA were prepared 0, 5, 10, 15, and 30 min later, and analyzed using probe oliE. (B) Autoradiogram in panel A was analyzed as in Fig. 11B.

A.



B.

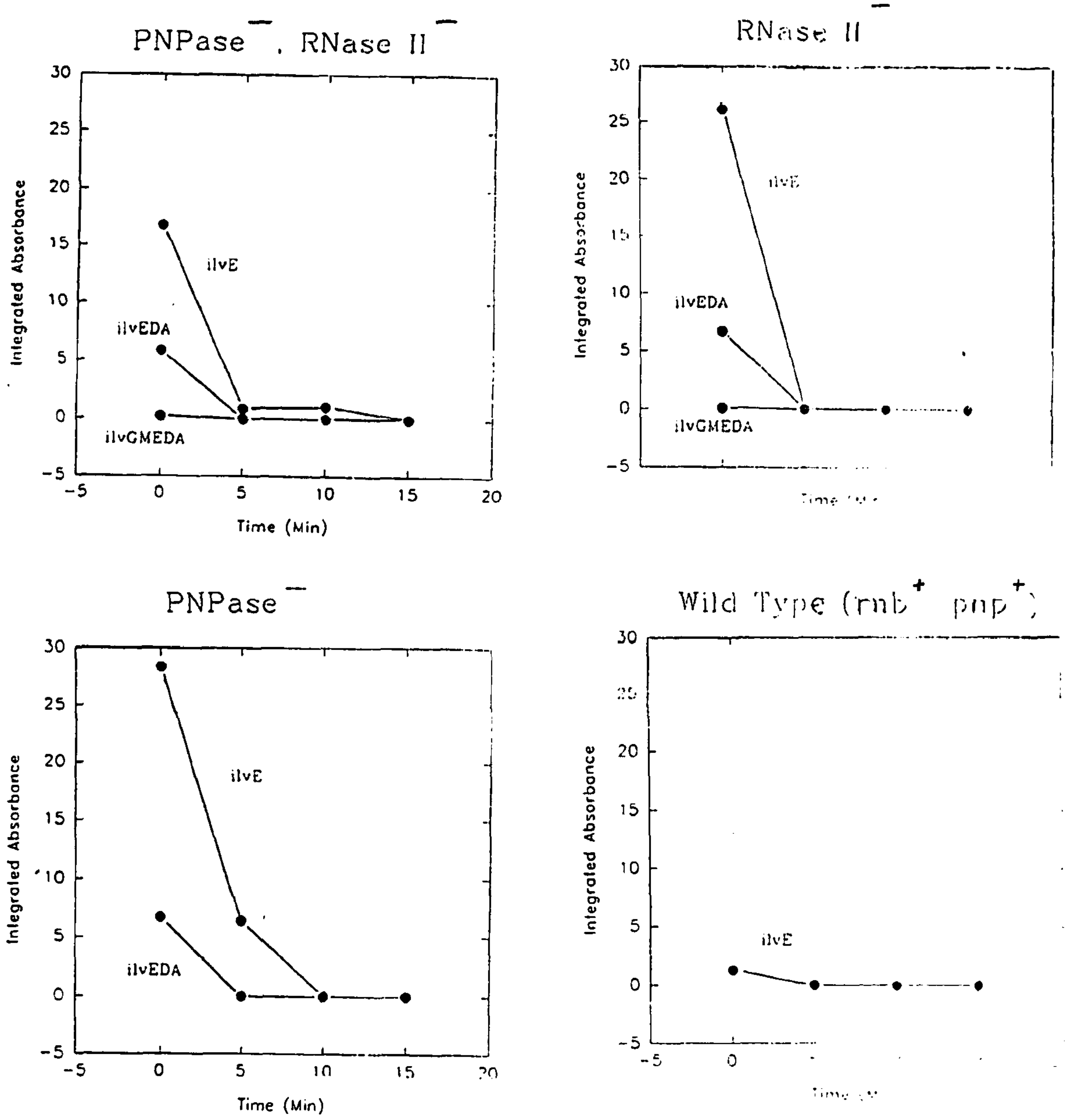
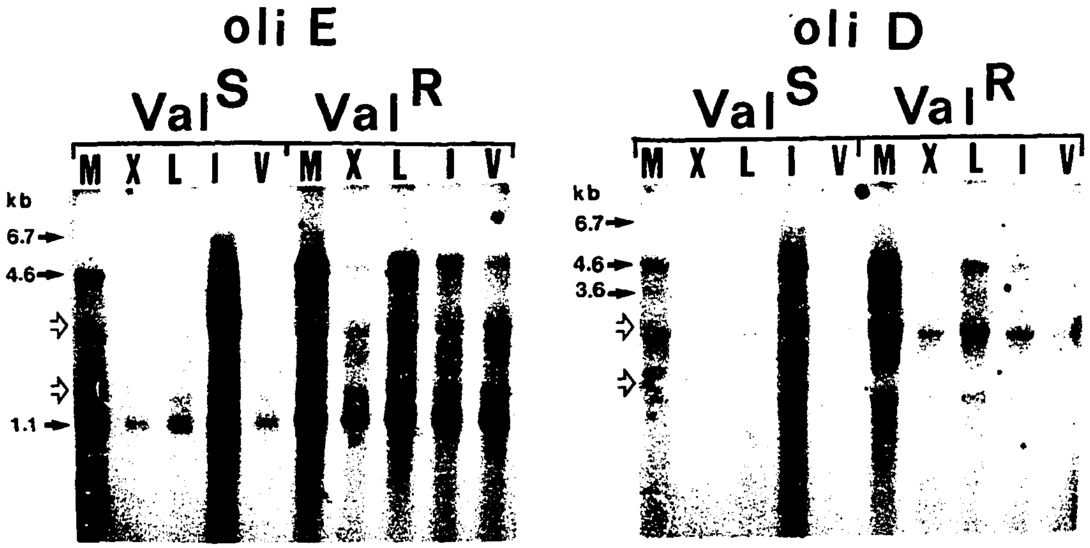
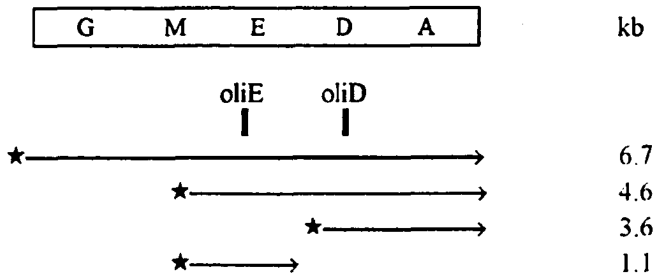

PNPase⁻, RNase II⁻
Integrated Absorbance
Time (Min)
ilvE
ilvEDA
ilvGMEDA
RNase II⁻
Integrated Absorbance
ilvE
ilvEDA
ilvGMEDA
PNPase⁻
Integrated Absorbance
Time (Min)
ilvE
ilvEDA
Wild Type (rnb⁺ pnp⁺)
Integrated Absorbance
ilvE
30
25
20
15
10
5
0
-5
0
5
10
15
20

Fig. 15. Effect of end product amino acids upon the *ilvGMEDA* transcripts. (A) RNA was isolated from derivatives of strain FD1054 containing plasmid pJG51 (Val^S) or pCC36 (Val^R) and hybridized to probe oliE (left) or oliD (right). Cells were grown in unsupplemented minimal medium (lane M), in the presence of excess (lane X) leucine, isoleucine, and valine, or with limiting concentrations of leucine, isoleucine, or valine (lanes L, I, and V, respectively). Symbols are described in Fig. 7. (B) Probes used and transcripts detected by each probe. (C) Comparison of four *ilv* transcripts at different growth conditions. Relative amounts of *ilv* mRNA were analyzed as in Fig. 11B.

A.



B.



C.

