JOURNAL OF BACTERIOLOGY, July 1985, p. 186–198
0021-9193/85/070186-13$02.00/0

Vol. 163, No. 1

Unusual Organization of the *ilvIH* Promoter of *Escherichia coli*

GEORGE W. HAUGHN,[1]† CHARLES H. SQUIRES,[2] MAURILIO DeFELICE,[3] CARMINE T. LARGO,[3] AND JOSEPH M. CALVO[1]*

Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, New York 14853[1]; Synergen, Boulder, Colorado 80301[2]; and International Institute of Genetics and Biophysics, Naples 80125, Italy[3]

Received 29 January 1985/Accepted 8 April 1985

Analysis of plasmids containing *ilvIH-galK* fusions indicated that the *Escherichia coli ilvIH* promoter and sequences sufficient to cause leucine repression lie within 363 base pairs (bp) of *ilvI*. Experiments designed to locate the promoter and regulatory sequences more precisely gave the following results. The positions of the 5′ endpoints of both unlabeled and pulse-labeled *ilvIH* mRNAs transcribed in vivo lie 30 bp upstream of *ilvI*. By contrast, the major in vitro RNA endpoints lie at positions further upstream. Several mutations which increase the expression of *ilvIH* lie 40 to 50 bp upstream of *ilvI*, within a putative promoter termed P1. Deletion of a 50-bp region immediately upstream of *ilvI*, which includes P1, resulted in the loss of all *ilvIH* promoter activity. Deletion of sequences more than 200 bp upstream of *ilvI* reduced *ilvIH* promoter activity by more than 80%. These results suggest that transcription of the *ilvIH* operon is initiated from promoter P1 but that sequences more than 200 bp upstream are required for optimal transcription of the operon.

The first reaction common to the biosynthesis of valine, isoleucine, and leucine is catalyzed by acetohydroxyacid synthase (AHAS) (11). *Escherichia coli* K-12 produces two AHAS isozymes, AHAS I and AHAS III, encoded by *ilvB* and *ilvIH*, respectively (20). *ilvI* and *ilvH* together comprise an operon and code for polypeptides with molecular weights of 61,000 and 17,000, respectively (21, 41, 42). The *ilvI* polypeptide is absolutely required for AHAS III activity, whereas the *ilvH* polypeptide enhances AHAS III activity and may also confer upon it sensitivity to inhibition by valine (42). The expression of the *ilvIH* operon is negatively controlled by leucine (12) at the level of transcription (42).

As a first step in determining the mechanism by which the *ilvIH* operon is regulated, we characterized sequences upstream of *ilvI*. We show that the *ilvIH* promoter and sequences involved in leucine regulation of *ilvIH* lay within 363 base pairs (bp) of the *ilvI* gene. Analyses of *ilvIH* mRNA 5′ endpoints and of deletion and point mutations suggest that transcription initiates within 30 bp of *ilvI*. However, optimal transcription of the *ilvIH* operon was markedly dependent upon sequences more than 190 bp upstream of *ilvI*. The organization of the *ilvIH* promoter is compared with that of other well-studied bacterial promoters, particularly with respect to the location and influence of upstream sequences.

MATERIALS AND METHODS

Strains and growth conditions. The bacterial strains, plasmids, and M13 bacteriophages used in this study are described in Table 1. L broth (LB [30]) and SSA (7) supplemented with 0.2% glucose and with amino acids (50 μg/ml) and vitamins (5 μg/ml) where needed were used as rich and minimal media, respectively. SSA contains the following (grams per liter of distilled water): K_2HPO_4, 10.5; KH_2PO_4, 4.5; $(NH_4)_2SO_4$, 1.0; sodium citrate dihydrate, 0.97; $MgSO_4$, 0.05. MOPS (morpholinepropanesulfonic acid) minimal medium (32) supplemented with glucose (0.2%), amino acids (50 μg/ml), vitamins (5 μg/ml), $NaHCO_3$ (10 mM), and K_2HPO_4 (0.2 mM) was used to grow cells for isolation of RNA pulse-labeled with ^{32}P. Ampicillin (100 μg/ml) or chloramphenicol (30 μg/ml) was added to the media for selection of plasmid-containing strains.