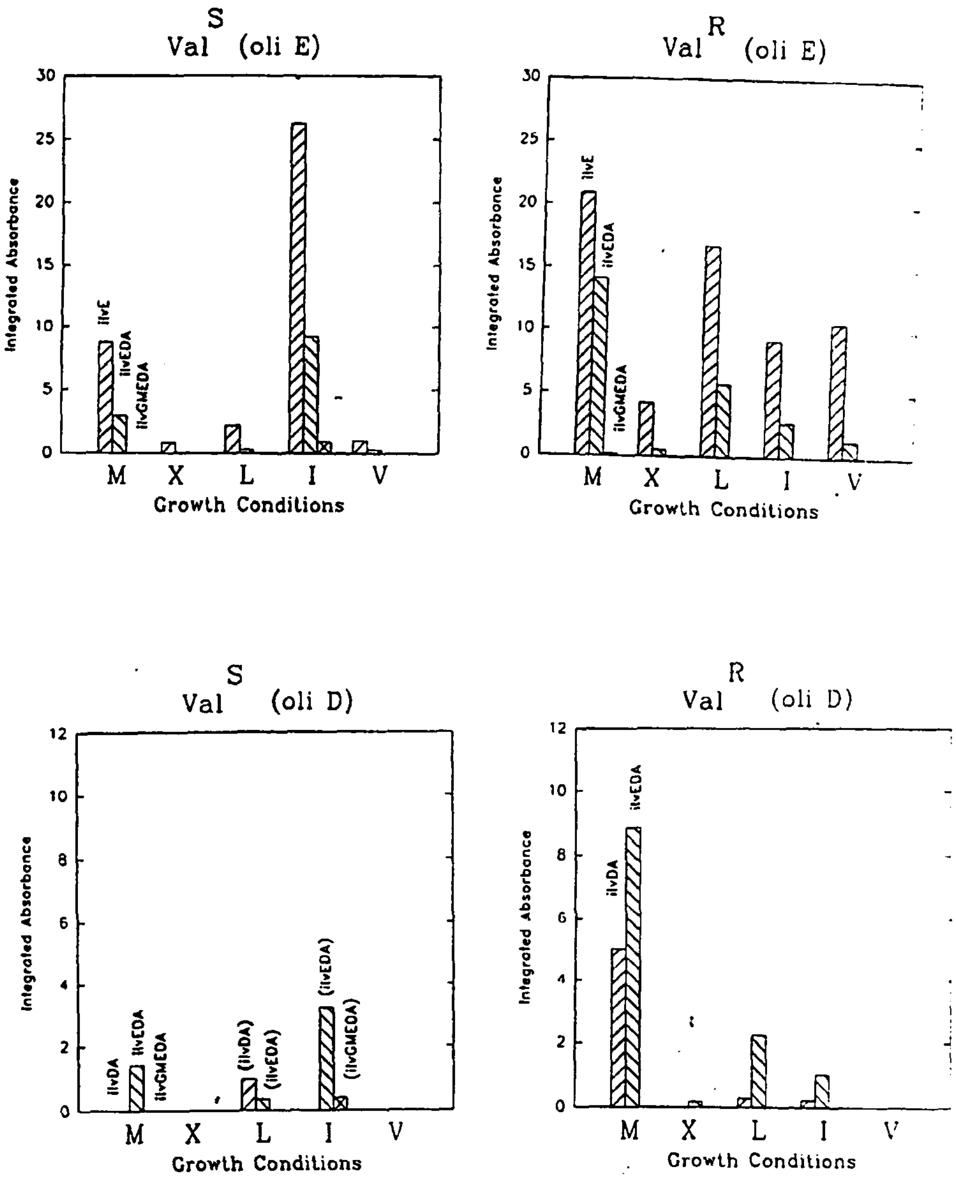
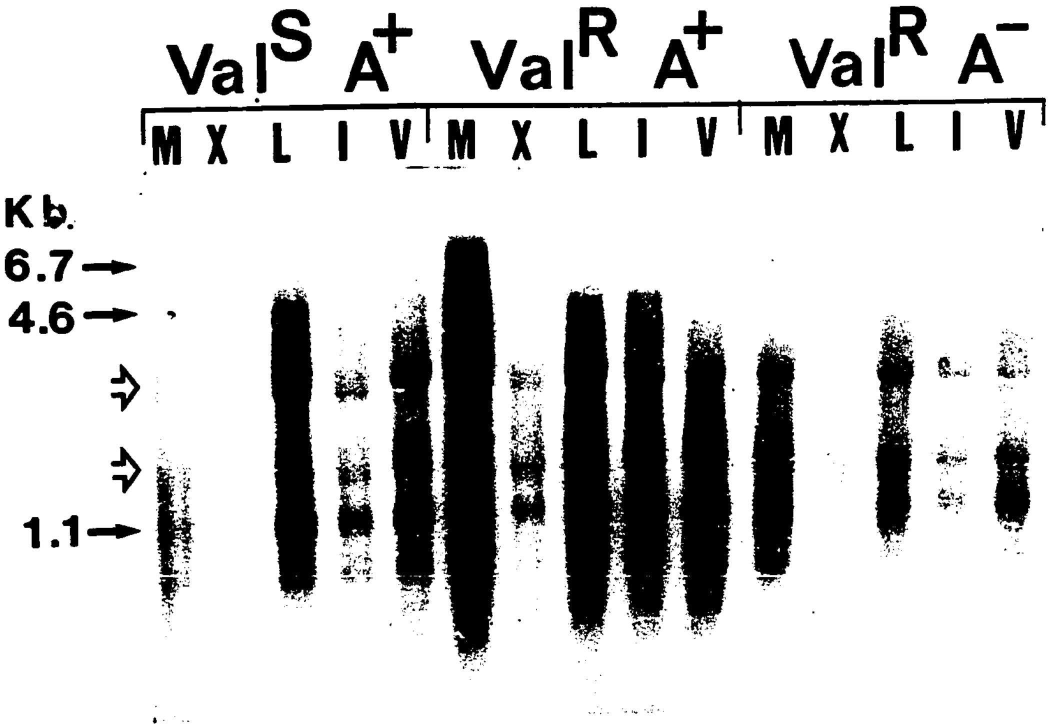

Val^S (oli E)
Integrated Absorbance
30
25
20
15
10
5
0
ilvE
ilvEDA
ilvGMEDA
M X L I V
Growth Conditions
Val^R (oli E)
Integrated Absorbance
ilvE
ilvEDA
ilvGMEDA
M X L I V
Growth Conditions
Val^S (oli D)
Integrated Absorbance
12
10
8
6
4
2
0
ilvDA
ilvEDA
ilvGMEDA
(ilvDA)
(ilvEDA)
(ilvEDA)
(ilvGMEDA)
M X L I V
Growth Conditions
Val^R (oli D)
Integrated Absorbance
ilvDA
ilvEDA
M X L I V
Growth Conditions

Fig. 16. Effects of the *ilvA538* regulatory allele and the *ilvG*$^-$ (Val^R) allele upon the *ilvGMEDA* transcripts. (A) Strain CU505 (*ΔilvGMEDAYC2014, leu*$^-$) were transformed with plasmids pJG51 (Val^S, *ilvA*$^+$), pCC40 (Val^R, *ilvA*$^+$), and pJG45 (Val^R, *ilvA538*). Cells were grown in unsupplemented minimal medium (lane M), in the presence of excess (lane X) leucine, isoleucine, and valine, or with limiting concentrations of leucine, isoleucine, or valine (lane L, I, and V, respectively). RNA samples were analyzed using probe oliE. (B) Autoradiogram in panel A was analyzed as in Fig. 11B.

A.



B.

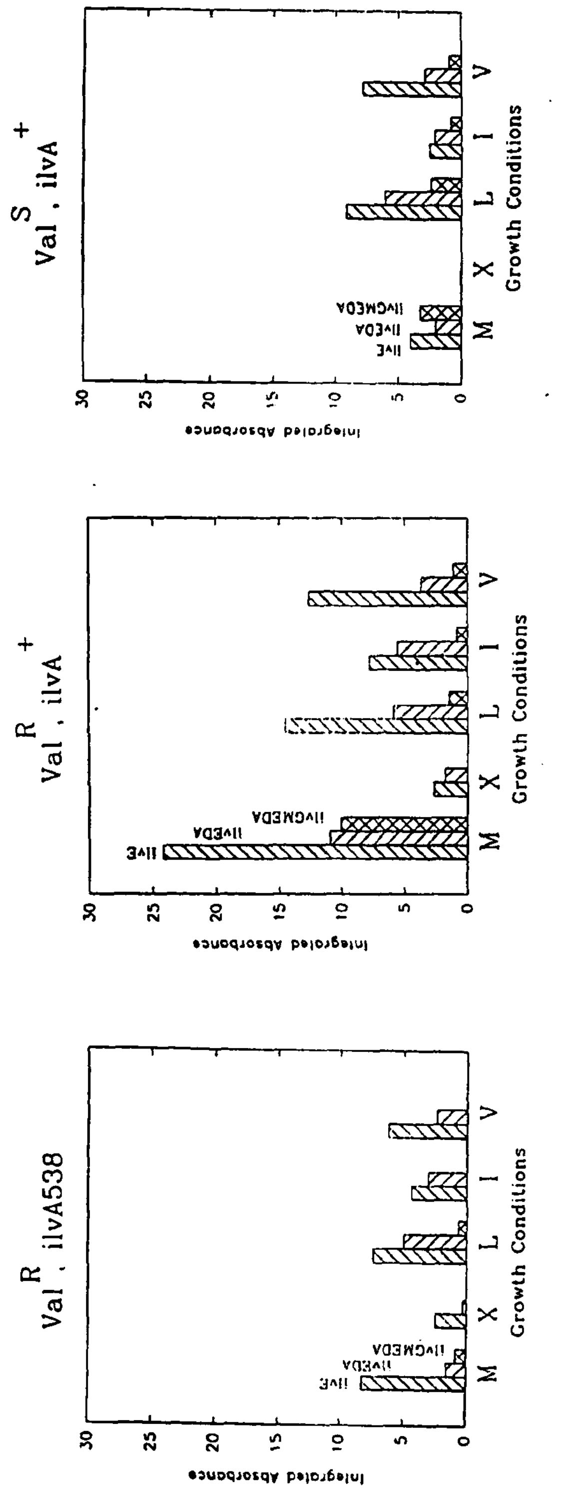
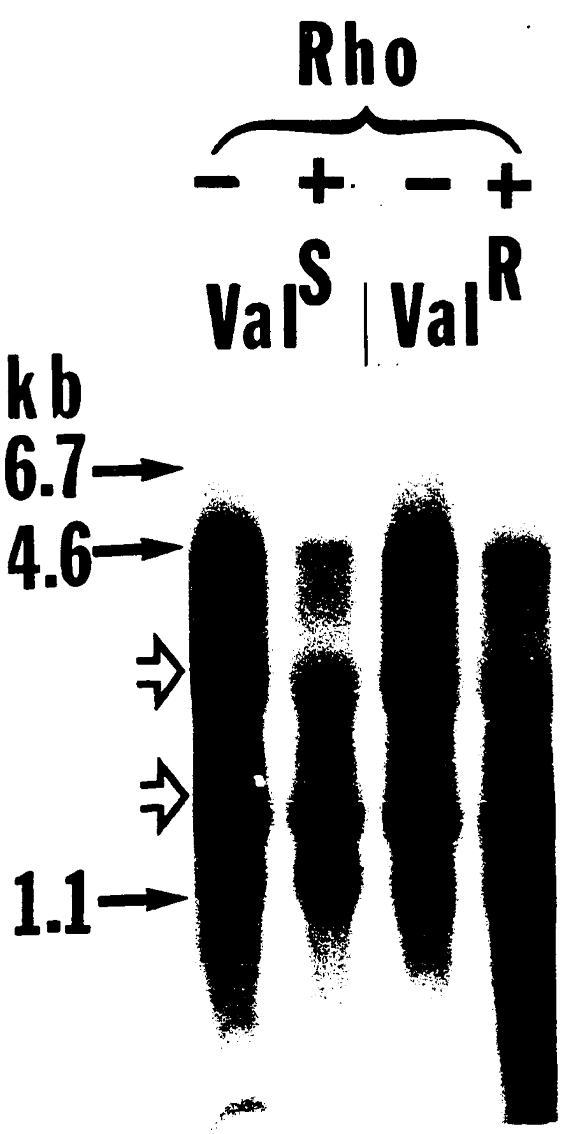

Val^S, ilvA^+
Val^R, ilvA^+
Val^R, ilvA538
Integrated Absorbance
0
5
10
15
20
25
30
ilvE
ilvEDA
ilvGMEDA
M
X
L
I
V
Growth Conditions

Fig. 17. Effect of the *rho-115* allele upon the *ilvGMEDA* transcripts. (A) Strains TN1276 (*rho*$^+$) and TN1271 (*rho-115*) were transformed with plasmids pJG51 (Val^S) and pCC40 (Val^R). Cells were grown in minimal medium supplemented with leucine. RNA samples were prepared and analyzed using probe oliE. (B) Autoradiogram in panel A was analyzed as in Fig. 11B. The asterisks in the diagram of Fig. 17B indicate the three Rho dependent polarity sites in the *ilvGM* region of the Val^S strain, and the single postulated Rho dependent site in both the Val^R and Val^S strains between *ilvE* and *ilvD*. The data presented here are average values of two independent experiments and three different times of exposure to X-ray film.

A.



B.

ilvEDA (Rho$^-$/Rho$^+$)	3.4	2.3.
ilvE (Rho$^+$/Rho$^-$)	1.2	3.2

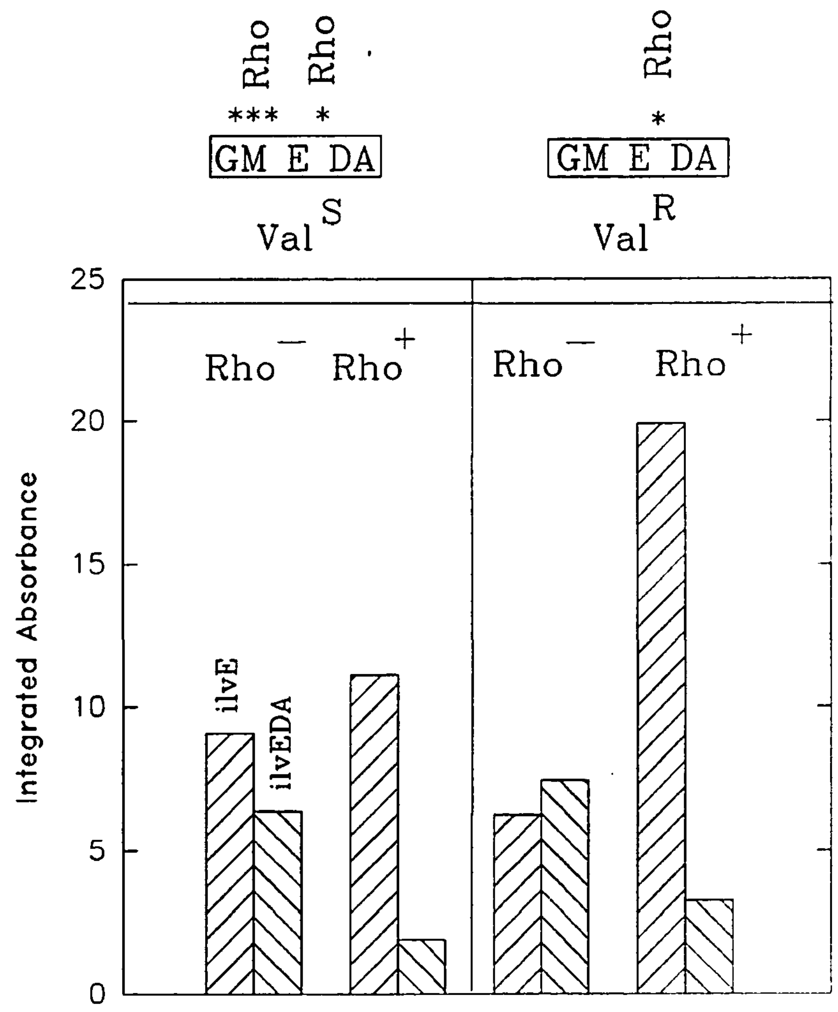
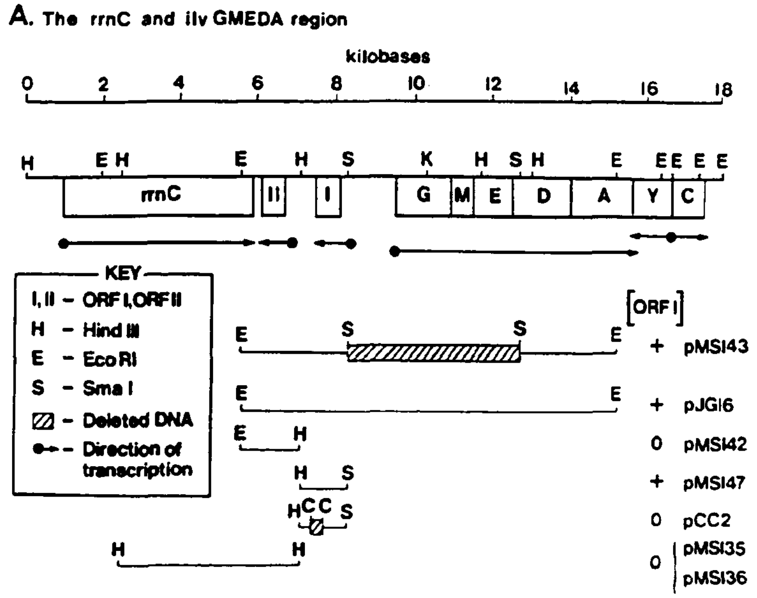
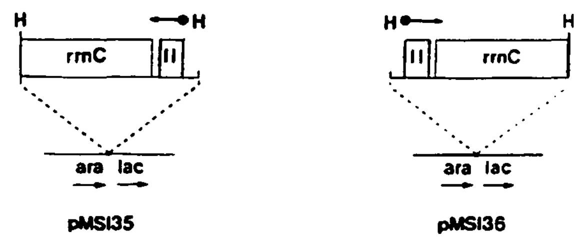


Fig. 18. The *rrnC* and *ilvGMEDA* region of the *E. coli* K-12 chromosome. 18-kb segment of the *E. coli* K-12 chromosome containing the *rrnC* and *ilvGMEDAYC* genes that flank the *ilv*-related gene encoding ORFI and the gene encoding ORFII, both have unknown functions. For clarity, ORFIII (Fig. 19) is not shown. The locations of selected restriction enzyme sites are indicated. Plasmids containing the indicated restriction fragments were tested in maxicells for the production of the *ilv*-related ORFI peptide. The presence or absence of ORFI peptide in maxicell analysis of plasmid shown is indicated by + or 0, respectively, in the column under [ORFI], plasmid pCC2 is a derivative of plasmid pMS147 that contains a deletion of a *Cla*I fragment containing the C-terminal coding region of ORFI, and this results in a decreased size of the ORFI peptide seen in maxicell analysis (Fig. 20) and a decreased size of mRNA seen in Northern blots (Fig. 21). The region upstream of *ilvG* is complex and contains two promoters directed towards *ilvG*, at positions designated +1 for *ilvGp2* and -72 for *ilvGp1*, and a divergently transcribed promoter directed towards ORFI, designated ORF*p*, at -112. The region can be diagrammed as follows, where the direction of transcription is indicated by the arrows and the letter *p* indicates a promoter site (Coppola *et al.*, 1991).

ORFI	← ORF*p*	*ilvGp1* →	*ilvGp2* →	*ilvGMEDA*
	[-112]	[-72]	[+1]	