Recombinant DNA techniques. DNA restriction and modification enzymes were used according to the conditions specified by their manufacturers (New England Biolabs and Bethesda Research Laboratories). The following techniques were carried out by published procedures: isolation of plasmid DNA (38); isolation of single-stranded phage M13 DNA (37); isolation of l-μg amounts of plasmid or M13 replicative-form (RF) DNA (5); bacterial transformation (8; during preparation of competent cells, the NaCl wash was omitted, and 50 mM rather than 30 mM $CaCl_2$ was used); polyacrylamide gel electrophoresis (39); labeling of DNA fragments with ^{32}P at their 5′ (26) or 3′ (39) ends; DNA sequencing (26, 36); isolation of unlabeled DNA fragments from agarose gels by electroelution (25) and from 2-mm-thick polyacrylamide gels (39); isolation of radioactive DNA fragments from 0.4-mm-thick polyacrylamide gels (39; the maceration step was omitted, and DNA was not purified by DE52 chromatography); conversion of 5′-protruding ends to blunt ends, using reverse transcriptase (39; all necessary nucleoside triphosphates were added to reaction mixture); and phage M13 manipulations (29).

Enzyme assays. Cells were grown at 32°C and assayed for AHAS by the procedure of DeFelice et al. (10). For galactokinase assays, strains carrying pKO1-derived plasmids were grown to an A_{550} of 0.6 to 0.7 in 25 ml of minimal medium containing ampicillin (100 μg/ml) and chloramphenicol (30 μg/ml). Cells in 5-ml samples were centrifuged at 9,500 × *g* for 10 min at 4°C and stored at −20°C. Cell pellets were resuspended in 2 ml of lysis buffer (1 mM EDTA, 1 mM dithiothreitol [DTT], 20 mM Tris hydrochloride [pH 8]) and lysed at 3,000-lb pressure in a French pressure cell (0.95-cm-diameter piston) at 4°C. Debris was removed by centrifugation at 18,700 × *g* for 15 min at 4°C, and the crude extract was diluted 30-fold in lysis buffer containing 0.1 mg of bovine serum albumin per ml. Extracts were assayed for galactokinase activity by a modification of the procedure of McKenney et al. (28). The reaction mixture contained 1 mM [^{14}C]galactose (4 × 10^6 cpm/mmol; Amersham Corp.), 1 mM DTT, 3.2 mM NaF, 4 mM $MgCl_2$, 100 mM Tris hydro-

* Corresponding author.
† Present address: Plant Research Laboratory, Michigan State University, East Lansing, MI 48824.