A. The rrnC and ilv GMEDA region



B. Plasmid pMC81 derivatives

Fig. 19. The DNA sequence of the region between *rrnC* and *ilvGMEDA*. The sequence extending from the *Kpn*I in *ilvG* to the *EcoR*I site in *rrnC* is shown. The N-terminal segment of the *ilvG* encoded peptide extends from the ATG at nt 1206 to the *Kpn*I site nt=1. The ORFII, ORFIII regions are indicated. The control region between *ilvG* and ORFIII shows the transcriptional start points for the *ilvGp1*, *ilvGp2* and ORF*p* transcripts. The two binding sites for integration host factor and the two overlapping *trp* operator consensus sites are shown. The potential ribosome binding sites (RBS) are indicated at nt 1753 for ORFIII and at nt 2812 for ORFI. The ORFI ATG start codon is at nt 2826. The potential stem-and-loop structures are indicated for the *ilvGMEDA* attenuator at nt 1297 to 1320 and for the *rrnC* terminator at nt 4703 to 4724 (Coppola *et al.*, 1991).

```
[KpnI]       10        20        30        40        50        60
              |         |         |         |         |         |
   1 ggtacctaaaccgctggaggtgatgaaatttccgggcgagtgtgggcgatgtgctgcgc
      T  G  L  G  S  S  T  I  F  N  E  P  R  T  H  A  I  H  Q  A
  61 agcccacatctggtgctgccccacatctgtggtcacgacgcaatccgcaggtttacgatc
      A  W  N  Q  H  Q  G  V  D  T  T  V  V  C  D  A  P  K  R  D
 121 cgacagttgttttaacaacaacggcgcgtagatagcgtcaccgggatggtcgtaacgcca
      S  L  Q  K  L  L  L  P  A  Y  I  A  D  G  P  H  D  Y  R  W
 181 ggaatgttcatcacgcagctgcgcgcagtgttgctgccagtcatattgatttaacggctg
      S  H  E  D  R  L  Q  A  C  H  Q  Q  W  D  Y  Q  N  L  P  Q
 241 ctgtaatgctggtaacagagcatttaaatcaccttgtaatgccacatgtgcctgacgcag
      Q  L  A  P  L  L  A  N  L  D  G  Q  L  A  V  H  A  Q  R  L
 301 cttgttcatttctgccgggtcgatatccatatggataacactggcgtgtggcgcgaaggt
      K  N  M  E  A  P  D  I  D  M  H  I  V  S  A  H  P  A  F  T,
 361 gttcagtttgccggtcacccggtcatcaaaacgtgcgcccacggcgatcagcaggtcaca
      N  L  K  G  T  V  R  D  D  F  R  A  G  V  A  I  L  L  D  C
 421 ctcctgcaccgcgaagtttgccgctttggtgccgtgcatccccagcatgcccagatagta
      E  Q  V  A  F  N  A  A  K  T  G  H  M  G  L  M  G  L  Y  Y
 481 cggataatctgcttctactgcgcccagccctttcagcgtacaggtggcaggcattttgt
      P  Y  D  A  E  V  A  G  L  G  K  L  T  C  T  A  P  M  K  T
 541 ggcagcgagaaattcacgcaaagccggaactgcctgcgccatacccacgccaccgccaac
      A  A  L  F  E  R  L  A  P  V  A  Q  A  M  G  V  G  G  G  V
 601 gtacagcatcggtttttgcgcttttgccagcatctggcgcgcttgctcaacttcggcatg
      Y  L  M  P  K  Q  A  K  A  L  M  Q  R  A  Q  E  V  E  A  H
 661 tgggaaagtcacttcgttttcaacggtggtgaaccacggttccaggtcaccgctggctaa
      P  F  T  V  E  N  E  V  T  T  F  W  P  E  L  D  G  S  A  L
 721 ctggatatcttttgggatatcgaccagaaccggaccaggacgacctgagcaggcaacgtc
      Q  I  D  K  P  I  D  V  L  V  P  G  P  R  G  S  C  A  V  D
 781 gaatgcttcagccatgatgcgcggcaactcttccagcgactgcaccagaaagctatgctt
      F  A  E  A  M  I  R  P  L  E  E  L  S  Q  V  L  F  S  H  K
 841 ggtacaggctaacgacaatcccaggacatccacttcctgaaatgcgtcagtgccgataaa
      T  C  A  L  S  L  G  L  V  D  V  E  O  F  A  D  T  G  I  F
 901 cggtgcggacacttgaccggtgatggcaacaacagggatggaatctaacagtgcgtccgc
      P  A  S  V  Q  G  T  I  A  V  V  P  I  S  D  L  L  A  D  A
 961 aagcccggttatcaggttggttgcgcccggaccagacgtggcgatacatacgccagtttt
      L  G  T  I  L  N  T  A  G  P  G  S  T  A  I  C  V  G  T  K
1021 gccggtagcacgagcataaccgatagccgccattgccgcaccctgctcatgtcggcatag
      G  T  A  R  A  Y  G  I  A  A  M  A  A  G  Q  E  H  R  C  L
1081 caagtgctccacgccgccgtcatacaatgcatcgtaaaccggcataattgcgccacccgg
      L  H  E  V  G  G  D  Y  L  A  D  Y  V  P  M  I  A  G  G  P
1141 ataaccgaaaacggtgttcacaccctgtgcccgcaacgcatgtaccacccactgtgcgcc
      Y  G  F  V  T  N  V  G  Q  A  R  L  A  H  V  V  W  Q  A  G      ↑
1201 attcatagttagttccccgtcctgaatcttgagaaacagaattttgtgctgttattcatt
      N  M                                                           ilvG
                                         [← ilvGMEDA attenuator ]
1261 gtctgctcctcggttatgtttttaaggtcaaaaaaaacccccggacctttcggtgcgggg

          [ integration host factor site 1   ]
1321 gtcttagttcgttaaggcttgatctctaagcctttcctcgtccaagtgcagccccgcacg

1381 gtgggataataatcaccaccacgctaatcacgaccaggctaatcactcgtagaagggctg

                                                      [ trp
                                        [trp operator site 1 ]
1441 tcattttgtcttttcttgcatcttgttcgaaggaatgcctaaagagttaccataggtttt
                                          -[* ilvGp2]
   operator site 2]
                                                         [integration
1501 gtaaatagtacaagatatttttggccgattttatgatgcgtttaataatttcaatata
                                                     -[* ilvGp1]

        host factor site 2          ]
1561 aattgttgtttttttaaaaataaaatagtgagtgaaaaaaattcaatttcagtatgaatgt
                              [ORFp *]-

1621 tgttcccttcctcgtagttcatcctttgtccgctgcattttgcgagcatcattccaccg

1681 ccagaacaactcaacttcacttgcacaaataaacaggaatcagccactaaaattcacgaa

1741 ttagcgggctacagcagcgatatcttatgcacataatggtgactcaaggagggcttatgt
                 [RBS]                 M  H  I  M  V  T  Q  G  G  L  M   ORFIII
1801 cactgtcaattgttcatacccgcgcagccctgggagtaaatgcgcccccgatcactgttg     ↓
     S  L  S  I  V  H  T  R  A  A  L  G  V  N  A  P  P  I  T  V
1861 aggtacatatcagtaaaggtctacccggcttaacgatggtgggcttaccagaaacaacgg
     E  V  H  I  S  K  G  L  P  G  L  T  M  V  G  L  P  E  T  T
1921 taaaagaagctcgcgatcgcgtgcgcagcgccattatcaatagcggatatgaatatccgg
     V  K  E  A  R  D  R  V  R  S  A  I  I  N  S  G  Y  E  Y  P
1981 cgaaaaaaatcaccatcaacctggctccagctgatctgccaaaagaaggggacgatatg
     A  K  K  I  T  I  N  L  A  P  A  D  L  P  K  E  G  G  R  Y
2041 atttacctatcgccattgcgttgctggcggcctcagaacagcttacagccaataagttag
     D  L  P  I  A  I  A  L  L  A  A  S  E  Q  L  T  A  N  K  L
```

```
2101 atgaatatgaattagtcggagaactggcgcttacaggcgctctgcgtggcgttcccggcg
     D  E  Y  E  L  V  G  E  L  A  L  T  G  A  L  R  G  V  P  G
2161 caatctccagtgcaactgaagctattaagtcgggcagaaaaattatcgtcgcgaaagata
     A  I  S  S  A  T  E  A  I  K  S  G  R  K  I  I  V  A  K  D
2221 acgaagatgaagtggggctaattaacggtgaaggatgcctgatagccgatcatctgcaag
     N  E  D  E  V  G  L  I  N  G  E  G  C  L  I  A  D  H  L  Q
2281 ctgtctgtgcgtttctggaaggtaagcacgctctcgaacgcccgaaaccaactgatgcag
     A  V  C  A  F  L  E  G  K  H  A  L  E  R  P  K  P  T  D  A
2341 tatcccgggcgctacaacatgatctcagtgatgttatcggtcaggaacaaggaaagcgag
     V  S  R  A  L  Q  H  D  L  S  D  V  I  G  Q  E  Q  G  K  R
2401 gactggaaattaccgctgctggcgggcacaaccttttactgattgggccgccgggaacag
     G  L  E  I  T  A  A  G  G  H  N  L  L  L  I  G  P  P  G  T
2461 gtaaaacaatgctcgccagccgtattaatggcctgttgccagatttaagcaatgaagagg
     G  K  T  M  L  A  S  R  I  N  G  L  L  P  D  L  S  N  E  E
2521 cactggagagtgctgcgatattaagtctggtaaatgctgaatcagtacaaaaacaatggc
     A  L  E  S  A  A  I  L  S  L  V  N  A  E  S  V  Q  K  Q  W
2581 ggcagcgcccgttccgctcacctcatcacagtgcatcgttaactgcgatggtaggcggtg
     R  Q  R  P  F  R  S  P  H  H  S  A  S  L  T  A  M  V  G  G
2641 gcgcaattccagggcccggtgaaatttcgctggcgcataacggcgtgctttttcttgatg
     G  A  I  P  G  P  G  E  I  S  L  A  H  N  G  V  L  F  L  D
2701 agctacctgaatttgaacggcgtacactggatgccttgcgagagccgattgaatccgggc
     E  L  P  E  F  E  R  R  T  L  D  A  L  R  E  P  I  E  S  G
2761 agatccatctttcacgcacacgagcaaaaataacctatccagcccgtttccagcttgttg [RBS]
     Q  I  H  L  S  R  T  R  A  K  I  T  Y  P  A  R  F  Q  L  V
2821 cggcgatgaatcccagccctaccggacattatcagggaaaccataaccgctgcacgccag [2826]
     A  A  M  N  P  S  P  T  G  H  Y  Q  G  N  H  N  R  C  T  P    ORFI
2881 aacagacattacgttatctcaaccggctctcggggccctttctcgaccgcttcgatctct    ↓
     E  Q  T  L  R  Y  L  N  R  L  S  G  P  F  L  D  R  F  D  L
2941 cactggagatcccattaccaccccccggcattttgagtaaaacggtagtgccgggagaaa
     S  L  E  I  P  L  P  P  P  G  I  L  S  K  T  V  V  P  G  E
3001 gcagcgccaccgttaaacaacgcgtaatggccgccagagagcgccaatttaagcggcaga
     S  S  A  T  V  K  Q  R  V  M  A  A  R  E  R  Q  F  K  R  Q
3061 ataaactgaatgcctggctggatagtccggaaatacgccaattctgcaaacttgagagcg
     N  K  L  N  A  W  L  D  S  P  E  I  R  Q  F  C  K  L  E  S
3121 aagatgcgatgtggttggaaggaacactgatccatctggggttatcgattcgtgcctggc
     E  D  A  M  W  L  E  G  T  L  I  H  L  G  L  S  I  R  A  W
3181 agcggttattgaaagttgcacgaaccattgctgatattgatcagtctgacattatcacac
     Q  R  L  L  K  V  A  R  T  I  A  D  I  D  Q  S  D  I  I  T
3241 gtcagcatttgcaggaggcagttagctatcgagcgattgaccgtttgctcatccatctgc
     R  Q  H  L  Q  E  A  V  S  Y  R  A  I  D  R  L  L  I  H  L
3301 agaaactactgacataaaaaaagggcatctcgcccttttttattaatcgtcagaatcggtg
     Q  K  L  L  T
3361 tagtcttcagcaccttcaacctgcggtttaccgccggaaagggtgtgaaaacgttttgga
3421 cgcttaatacgcgtcatatacttggaccacacgcgttctgcttctgtcactggctcacgt
3481 tcgccacggcatactgctacgaagagtttctcttcctcggtaaccggctcgcgtttgcca
3541 agatccaactcattgaaggcataaccatgacgctcaagcagttgtgcctctttgatggtg
3601 aaatcaccatgacgagagaatccacgtggataatgtttattgtcgaaatatcgattagtc
3661 gtcgtaaagcttctccgccatcctgcacgctcctaattctttgaccgagctagttatggc
3721 gcggagtattagttacgcttgacagagtgtaaaacaaaacatttaaatcataacgacaaa
3781 taattttgcggagagcactgtggatacggaattgttaaaaactttcctggaagttagccg
3841 aacgcgtcactttggtcgagcggctgaatcgctctatctgacccagtcagcagtgagctt
3901 tcgaatcagacaactggaaaatcaactgggtgtgaaccttttcacccgccacagaaacaa
3961 tatccgtttaaccgctgccggtgaaaaactactgccttatgcagaaacgctcatgagcac
4021 tggcaggccgccgtaagaggtggcgcatactcacgacataacgagttttctatcggtgcc
4081 agcgcctcgttgtgggaatgtatgcttaatcagtggctgggacgcttgtatcaaaatcag
                          M  L  N  Q  W  L  G  R  L  Y  Q  N  Q    ORFII
4141 gatgcccatacaggcttacagttcgaagcgcgaattgcccaacggcagtctctggtaaaa     ↓
      D  A  H  T  G  L  Q  F  E  A  R  I  A  Q  R  Q  S  L  V  K
4201 cagctgcatgaacgccagcttgatcttcttattaccactgaagcgcccaaaatggacgaa
      Q  L  H  E  R  Q  L  D  L  L  I  T  T  E  A  P  K  M  D  E
4261 tttagtagtcagttgctgggatatttcactttagcgctttataccagtgccccttcaaaa
      F  S  S  Q  L  L  G  Y  F  T  L  A  L  Y  T  S  A  P  S  K
4321 ctaaagggagatcttaattatctgcgacttgagtgggggccagattttcaacagcatgag
      L  K  G  D  L  N  Y  L  R  L  E  W  G  P  D  F  Q  Q  H  E
4381 gcaggtttgatcggtgctgacgaagtgcccattctgacaaccagttctgctgaactggca
      A  G  L  I  G  A  D  E  V  P  I  L  T  T  S  S  A  E  L  A
4441 cagcaacagattgcgatgcttaatggttgcacctggctacccgtcagctgggcgcgtaaa
      Q  Q  Q  I  A  M  L  N  G  C  T  W  L  P  V  S  W  A  R  K
4501 aaaggcggcctgcataccgttgtcgatagcacaacactttcacggccgctttatgccata

      K  G  G  L  H  T  V  V  D  S  T  T  L  S  R  P  L  Y  A  I
4561 tggctgcaaaatagcgataaaaatgcgttgattcgcgatcttttgaaaattaacgtgctg
      W  L  Q  N  S  D  K  N  A  L  I  R  D  L  L  K  I  N  V  L
4621 gatgaagtgtattaatatgaatggctggcaaggatgccggtagaaggatttacttcggag
      D  E  V  Y
                           [   rrnC terminator  ]
4681 agggttatttcagataaaaaaaatccttagctttcgctaaggatgatttctggcaggggc [rrnC]
4741 ggagagactcgaactcccaacacccggttttggagaccggtgctctaccaattgaactac    ↑
4801 gcccctaattagggtggcggaacggacgggactcgaacccgcgaccccctgcgtgacagg
4861 caggtattctaaccgactgaactaccgctccaccgaattc
                                       [EcoRI]
```

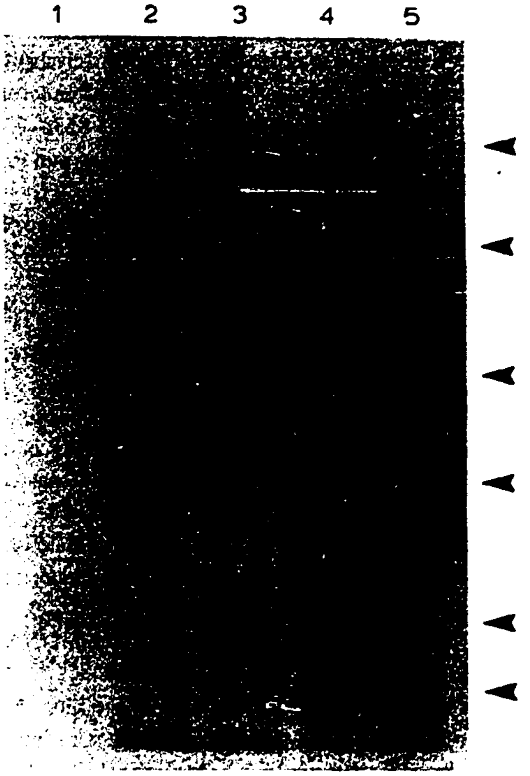


Fig. 20. Maxicell analysis of the *ilv*-related ORFI peptide. Derivatives of the maxicell strain CSR603 containing no plasmid (lane 1), pBR322 (lane 2), pMS143 (lane 3), pMS147 (lane 4), and pCC2 (lane 5) were used to detect plasmid coded proteins. Cells containing plasmid pMS143 and pMS147 produce the ORFI peptide of 18,751 daltons. Plasmid pCC2 is a derivative of plasmid pMS147 with a deletion of a *Cla*I fragment that should reduce the size of the ORFI peptide to 16,930 daltons. The arrows on the right indicate molecular weight markers of 92,000, 66,200, 45,000, 31,000, 21,500, and 14,400 daltons (Coppola *et al.*, 1991).

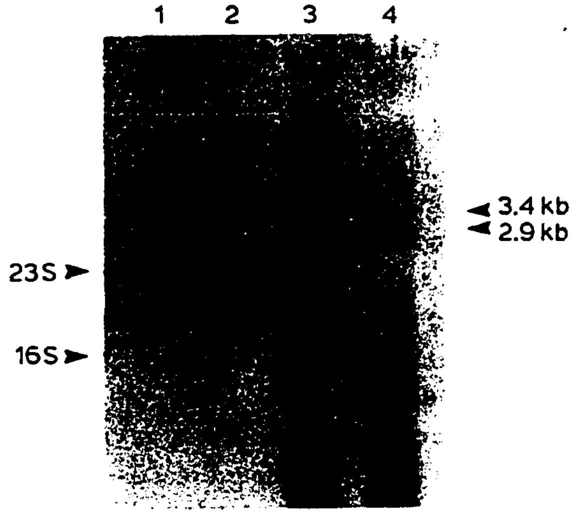


Fig. 21. Northern blot detection of ORFI specific transcripts. RNA was analyzed on 1.2% agarose gels containing 6.6% formaldehyde, transferred to nylon filters(Nytran, 0.45 µm, Schleicher & Schuell, Inc., Keene, NH.), and hybridized to the [^{32}p]-labeled probe V, specific for ORFI mRNA. RNA came from cells containing no plasmid (lane 1), plasmid pBR322 (lane 2), plasmid pMS147 (lane 3), and plasmid pCC2 (lane 4). The position of the 16S and 23S ribosomal RNA are indicated on the left, and the approximate size of the transcripts in lane 3 are indicate on the right. The decreased size of the transcript is in good agreement with the deletion of 504 bp deleted by *Cla*I treatment of plasmid pMS147 (Coppola *et al.*, 1991).

```
ECORRNILV intergenic region translated (+fs) ORF II
E. coli ilvY regulatory protein

MDTELLKTFLEVSRTRHFGRAAESLYLTQSAVSFRIRQLENQLGVNLFTRHRNNIRLTAA     60
::   ::::: .  .::::: :    .. : .:  :  ::  ::  :: :    . :: :
MDLRDLKTFLHLAESRHFGRSARAMHVSPSTLSRQIQRLEEDLGQPLFVRDNRTVTLTEA     60

GEKLLPYA-ETLMSTG--RAAVRGGAYSRHNEFSIGASASLWECMLNQWLGRL-YQNQDA    116
:: :  .:  ::      :  .   . :   :  :  : .     :   : :
GEELRVFAQQTLLQYQQLRHTIDQQGPSLSGELHIFCSVTAAYSHLPPILDRFRAEHPSV    120

HTGLQFEARIAQRQSLVKQLHERQLDLLITTEAPKMDEFSSQLLGYFTLALYTSAPSKLK    176
   :           .:    .  .     :            :.   .:     : .
EIKLTTGDAADAMEKVVTGEADLAIAGKPETLPGAVAFSMLENLAVVLIAPALPCPVRNQ    180

GDLNYLRLEWGPDFQQHEAGLIGADEVPILTTSSAELAQQQIAMLNGCTWLPVSWARKKG    236
  .      . :      . .    :.                         :  .   .
VSVEKPDWSTVPFIMADQGPVRRRIELWFRRNKISNPMIYATVGGHEAMVSMVALGCGVA    240

GLHTVVDSTTLSRPLYAIWLQNSDKNALIRDLLKINVLDEVY                      278
 :  ::   .       . .          .:        .
LLPEVVLENSPEPVRNRVMILERSDEKTPFELGVCAQKKRLHEPLIEAFWKILPNHK       297

Aligned 282, Matches 53, Mismatches 229, Score 83, Homology 18%
```

Fig. 22. Homology between ORFII and *ilvY* coded proteins. Standard one-letter amino acid abbreviations are used. Pairs of sequences are aligned with gaps (-) inserted to obtain maximal similarities. Residues identical in two sequences are shown (:), residues similar are shown (·) (Henikoff, 1991, personal communication).

```
             1        10        20        30        40        50
ILRE         MDTELLKTFLEVSRTRHFGRAAESLYLTQSAVSFRIRQLENQLGVNLFTRHRNN..
                               | helix-turn-helix |
RBCR       MHVSLRQLRVFEAVARHNSYTRAAEELHLSQPAVSMQVRQLEDEIGLSLFERLGKQVV
CFXO  (M) APHWTLRQLRLVALAAASGSYAKAAQDMGLSPPAVTAQMKALEEDIGVPMFERVDGRLR
LYSR       MAAVNLRHIEIFHAVMTAGSLTEAAHLLHTSQPTVSRELARFEKVIGLKLFERVRGR.L
CYSB      MKLQQLRYI..VEVVNHNLNVSSTAEGLYTSQPGISKQVRMLEDELGIQIFSRSGKH.L
ILVY         MDLRDLKTFLHLAESRHFGRSARAMHVSPSTLSRQIQRLEEDLGQPLFVRDNRT.V
GLTC  MDVRQTWSLRQLRYFMEVAEREHVSEAADHLHVAQSAISRQIANLEEELNVTLFEREGRN.I
OXYR         MNIRDLEYLVALAEHRHFRRAADSCHVSQPTLSGQIRKLEDELGVMLLERTSRK.V
MLER       MSLNLRDLEYFYQLSKLRSFTNVAKHFRVSQPTISYAIKRLETYYDCDLFYKDSSHQV
METR        MIEIKHLKTLQALRNSGSLAAAAAVLHQTQSALSHQFSDLEQRLGFRLFVRKSQP..
AMPR    MTRSYLPLNSLRAFEAAARHLSFTHAAIELNVTHSAISQHVKTLEQHLNCQLFVRVSRG..
TRPI    MSRDLPSLNALRAFEAAARLHSISLAAEELHVTHGAVSRQVRLLEEDLGVALFGRDGRG..
NODD    MRFKGLDLNLLVALDALMTERKLTAAARSINLSQPAMSAAIGRLRAYFNDELFLMQQRR.L
NAHR    MELRDLDLNLLVVFNQLLVDRRVSITAENLGLTQPAVSNALKRLRTSLQDPLFVRTHQG.M
LEUO        MVDLNLLTVFDAVMQEQNITRAAHVLGMSQPAVSNAVARLKVMFNDELFVRYGRG.I
CATR         MELRHLRYFKVLAETLNFTRAAELLHIAQPPLSRQISQLEDQLGTLLVVRE.RP..
CATM         MELRHLRYFVTVVEEQSISKAAEKLCIAQPPLSRQIQKLEEELGIQLFERGFRP..
ANTO    MSMSHINYNHLYYFWHVYKEGSVVGAAEALYLTPQTITGQIRALEDALQAKLFKRKGTW..
SVIR      MDFLINKKLKIFITLMETGSFSIATSVLYITRTPLSRVISDLERELKQRLFIRKNGT..
CHVO           MSHLRMLVMIEEHGQVSAAAAAMNMTQPAASRMLSEMEAIVKSPLCQRASRG..
TFDS         MEFRQLRYFVAAAEEGNVGAAARRLHISQPPVTRQIHALEQHLGVLLFERSARG..
             1        10        20        30        40        50
```