TABLE 1. Bacteria, plasmids, and phage

Strain	Genotype or comments	Source (reference)
Bacteria		
W3102	*galK*	(3)
PS1283	HfrC *thi-1 trpR rbs-115* Δ(*ara-leu-ilvIH*) *ilvB alg-8*::Mu-1; AHAS I^-, AHAS III^-	M. Levinthal
PS1035	*thi-1 glyA ilvB619*; AHAS I^-, AHAS III^+	(12)
MI316	*thi-1 glyA ilvB619 ilvH612 ilvI614 ara bgl*: AHAS I^-, AHAS III^-	(12)
JM103	Δ(*lac-pro*) *thi rpsL supE endA subcB15 hspR4*/F′ *traD36 proAB lacI*^{q} Δ*lacZ15*	(29)
ID60, ID61, ID62, ID63, ID64, ID65, ID66, and ID69	MI316 containing plasmids pCV93 to pCV100, respectively (Table 3)	This study
ID86	MI316(pCV87)	(21)
CV696	PS1283(pCV35)	This study
CV670	MI316(pCV7)	(42)
CV709	W3102(pCV53)	This study
CV781	W3102(pK01, pSE150)	This study
CV782	W3102(pK0110, pSE150)	This study
CV787	W3102(pCV53, pSE150)	This study
CV788	W3102(pCV65, pSE150)	This study
CV794, CV795, CV798, CV801, CV803, CV804, CV805, and CV806	Strain W3102 carrying, respectively, pCV75, pCV76, pCV79, pCV82, pCV84, pCV85, pCV86 and pCV89	This study
CV808, CV809, CV810, CV811, CV812, CV813, CV814, CV816, CV817, CV818, CV819, an CV820	Strain W3102(pSE150) carrying, respectively, pCV75, pCV76, pCV82, pCV84, pCV85, pCV89, pCV101 pCV52, pCV66, pCV77, pCV79, and pCV86	This study
Plasmids		
pK01	*galK* fusion vector	(28)
pSE150	Low-copy-number plasmid	G. Walker
pK0110	pK01 derivative, *galK* transcribed from *lacUV5* promoter	(46)
pCV7	pBR322 derivative carrying *ilvI* and part of *ilvH*	(42)
pCV35	10.9-kilobase *Eco*RI fragment carrying the *E. coli leuABCD-ilvIH* region	(42)
pCV52	*Hinc*II$_{1733}$ fragment ligated into *Sma*I site of pK01, *ilvP* oriented toward *galK* (Fig. 1)	This study
pCV53	*Hae*III$_{361}$ fragment ligated into *Sma*I Site of pK01, *ilvP* oriented toward *galK* (Fig. 1)	This study
pCV65	pCV53 with *Hae*III$_{361}$ fragment in opposite orientation	This study
pCV66	pCV52 with *Hinc*II$_{1733}$ fragment in opposite orientation	This study
pCV74, pCV75, pCV76, pCV77, pCV78, pCV79, pCV80, pCV81, pCV82, pCV83, pCV84, pCV85, and pCV86	Deletion derivatives of pCV53 (Fig. 5)	This study
pCV87	pCV7 derivative carrying the entire *ilvIH* operon	(21)
pCV89	*Hin*fI-*Hae*III$_{178}$ fragment ligated into the *sma*I site of pK01, *ilvP* oriented toward *galK* (Fig. 1)	
pCV93, pCV94, pCV95, pCV96, pCV97, pCV98, pCV99, and pCV100	pCV7 derivatives carrying mutations that increase expression of *ilvIH* (Table 3)	This study
pCV101	*Hae*III-*Sau*3A$_{312}$ fragment ligated into the *Sma*I site of pK01, *ilvP* oriented toward *galK* (Fig. 1)	This study
Bacteriophage		
M13mp8	M13 cloning vector	(29)
M13GH1, M13GH2, M13GH3, M13GH4, M13GH5, and M13GH6	Phage (+) strand carries the coding strand of *ilVIH* fragments *Hin*fI$_{457}$, *Hin*fI-*Nde*I$_{349}$, *Nde*I-*Hin*fI$_{110}$, *Hae*III$_{174}$, *Hae*III-*Sau*3A$_{312}$, and *Sau*3A$_{89}$, respectively (Fig. 1)	This study
M13GH9	Phage (+) carries noncoding strand of *Hin*fI$_{457}$ (Fig. 1)	This study

chloride (pH 8), and 3.2 mM ATP in a total reaction volume of 100 μl. The reaction was started by addition of 20 μl of diluted crude extract and incubated at 30°C for 15 min. Samples (25 μl), taken immediately and 15 min after addition of crude extract, were spotted on Whatman DE81 paper strips (15 by 1 cm) lying on a dry ice block. Phosphorylated and unphosphorylated galactose were separated by descending paper chromatography (until front was within 2 cm of bottom), using H_2O as the solvent. The amount of phosphorylated [^{14}C]galactose remaining at the origin was determined by scintillation counting. Protein in crude extracts was determined spectrophotometrically by the method of Groves et al. (17).

Plasmid copy number. Strains containing pKO1-derived plasmids and low-copy-number plasmid pSE150 (from G. Walker) were grown and harvested as described for galactokinase assays. Plasmid DNA was isolated from pellets containing 10^9 cells by the procedure of Birnboim and

Doly (5) and treated with *Eco*RI (20 U) for 30 min at 37°C in a total volume of 20 μl. After addition of 4 μl of solution containing 50% glycerol, 0.2% bromophenol blue, and 0.2% xylene cyanol, 1- and 4-μl samples were fractionated on an 0.8% horizontal agarose gel by electrophoresis in Tris borate buffer (90 mM Tris, 90 mM boric acid, 3 mM EDTA [pH 8.3]) at 160 V for 5 h. The gel was stained with 0.4 μg of ethidium bromide per ml and photographed during illumination with UV light. Plasmid DNA was quantitated by scanning photographs with a densitometer and weighing the paper beneath the peaks. The amount of pKO1-derived plasmid divided by the amount of plasmid pSE150 is the relative copy number. X. Yang and J. Roberts (personal communication) found that plasmid pSE150 served as a useful internal standard for plasmid copy number determinations of λ 6S promoter-pKO1 fusions.