Fig. 23. Comparison of N-terminal sequences of the LysR family of transcriptional regulatory proteins. *ilvY* and ORFII are members of this family. Conserved amino acids are underlined. A helix-turn-helix region is indicated (Henikoff, 1991, personal communication).

```
               250       260        270       280       290  --ATP?-
  CRTA  CMTTAVARLQPSASGAKTRPV.FPFSAIVGQEDMKLALLLTAVDPGIGGVLVFGDRGTGK
            *             *** :** *****::*** ***  :** ****:: *******
  CPCS  SKKSA.............RPV.YPFAAIVGQDEMKLCLLLNVIDPKIGGVMIMGDRGTGK
                          *        ::**:: *  *            :  *  ****
ILVREL  PTDAV............SRALQHDLSDVIGQEQGKRGLEITAAGGH..NLLLIGPPGTGK
        190                   200       210       220         230

        -       310       320           330          340       350
  CRTA  STAVRALAALLPEIEAVEGCPVSSPNVEM....IPDWATVLSTN...VIRKPTPVVDLPL
        ** ** *  *****: * * *  *  ::     :     *   :   **      *****
  CPCS  STTVRSLVDLLPEINVVAGDPYNSDPIDPEFMGVEVRERVEKGEQVPVIATKINMVDLPL
                 ***:      :  :*  *                *    :       : *
ILVREL  TMLASRINGLLPDLS..NEEALESAAI..........LSLVNAE...SVQKQWR..QRPF
            240         250                 260          270

             360       370       380       390       400       410
  CRTA  GVSEDRVVGALDIERAISKGEKAFEPGLLARANRGYLYIDECNLLEDHIVDLLLDVAQSG
        *  **** *  *** *   * ********* **** **:** ***:** ** *** * **
  CPCS  GATEDRVCGTIDIEKALTEGVKAFEPGLLAKANRGILYVDEVNLLDDHLVDVLLDSAASG
                 *     *:  *     **    *  *:*: **    :    * * :   **
ILVREL  RSPHHSASLT.....AMVGGGAIPGPGEISLAHNGVLFLDELPEFERRTLDALREPIESG
         280            290       300       310       320       330

             420       430       440       450       460        470
  CRTA  ENVVERDGLSIRHPARFVLVGSGNPEEGDLRPQLLDRFGLSVEVLSPRDVETR.VEVIRR
         * ***:* ** *****:*:********:**********:  :*   ** : * * ::
  CPCS  WNTVEREGISISHPARFILIGSGNPEEGELRPQLLDRFGMHAQVGTVRDADLR.VKIVEE
             *    *  **** *:   **          :*                     :
ILVREL  QIHLSRTRAKITYPARFQLVAAMNPSPTGHYQGNHNRCTPEQTLRYLNRLSGPFLDRFDL
              340       350       360       370       380       390

              480       490          500       510       520
  CRTA  RDTYDADPKAFLEEWRPKDMD...IRNQILEARERLPKVEAPNTALYDCAAL...CIALG
        *  :* :** * :     :       :**  **  *  *:               *  *
  CPCS  RARFDSNPKDFRDTYKTEQDK...LQDQISTARANLSSVQIDRELKVKISRV...CSELN
               *           :      : :  **      *   :       :   *
ILVREL  SLEIPLPPPGILSKTVVPGESSATVKQRVMAARERQFKRQNKLNAWLDSPEIRQFCKLES
              400       410       420       430       440       450

          530       540       550       560       570       580
  CRTA  SDGLRGELTLLRSARALAALEGATAVGRDHLKRVATMALSHRLRRDPLDEAGSTARVART
         *****:    * * ***** *   *  *    *    * **** ***:   *   *
  CPCS  VDGLRGDIVTNRAAKALAALKGKDRVTPDDVATVIPNCLRHRLRKDPLESIDSGVLVSEK
         * :  :           *      *          :        :         *
ILVREL  EDAMWLEGTLIHLGLSIRAWQRLLKVARTIADIDQSDIITRQHLQEAVSYRAIDRLLIH.
              460       470       480       490       500       510

          590
  CRTA  VEETLP
          *
  CPCS  FAEIFS

ILVREL  LQKLLT
           516
```

Fig. 24. Homology among ORFIII, CRTA and CPCS proteins. ORFIII codes for *ilv*-related protein; CrtA codes for spheriodene monooxygenase; CPCS is a "novel" *Arabadopsis* chloroplast protein. Standard one-letter amino acid abbreviations are used. Gaps (---) are introduced to achieve maximum homology. Identical residues are shown by (*), similar residues are shown by (:). An ATP-binding site motif is indicated (Henikoff, 1991, personal communication).

BIBLIOGRAPHY

Adams, C.W., Rosenberg, M., and Hatfield, G.W. (1985). Analysis of the *in vivo* RNA transcription products of the *ilvGMEDA* attenuator region of *Escherichia coli* K-12. J. Biol. Chem. 260: 8538-8544.

Albrechtsen, B., Ross, B.M., Squires, C., and Squires, C.L. (1991). Transcriptional termination sequence at the end of the *Escherichia coli* ribosomal *rrnG* operon: Complex terminators and antitermination. Nucl. Aci. Res. 19: 1845-1852

Appleyard, R.K. (1954). Segregation of new lysogenic types during growth of a doubly lysogenic strain derived from *Escherichia coli* K-12. Genetics 39: 440-449.

Armstrong, G.A., Albert, M., Leach, F., and Hearst, J.E. (1989). Nucleotide sequence, organization and nature of the protein products of the carotenoid biosynthesis gene cluster of Rhodobacter capsulatus. Mol. Gen. Genet. 216: 254-268.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (1991). Current Protocols in Molecular Biology. John Wiley and Sons, NY.

Bachman, B.J., Low, K.B., and Taylor, A.L. (1976). Recalibrated linkage map of *Escherichia coli* K-12. Bacteriol. Rev. 40: 116-167.

Baez, M., Patin, D.W., and Calhoun, D.H. (1979). Deletion mapping of the *ilvGMEDAC* genes of *Escherichia coli* K-12. Mol. Gen. Genet. 169:289-297.

Bairoch, A. (1989). PROSITE: A Dictionary of Protein Sites and Patterns. Second Edition, University of Geneva, Switzerland.

Bardwell, J.C.A., Regnier, P., Chen, S.M., Nakamura, Y., Grunberg-Manago, M., and Court, D.L. (1989). Autoregulation of RNase operon by mRNA processing. EMBO J 8: 3401-3407.

Belasco, J.G., and Higgins, C.F. (1988). Mechanisms of mRNA decay in bacteria: a perspective. Gene 72: 15-32.

Belasco, J.G., Nilsson, G., Gabain, A., and Cohen, S.N. (1986). The stability of *E. coli* gene transcripts is dependent on determinants located to specific mRNA segments. Cell 46: 245-251.

Bennett, D.C., and Umbarger, H.E. (1984). Isolation and analysis of two *Escherichia coli* K-12 *ilv* attenuator deletion mutants with high level constitutive expression of an *ilv-lac* fusion operon. J. Bacteriol. 157: 839-845

Berg, C.M., and Shaw, K.J. (1981). Organization and regulation of the *ilvGEDA* operon in *Salmonella typhimurium* LT2. J. Bacteriol. 145: 984-989.

Berg, C.M., Shaw, K.J., Vender, J., and Borucka-Mankiewicz, M. (1979). Physiological characterization of polar Tn5 induced isoleucine-valine auxotrophs in *Escherichia coli* K-12: evidence for an internal promoter in the *ilvOGEDA* operon. Genetics 93: 309-319.

Blatt, A.J., and Umbarger, H.E. (1972). On the role of isoleucyl-tRNA synthetase in multivalent repression. Biochem. Genet. 6: 99-118.

Blazey, D.L., and Burns, R.O. (1979). Transcriptional activity of the transposable element Tn10 in the *Salmonella typhimurium ilv* gene cluster. Mol. Gen. Genet. 177: 1-11.

Blundell, M., and Kennell, D. (1974). Evidence of endonucleolytic attack in decay of *lac* messenger RNA in *Escherichia coli*. J. Mol. Biol. 83: 143-161.

Bogosian, G., and Somerville, R.L. (1983). *Trp* repressor protein is capable of intruding into other amino acid biosynthetic systems. Mol. Gen. Genet. 191: 51-58.

Bolivar, F., Rodriguez, R.L., Greene, P.L., Betlach, M.C., Heynecker, B.L., Boyer, H.W., Crosa, J.H., and Falkow, S. (1977). Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene 2: 95-113.

Brawerman, G. (1987). Determinants of messenger RNA stability. Cell 48: 5-6.

Calhoun, D.H. (1976). Threonine deaminase from *Escherichia coli*: Feedback-hypersensitive enzyme from a genetic regulatory mutant. J. Bacteriol. 126: 56-63.

Calhoun, D.H., and Gray, J.E. (1982). Cloning of the *ilvA538* gene coding for feedback-hypersensitive threonine deaminase from *Escherichia coli* K-12. J. Bacteriol. 15: 274-280.

Calhoun, D.H., and Hatfield, G.W. (1975). Autoregulation of gene expression. Ann. Rev. Microbiol. 29: 275-299.

Calhoun, D.H., and Hatfield, G.W. (1973). Autoregulation: a role for a biosynthetic enzyme in the control of gene expression. Proc. Natl. Acad. Sci. USA 70: 2757-2761.

Calhoun, D.H., Wallen, J.W., Traub, L., Gray, J., and Kung, H.-F. (1985). Internal promoter in the *ilvGEDA* transcription unit of *Escherichia coli* K-12. J. Bacteriol. 161: 128-132.

Cannistraro, V.J., Subbarao, M.N., and Kennell, D. (1986). Specific endonucleolytic cleavage sites for decay of *Escherichia coli* mRNA. J. Mol. Biol. 192: 257-274.

Casadaban, M.J., and Cohen, S.N. (1980). Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J. Mol. Biol. 138: 179-207.

Chen, C.Y., Beatty, J.T., Cohen, S.N., and Belasco, J.G. (1988). An intercistronic stem-loop structure function as an mRNA decay terminator necessary but insufficient for *puf* mRNA stability. Cell 52: 609-619.

Cho, K-O., and Yanosky, C. (1988). Sequence changes preceding a Shine-Dalagarno region influence *trpE* mRNA translation and decay. J. Mol. Biol. 204: 51-60.

Chung, C.T., Niemela, S.L., and Miller,R.H. (1989). One-step preparation of competent *Escherichia coli*: Transformation and storage of bacterial cells in the same solution. Proc. Natl. Acad. Sci. U.S.A. 86: 2172-2175.

Coppola, G., Huang, F., Riley, J., Cox, J.L., Hantzopoulos, P., Znou, L.-B., and Calhoun, D.H. (1991). Sequence and transcriptional activity of the *Escherichia coli* chromosome region between *rrnC* and *ilvGMEDA*. Gene 97: 21-27.

Cortese, R., Landsberg,R., Vonder Hesr, R.A., and Umbarger, H.E. (1974). Pleiotrophy of *hisT* mutants blocked in pseublouridine synthesis in RNA: leucine and isoleucine-valine operons. Proc. Natl. Acad. Sci. USA 71: 1857-1861.

Cox J. L., Cox, B.J., Fidanza, V., and Calhoun, D. H. (1978). The complete nucleotide sequence of the *ilvGMEDA* cluster of *Escherichia coli* K-12. Gene 56: 185-198.

Craig, N., and Nash, N.A. (1984). *E. coli* integration host factor binds to specific sites in DNA. Cell 39: 707-716.

Dale, R.M.K., McClure, B.A., and Houcuins, J.P. (1985). A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondria 185 rDNA. Plasmid 13: 31-40.

Daniel, J., Dondon, L., and Danchin, A. (1983). 2-Ketobutyrate: a putative alarmone of *Escherichia coli*. Mol. Gen. Genet. 190: 452-458.

Defelice, M.J., Guardiola, J., Esposito, B., and Iaccarino, M. (1974). Structural genes for a newly recognized acetolactate synthase in *Escherichia coli* K-12. J. Bacteriol. 120: 1068-1077.

Donovan, W.P., and Kushner, S.R. (1986). Polynucleotide phosphorylase and ribonuclease II are required for cell viability and mRNA turnover in *Escherichia coli* K-12. Proc. Natl. Acad. Sci. USA 83: 120-124.

Driver, R.P., and Lawther, R.P. (1985). Physical analysis of deletion mutations in the *ilvGMEDA* operon of *Escherichia coli* K-12. J. Bacteriol. 162: 598-606.

Eidlic, L., and Neidhardt, F.C. (1965). Role of Valyl-tRNA synthetase in enzyme repression. Proc. Natl. Acad. Sci. USA 53: 539-543.

Emory, S.A., and Belasco, J.G. (1990). The *ompA* 5' untranslated RNA segment functions in *E. coli* as a growth-rate-regulated mRNA stabilizer whose activity is unrelated to translational efficiency. J. Bacteriol. 172: 4472-4481.

Ericson, M.L. (1990). Quick DNA recovery from agarose gels by ultracentrifuge run. Trends Genet.: Vol. 6, No. 9: 378.

Favre, R., Wiater, A., Puppo, S., Iaccarino, M., Noelle, R., and Freundlich, M. (1976). Expression of a valine-resistant acetolactate synthase activity mediated by the *ilvO* and *ilvG* genes of *Escherichia coli* K-12. Mol. Gen. Genet. 1453: 243-252.

Fickett, J.W. (1982). Recognition of protein coding regions in DNA sequences. Nucl. Acids Res. 10: 5303-5318.

Freundlich, M., Burns, R.O., and Umbarger, H.E. (1962). Control of isoleucine, valine and leucine biosynthesis. I. Multivalent repression. Proc. Natl. Acad. Sci. USA. 48: 1804-1808.

Friden, P., Voelkel, K., Sternglanz, R., and Freundlich, M. (1984). Reduced expression of the isoleucine and valine enzymes in integration host factor mutants of *Escherichia coli.* J. Mol. Biol. 172: 573-579.

Frieddberg, D., Rosenthal, E.R., Jones, J.W., and Calvo, J.M. (1985). Characterization of the 3' end of the leucine operon of *Salmonella typhimurium.* Mol. Gen. Genet. 199: 486-494.

Friedman, D.I., Olson,E.J., Carver, D., and Gellert, M. (1984). Synergistic effect of *himA* and *gyrB* mutations: evidence that Him functions control expression of *ilv* and *xyl* genes. J. Bacteriol. 157: 484-489.

Gayda, D.J., Leathers, T.D., Noti, J.D., Smith, F.J., Smith, J.M., Subrahmanyam, C.S., and Umbarger, H.E.(1980). Location of the multivalent control site for the *ilvEDA* operon of *Escherichia coli.* J. Bacteriol. 142: 556-567.

Ghora, B.K., and Apirion, D. (1978). Structural analysis and *in vitro* processing to p5 rRNA of a 9S RNA molecule isolated from an *rnc* mutant of *Escherichia coli.* Cell 15: 1055-1066.

Gilman, M. (1987). Rapid isolation of RNA from gram-negative bacteria. In Ausubel, R.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith,J.A. and Struhl, K. (Eds.) Current Protocols in Molecular Biology. John Wiley and Sons, New York, pp. 4.4.1.-4.4.4.

Gorski, K., Roch, J-M., Prentki, P., and Krisch, H.M. (1985). The stability of bacteriophage T_4 gene 32 mRNA: a 5' leader sequence that can stabilize mRNA transcripts. Cell 43: 461-469.

Gray, J.E., Bennett, D.C., Umbarger, H.E., and Calhoun, D.H. (1982). Physical and genetic localization of *ilv* regulatory sites in *ilv* bacteriophages. J. Bacteriol. 149: 1071-1081.

Gray, J.E., Patin, D.W., and Calhoun, D.H. (1981). Identification of the protein products of the *rrnC, ilv, rho* region of the *Escherichia coli* K-12 chromosome. Mol. Gen. Genet. 183: 428-436.

Gray, J.E., Wallen, J.W., and Calhoun, D.H. (1982). Identification of a protein of 15,000 daltons related to isoleucine-valine biosynthesis in *Escherichia coli* K-12. J. Bacteriol. 151: 127-134.

Hahn, J.E., and Calhoun, D.H. (1978). Suppressers of a genetic regulatory mutation affecting isoleucine-valine biosynthesis in *Escherichia coli* K-12. J. Bacteriol. 136: 117-124.

Hall, B.G., Yokoyama, S., and Calhoun, D.H. (1983). Role of cryptic genes in microbial evolution. Mol. Biol. Evol. 1 (1): 109-124.

Harms, E., Higgins, E., Chen, J.-W., and Umbarger, H.E. (1988). Translational coupling between the *ilvD* and *ilvA* genes of *Escherichia coli.* J. Bacteriol 170: 4798-4807.

Harms, E., and Umbarger, H.E. (1991). The absence of branched-chain amino and growth rate control at the internal *ilvEp* promoter of the *ilvGMEDA* operon. J. Bacteriol. 173: 6446-6452.

Harris, C.L., Liu, L., Sakallah, S., and Devore, R. (1983). Cysteine starvation, isoleucyl-$tRNA^{Ile}$, and the regulation of the *ilvGEDA* operon of *Escherichia coli*. J. Biol. Chem. 258: 7676-7683.

Higgins, C.F., McLaren, R.S., and Newburg, S.F. (1988). Repetitive extragenic palindromic sequences, mRNA stability and gene expression: Evolution by gene conversion?— a review. Gene 72:3-14.

Johnson, D. I., and Somerville, R.L. (1983). Evidence that repression mechanisms can exert control over the *thr, leu*, and *ilv* operons of *Escherichia coli* K-12. J. Bacteriol. 155: 49-55.

Johnson, D.I., and Somerville, R.L. (1984). New regulatory genes involved in the control of transcription initiation at the *chr* and *ilv* promoters of *Escherichia coli* K-12. Mol. Gen. Genet. 195: 70-76.

Kikuchi, Y., and Nash, H.A. (1978). The bacteriophage λ *int* gene product. J. Biol. Chem. 253: 7149-7157.

Kirsebon, L.A., Baer, M.F., and Altman, S. (1988). Differential effects of mutations in the protein and RNA moieties of RNase P on the efficiency of suppression by various tRNA suppressors. J. Mol. Biol. 204: 879-888.

Koncz, C., Mayerhofer, R., Koncz-Kalman, Z., Nawrath, C., Reiss, B., Redei, G.P., and Schell, J. (1990). Isolation of a gene encoding a novel chloroplast protein by T-DNA tagging in *Arabadopsis thaliana*. EMBO J. 9: 1337-1347.

Kovach, J.S., Berberich, M.A., Venetianer, P., and Goldberger, R.F. (1969). Repression of the histidine operon: effect of the first enzyme of the kinetics of repression. J. Bacteriol 97: 1283-1290.

Kuramitsu, S., Ogawa, H., Ogawa, T., and Kagimiyama, H. (1985). Branched-chain amino acid aminotransferase of *Escherichia coli*: nucleotide sequence of the *ilvE* gene and the deduced amino acid sequence. J. Biochem. 97: 993-999.

Lawther, R.P., and Hatfield, G.W. (1980). Multivalent translational control of transcription termination at the attenuator of *ilvGEDA* operon of *Escherichia coli* K-12. Proc. Natl. Acad. Sci. USA 77: 1862-1866.

Lawther, R.P., and Hatfield, G.W. (1978). A site of action for tRNA-mediated regulation of the *ilvGMEDA* operon of *Escherichia coli* K-12. Mol. Gen. Genet. 167: 227-234.

Lawther, R.P., Calhoun, D.H., Adams, C.W., Hauser, C.A., Gray, J., and Hatfield, G.W. (1981). Molecular basis of valine resistance in *Escherichia coli* K-12. Proc. Natl. Acad. Sci. USA 78: 922-925.

Lawther,R.P., Calhoun, D.H., Gray, J., Adams, C.W., Hauser, C.A., and Hatfield, G.W. (1982). DNA sequence fine-structure analysis of *ilvG* (IlvG$^+$) mutations of *Escherichia coli* K-12. J. Bacteriol. 149: 294-298.

Lawther, R.P., Wek, R.C., Lopes, J.M., Pereira, R., Taillon, B.E., and Hatfield, G.W. (1987). The complete nucleotide sequence of the *ilvGMEDA* operon of *Escherichia coli* K-12. Nucl. Acids Res. 15: 2137-2155.

Levinthal, M., Williams, L.S., and Umbarger, H.E. (1973). Role of threonine deaminase in the regulation of isoleucine and valine biosynthesis. Nature (London) 246: 65-68.

Lopes, J.M., Soliman, N., Smith, P.K., and Lawther, R.P. (1989). Transcriptional polarity enhances the contribution of the internal promoter, *ilvEp*, in the expression of the *ilvGMEDA* operon in wild-type *Escherichia coli* K-12. Molec. Microbiol. 3: 1039-1051.

Lopes, J.M., and Lawther, R.P. (1986). Analysis and comparison of the internal promoter, pE, of the *ilvGMEDA* operons from *Escherichia coli* K-12 and *Salmonella typhimurium*. Nucl. Acids Res. 14: 2779-2798.

Lopes, J.M., and Lawther, R.P.(1989). Physical identification of an internal promoter, *ilvAp*, in the distal portion of the *ilvGMEDA* operon. Gene 76: 255-269.

Maniatis, T., Fristsch,E.F., and Sambrook, J. (1989). Molecular cloning: A Laboratory Manual. Second Edition. Cold Spring Harbor Laboratory.

Maxam, A.M., and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. 65: 499-460.

McCabe, P.C. (1990). Production of single-stranded DNA by asymmetric PCR, p. 76-83. In Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J. (eds.) PCR protocols. A Guide to Methods and Applications. Academic Press, Inc., N.Y.

Messing, J., and Viera, J. (1982). A new pair of M13 vectors for selecting either DNA strand of double digest restriction fragments. Gene 19: 269-276.