S1 nuclease mapping of 5′ RNA endpoints. The strands of DNA fragments labeled at their 5′ ends with ^{32}P were separated as described by Maxam and Gilbert (26). Total cellular RNA from cells grown to an A_{550} of 0.6 to 0.8 in minimal medium was isolated by the method of Cooper et al. (9), except that spheroplasts in 2 ml of buffer (20 mM Tris hydrochloride [pH 7.5], 0.2 M NaCl, 5 mM EDTA) at 4°C were lysed by addition of an equal volume of 65°C buffer containing 1% sodium dodecyl sulfate. Contaminating DNA was removed with diethylpyrocarbonate-treated DNase I as described by Smith and Calvo (38).

RNA was synthesized in vitro as follows. Plasmid DNA (0.5 to 1.0 pmol) and RNA polymerase (a gift of Lester Lau; 1 pmol) were added to a reaction mixture containing 20 mM Tris acetate (pH 7.9), 0.1 mM EDTA, 0.1 mM DTT, 50 mM KCl, and 4 mM magnesium acetate in a total volume of 100 μl. After incubation at 37°C for 10 min, the reaction was started by adding nucleoside triphosphates (100 μM GTP, 100 μM ATP, 75 μM UTP, 75 μM CTP [final concentrations]) and rifampin to 5 μg/ml. After incubation at 37°C for 10 min, the reaction was stopped by adding 200 μl of 0.45 M sodium acetate (pH 5.6), 10 μg of *E. coli* tRNA, and 300 μl of buffer-equilibrated phenol. The aqueous phase was extracted with ether, and nucleic acids were precipitated by addition of three volumes of ethanol.

The S1 nuclease mapping techniques of Berk and Sharp (4) as modified by Favaloro et al. (16) were used. Hybridizations were for 3 h at 30°C, and S1 nuclease digestions were carried out for 2 h at 14°C.

RNase mapping of endpoints of pulse-labeled RNA. A 20-ml culture of strain ID86 was grown to an A_{550} of 0.6 in 125 ml of MOPS medium at 32°C. Cells were centrifuged at 10,600 × *g* for 10 min at 4°C, washed twice with 8 ml of MOPS lacking phosphate, and then resuspended in the same volume of medium either with or without 50 μg of leucine per ml. Cultures were incubated for 15 min at 32°C and pulse-labeled for 3 min by addition of carrier-free P_i (1 mCi; Amersham). Labeling was stopped by mixing the culture with 8 ml of ice-cold MOPS containing 10 mM NaN_3 and 400 μg of chloramphenicol per ml. Total cellular RNA was isolated as described above.

The procedure for RNase mapping of RNA endpoints was a modification of the technique described by Salditt-Georgieff and Darnell (35). Pulse-labeled [^{32}P]RNA (10 to 50 μg, 2×10^5 to 3×10^5 cpm/μg; RNA was brought to a total of 50 μg with *E. coli* tRNA) isolated from strain ID86 grown in the presence or absence of leucine (50 μg/ml) and 3 to 6 μg of M13 DNA (plus strand; approximately 1 pmol) were hybridized at 30°C for 3 h under conditions described by Favaloro et al. (16) for S1 nuclease mapping. Nucleic acid was then treated with 40 μg of RNase A and 40 μg of RNase T1 per ml in 300 μl of buffer (0.1 M NaCl, 10 mM Tris hydrochloride [pH 8], 1 mM EDTA) for 30 min at 30°C. The reaction mixture was extracted once with an equal volume of phenol and once with an equal volume of ether, and 6 μl of 10% sodium dodecyl sulfate was added. Intact M13 DNA (including DNA-RNA hybrids) was purified by centrifuging the 300-μl sample through a 4-ml 15 to 30% sucrose gradient (sucrose in 10 mM Tris hydrochloride [pH 8.0] plus 1 mM EDTA) in a Beckman SW60 Ti rotor at 40,000 rpm for 5 h. Nucleic acid hybrids, recovered by ethanol precipitation after addition of 10 μg of *E. coli* tRNA, were examined by one or both of the following procedures. Samples were resuspended in 10 μl of formamide containing 0.04% xylene cyanol and 0.04% bromophenol blue, heated at 100°C for 1 min, and fractionated on a 5 or 10% polyacrylamide–50% urea gel in Tris borate buffer for 3 to 4 h at 1,000 V. Alternatively, RNA-DNA hybrids were treated with 300 μl of S1 nuclease buffer under conditions described for S1 mapping of RNA endpoints, resuspended in buffer (10 mM Tris hydrochloride [pH 8.3], 10% glycerol, 0.04% bromophenol blue, 0.04% xylene cyanol), and fractionated on a 5% polyacrylamide gel in Tris borate buffer for 3 to 4 h at 1,000 V. Results were visualized by autoradiography with Kodak XAR-5 film.