Miller, H.I., Kikuchi, Y., Nash, H.A., Weisberg, R.A., and Friedman, D.I. (1979). Site-specific recombination of bacteriophage λ: the role of host gene products. Cold Spring Harbor Symp. Quant. Biol. 43: 1121-1126.

Miller, J.H.(1974). Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, New York, pp. 352-355.

Mott, J.E., Calloway, J.L., and Platt, T. (1985). Maturation of *Escherichia coli* tryptophan operon mRNA: evidence for 3' exonucleolytic processing after *rho*-dependent termination. EMBO J. 4: 1887-1891.

Mudd, E.A., Krisch, H.M., and Highins, C.F. (1990). RNase E, an endoribonuclease, has a general role in the chemical decay of *Escherichia coli* mRNA: evidence that *rne* and *ams* are the same genetic locus. Mol. Microbiol. 4: 2127-2135.

Mudd, E.A., Prentki, P., Belin, D., and Kirsch, H.M. (1988). Processing of unstable bacteriophage T_4 gene 32mRNAs into a stable species requires *Escherichia coli* ribonuclease E. EMBO J. 7: 3601-3607

Murakawa, G.J., Kwan, C., Yamashita, J., and Nierlich, D.P. (1991). Transcription and decay of the *lac* messenger: role of an intergenic terminator. J. Bacteriol. 173: 28-36.

Nargang, F.E., Subrahmanyam, C.S., and Umbarger, H.E. (1980). Nucleotide sequence of *ilvGEDA* attenuator of *Escherichia coli*. Proc. Natl. Acad. Sci. USA 77 (1980) 1823-1827.

Newburg, S.F., Smith, N.H., Robinson, E.C., Hiles, I.D., and Higgins, C.F. (1987). Stabilization of translationally active mRNA by prokaryotic REP sequences, Cell 48:297-310.

Newman, T.C., and Levinthal,M. (1980). A new map position for the *ilvB* locus of *Escherichia coli*. Genetics 96:59-77.

Nikolaen, N., Glazier, K., and Schlessinger, D. (1985). Cleavage by ribonuclease III of the complex of 30S preribosomal RNA and ribosomal proteins of *Escherichia coli*. J. Mol. Biol. 94: 301-304.

Nilsson, P., and Uhlin, B.E.(1991). Differential decay of a polycistronic *Escherichia coli* transcript is initiated by RNase E-dependent endonucleolytic processing. Mol. Microbiology 5: 1791-1799.

Ortuno, M. J., and Lawther, R.P. (1987). Effect of the deletion of upstream DNA sequences on expression from the *ilvGp2* promoter of the *ilvGMEDA* operon of *Escherichia coli* K-12. Nucl. Acids Res. 15: 1521-1542.

Pagel, J.M., and Hatfield, G.W. (1991). Integration host factor-mediated expression of the *ilvGMEDA* operon of *Escherichia coli*. J. Biol. Chem. 266: 1985-1996.

Pereira, R.F., Ortuno, M.J., and Lawther, R.P. (1988). Binding of integration host factor (IHF) to the *ilvGp1* promoter of the *ilvGMEDA* cluster of *Escherichia coli* K-12. Nucl. Acids Res. 16: 5873-5989.

Pledger, W.J., and Umbarger, H.E. (1973). Isoleucine and valine metabolism in *Escherichia coli*. XXII. A pleiotropic mutation affecting induction of isomeroreductase activity. J. Bacteriol. 114: 195-207.

Portier, C., Dondon, L., Grunberg-Manago, M., and Regnier, P. (1987). The first step in the functional inactivation of the *Escherichia coli* polynucleotide phosphorylase message is a ribonuclease III processing at the 5' end. EMBO J6: 2165-2170.

Primerano, D.A., and Burns, R.O. (1982). Metabolic basis for the isoleucine, pantothenate or methionine requirement of *ilvG* strains of *Salmonella typhimurium*. J. Bacteriol 150: 1202-1211.

Robinson, C.L., and Jackson, J.H. (1982). New acetohydroxy acid synthase activity from mutational activation of a cryptic gene in *Escherichia coli* K-12. Mol. Gen. Genet. 186: 240-246.

Rosner, J.L. (1972) Formation, induction and curing of bacteriophage P1 lysogens. Virology 49: 676-689.

Ratzkin, B., Arfin, S., and Umbarger, H.E. (1972). Isoleucine and valine metabolism in *Escherichia coli*. XVIII. Induction of acetohydroxy acid isomeroreductase. J. Bacteriol. 112: 131-141.

Sameshima, J.H., Wek, R.C., and Hatfield, G.W. (1989). Overlapping transcription and termination of the convergent *ilvA* and *ilvY* genes of *Escherichia coli*. J. Biol. Chem. 264: 1224-1231.

Sancar, A., Hack, A.M., and Rupp, W.D. (1979). Simple method for identification of plasmid-coded protein. J. Bacteriol. 137: 692-693.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74: 5463-5467.

Schmeissner, U., Mckenney, K., Rosenberg, M., and Court, D. (1984). Removal of a terminator structure by RNA processing regulates *int* gene expression. J. Mol. Biol. 176:39-53.

Selden, R.F.(1987). Sieving agarose gel electrophoresis. In Ausubel, R.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (Eds.) Current Protocols in Molecular Biology. John Wiley and Sons, New York, pp.2.8.1.-2.9.10.

Singer, P.A., Levinthal, M,. and Williams, L.S. (1984). Synthesis of the isoleucyl and valyl-tRNA synthetases and the isoleucine-valine biosynthetic enzymes in a threonine deaminase regulatory mutant of *Escherichia coli* K-12. J. Mol. Biol. 175: 39-55.

Smith, J.M., Smolin, D.E., and Umbarger, H.E. (1976). Polarity and the regulation of the *ilv* gene cluster in *Escherichia coli* K-12. Mol. Gen. Genet. 148: 111-124.

Smolin, D.E., and Umbarger, H.E. (1975). Specificity of the stimulation of *in vitro* ribonucleic acid synthesis by guanosine 5'-diphosphate 3'-diphosphate. Mol. Gen. Genet. 141: 277-284.

Squires, C.H., Defelice, M., Wessler, S.R., and Calvo, J.M. (1981). Physical Characterization of the *ilvHI* operon of *Escherichia coli* K-12. J. Bacteriol. 147: 797-804.

Studier, F.W. (1975). Genetic mapping of a mutation that causes ribonuclease III deficiency in *Escherichia coli*. J. Bacteriol. 170: 2860-2865.

Subrahmanyam, C.S., Noti, J.D., and Umbarger, H.E. (1980). Regulation of *ilvEDA* expression occurs upstream of *ilvG* in *Escherichia coli*: additional evidence for an *ilvGEDA* operon. J. Bacteriol. 144:279-290.

Taillon, B.E., Little, R., and Lawther, R.P. (1988). Analysis of the functional domains of biosynthetic threonine deaminase by comparison of the amino acid sequences of three wild-type alleles to the amino acid sequence of biodegradative threonine deaminase. Gene 63: 245-252.

Tomcsanyi, T., and Apirion, D. (1985). Processing enzyme ribonuclease E specifically cleaves RNAI: an inhibitor of primer formation in plasmid DNA synthesis. J. Mol. Biol. 185: 713-720.

Tsui, P., and Freundlich, M. (1990). Integration host factor bends the DNA in the *Escherichia coli ilvBN* promoter region. Mol. Gen. Genet. 223: 349-352.

Tsui, P., and Freundlich, M.(1988). Integration host factor binds specifically to sites in the *ilvGMEDA* operon in *Escherichia coli*. J. Mol. Biol. 203: 817-820.

Umbarger, H.E. (1987). Biosynthesis of the branched-chain amino acids. In Neidhardt, F.C., Ingraham, J.L., Low, K.B., Magasanik, B., Schaechter, M. and Umbarger, H.E. (Eds), *Escherichia coli* and *Salmonella typhimurium* Cellular and Molecular Biology. American Society for Microbiology, Washington, D.C., pp. 352-367.

Umbarger, H.E. (1971). The regulation of enzyme levels in the pathways to the branched-chain amino acids. In: Metabolic regulation. Vol. 5, p. 447-462 (ed. H.J. Vogel). New York and London: Academic Press.

Vierra, J., and Messing, J. (1982). The pUC plasmids: an M13mp7 derived system for insertion, mutagenesis and sequencing with universal primers. Gene 19: 259-268.

von Gabain, A., Belasco, J.G., Schottel, J.L., Chang, A.C.Y., and Cohen, S.N. (1983) Decay of mRNA in *Escherichia coli*: investigation of the fate of specific segments of transcripts. Proc. Natl. Acad. Sci. USA 80: 653-657.

Vonder Haar, R. A., and Umbarger, H.E. (1972). Isoleucine and valine metabolism in *Escherichia coli*. XIX. inhibition of isoleucine biosynthesis by glycyl-leucine. J. Bacteriol. 112: 142-147.

Wasmuth, J.J., and Umbarger, H.E. (1973). Effect of isoleucine, valine, or leucine starvation on the potential for formation of the branched-chain amino acids biosynthetic enzymes. J. Bacteriol 116: 548-561.

Watson, M.D., Wild, D.J., and Umbarger, H.E. (1979). Positive control of *ilvC* expression in *Escherichia coli* K-12: identification and mapping of regulation gene *ilvY*. J. Bacteriol. 139: 1014-1929.

Weiss, D.L., Jonhson, D.I., Weith, H.L., and Somerville, R.L. (1986). Structural analysis of the *ileR* locus of *Escherichia coli* K-12. J. Biol. Chem. 261: 9966-9971.

Wek, R.C., and Hatfield, G.W. (1986). Examination of the internal promoter, P_E, in the *ilvGMEDA* operon of *Escherichia coli* K-12. Nucl. Acids Res. 14: 2763-2777.

Wek, R.C., and Hatfield, G.W. (1986). Nucleotide sequence and *in vivo* expression of the *ilvY* and *ilvC* genes in *Escherichia coli* K-12. Transcription from divergent overlapping promotors. J. Biol. Chem. 261: 244-2450.

Wek, R.C., Sameshima, J.H., and Hatfield, G.W. (1987). Rho-dependent transcriptional polarity in the *ilvGMEDA* operon of wild-type *Escherichia coli* K-12. J. Biol. Chem. 262: 15256-15261.

Winkelman, J.W., and Hatfield, G.W. (1990). Characterization of the integration host factor binding site in the *ilvGp1* promoter region of the *ilvGMEDA* operon of *Escherichia coli*. J. Biol. Chem. 265:10055-10060.

Wu, A.M., Christie, G.E., and Platt, T. (1981). Tandem termination sites in the tryptophan operon of *Escherichia coli*. Proc. Natl. Acad. Sci. 77: 5201-5205

Young, R.A. (1979). Transcription termination in the *Escherichia coli* operon *rrnC*. J. Biol. Chem. 254: 12725-12731.