Construction of plasmids and bacteriophages. Plasmids pCV52 and pCV66 were constructed by blunt-end ligation of fragment *Hinc*II$_{1733}$ (Fig. 1) into the *Sma*I site of plasmid pK01. Plasmids pCV53 and pCV65, containing fragment *Hae*III$_{361}$, were constructed in the same way. Plasmids pCV89 and pCV101 were similarly constructed, except that the ends of fragments *Hin*fI-*Hae*III$_{178}$ and *Hae*III-*Sau*3A$_{312}$ were made blunt by treatment with avian myeloblastosis virus reverse transcriptase and nucleoside triphosphates before ligation. The desired plasmids from ampicillin-resistant transformants of strain W3102 were identified by restriction analysis.

Derivatives of phage M13mp8 were constructed as follows. Fragments *Hin*fI$_{457}$, *Hin*fI-*Nde*I$_{349}$, *Nde*I-*Hin*fI$_{110}$, *Hae*III-*Sau*3A$_{312}$ (ends made blunt with reverse transcriptase), and *Hae*III$_{174}$ were separately cloned into the *Sma*I site of M13mp8 RF DNA. Fragment *Sau*3A$_{89}$ was cloned into the *Bam*HI site of M13mp8 RF DNA. Recombinant bacteriophage were identified after transfection of strain JM103 by their white plaque phenotype (29), and the desired isolates were identified by restriction analysis and by the Sanger sequencing procedure.

For construction of deletions, plasmid pCV53 DNA was linearized by digestion with *Hin*dIII and treated with nuclease BAL31 at 20°C for 40 min (5 to 10 bp digested per min per end of DNA). Aliquots were taken at 1-min intervals, frozen on dry ice, and pooled. BAL31 nuclease was removed by extraction with phenol, and after treatment with reverse transcriptase and nucleoside triphophates to make ends blunt, plasmids were recircularized by blunt-end ligation. The sample was treated with *Hin*dIII to eliminate undigested molecules before transforming strain W3102 and selecting colonies resistant to ampicillin. Plasmid DNA was isolated by the method of Birnboim and Doly (5) and screened by restriction analysis to estimate deletion size. Deletion endpoints were determined by nucleotide sequence analysis by the method of Maxam and Gilbert (26).

Mutants with increased expression of *ilvIH*. Plasmid pCV7 DNA was mutagenized in vitro by the procedure of Hashimoto and Sekiguchi (18) by incubating it for 25 h at 37°C in 0.4 M freshly prepared hydroxylamine containing 50

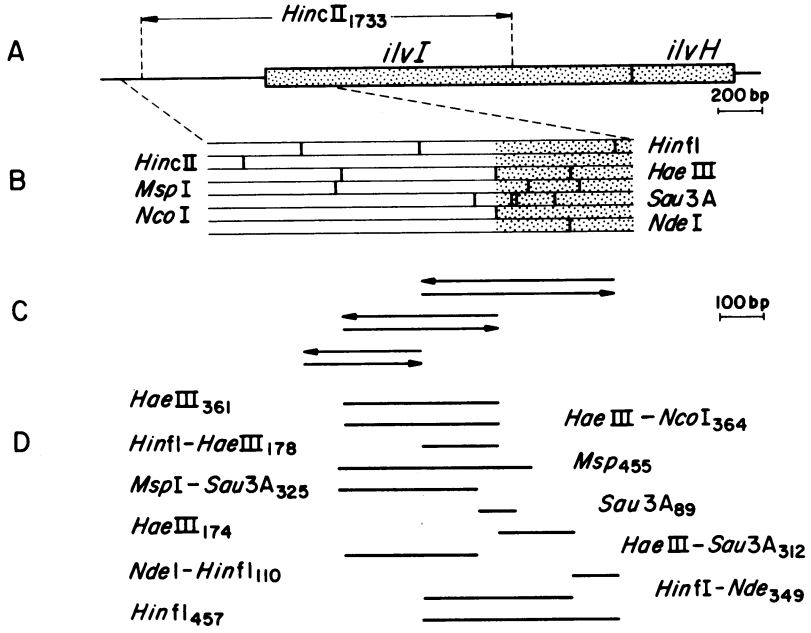

FIG. 1. Restriction map of *ilvIH*, DNA fragments used, and sequencing strategy. (A) Diagram of *ilvIH*. Transcription is rightward. (B) A more detailed restriction endonuclease map of the promoter-proximal end of *ilvI* and upstream sequences. The restriction sites were deduced from sequence data and by standard mapping techniques and have been presented, in part, elsewhere (41). (C) Sequencing strategy. Arrows indicate the fragments sequenced and point away from the 3'-labeled ends. (D) Restriction endonuclease fragments used for S1 nuclease mapping, DNA sequencing, or insertion into expression vector pKO1 or bacteriophage M13mp8. Panels B to D are drawn to the same scale.

mM sodium phosphate (pH 6.0) and 1 mM EDTA. Under these conditions, the ability of the DNA sample to transform cells to ampicillin resistance was reduced by 90%. After the sample was dialyzed (18), it was used to transform strain MI316, with selection being made simultaneously for isoleucine and valine prototrophy, resistance to leucine, and resistance to ampicillin (minimal agar plates contained 250 μg of leucine and 100 μg of ampicillin per ml and were incubated at 32°C). An amount of unmutagenized DNA that gave 10,000 ampicillin-resistant transformants yielded 5 to 10 leucine-resistant mutants after mutagenesis. DNA prepared from single colonies (5) was used to retransform strain MI316 under the same conditions to identify those mutants in which leucine resistance was caused by a mutation on the plasmid.

RESULTS

Nucleotide sequence of the region upstream of *ilvI*. The nucleotide sequences of *ilvI* and of *ilvH* were reported earlier (41). The sequence of a 365-bp region immediately upstream of *ilvI* was determined by the procedures of Maxam and Gilbert (26) (Fig. 1 and 2). Lago et al. (21) established the translation initiation codon as the AUG at position +1. The region immediately upstream of *ilvI* is rich in AT base pairs; for example, sequences from −123 to −273 contain 71% AT.

Localization of the *ilvIH* control region. A 1,733-bp *Hin*cII fragment (Fig. 1) containing part of *ilvI* and 592 bp upstream from *ilvI* provided a promoter for *galK* on plasmid pKO1 (28) (Table 2 [plasmid pCV52 row, −Leu column]). This promoter activity, which was only observed for fragment *Hin*cII$_{1733}$ in the proper orientation relative to *galK* (Table 2, compare plasmids pCV52 and pCV66 [−Leu column]), was severalfold higher than that conferred by the *lacUV5* promoter (Table 2, compare plasmids pCV52 and pKO110). Among *Hae*III fragments derived from fragment *Hin*cII$_{1733}$, only a single fragment (*Hae*III$_{361}$ [Fig. 1]) had promoter activity when cloned into plasmid pKO1. A strain carrying plasmid pCV53 (*Hae*III$_{361}$ oriented in the same direction as *galK*) produced high levels of galactokinase activity, whereas the same strain containing plasmid pCV65 (*Hae*III$_{361}$ oriented in the opposite direction) had no galactokinase activity (Table 2).

For some strains of *E. coli*, expression of the *ilvIH* operon is repressed about 8- to 10-fold when cells are grown in the presence of leucine, whereas in other strains, repression is 2- to 3-fold (12, 42). For the experiment described in Table 2, galactokinase assays were performed on cells grown both with and without exogenous leucine. Leucine caused a two to threefold repression of galactokinase activity in W3102 strains carrying plasmid pCV52 or pCV53 but no repression of galactokinase activity in strains carrying plasmid pKO1 (no promoter) or pKO110 (*lacUV5* promoter; Table 2). The level of galactokinase in a strain carrying plasmid pCV53 was not significantly repressed when cells were grown in the presence of 50 μg of isoleucine, proline, or threonine per ml (data not shown).

The copy number of plasmids such as pKO1 that carry the ColE1 replicon can decrease more than 10-fold with the insertion of some promoters (1, 43). We measured the copy number of pKO1-derived plasmids relative to plasmid pSE150, a low-copy-number plasmid derived from plasmid pDPT427 (R plasmid replicon [40]). The relative copy number of pKO1-derived plasmids did not vary greatly, usually less than twofold, in cells grown with or without leucine (Table 2).

The data in Table 2 indicate that both the *ilvIH* promoter and sequences involved in leucine regulation of *ilvIH* are contained on a 361-bp *Hae*III fragment. Further, the data suggest that leucine exerts an effect upon some aspect of *ilvIH* transcription because (i) the *galK* gene provides all signals required for translation of galactokinase (28) and (ii)

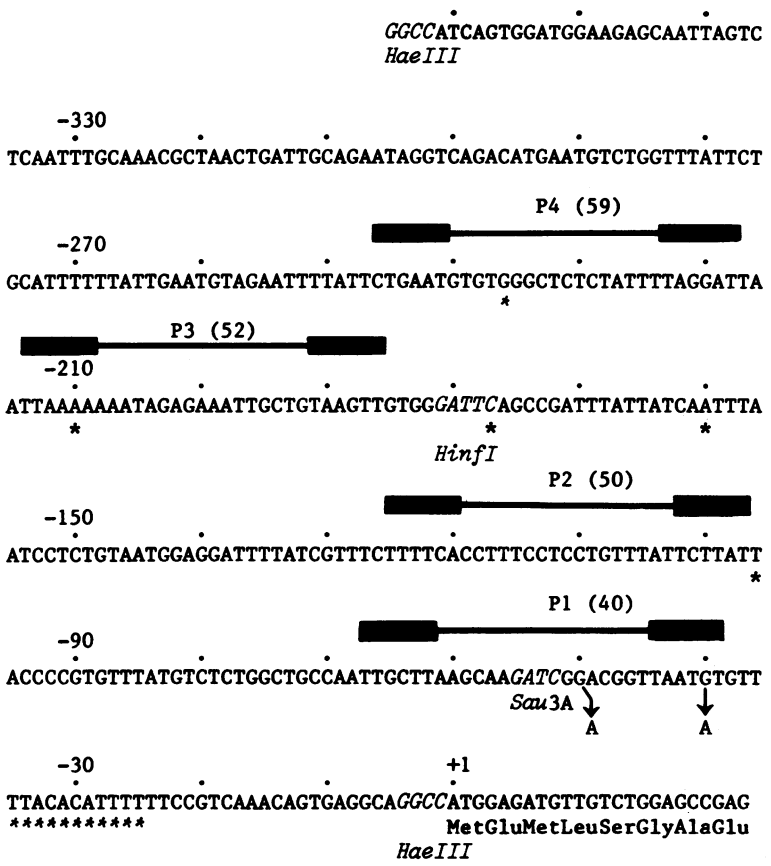

FIG. 2. Nucleotide sequence of the region upstream of *ilvI*. Position 1 marks the first nucleotide of the translational start codon of *ilvI*. The first eight amino acids of *ilvI* determined from polypeptide sequencing (21) are shown. Lines join bars identifying the positions of −10 and −35 regions of putative promoters (Fig. 7) with a homology rating greater than or equal to 50 (31). An additional promoter, P1, is also shown. An asterisk underscores the positions of the major 5′ endpoints of RNA transcribed in vitro (−210, −177, −160, and −96) or in vivo (−236, −177, −96, and −35 to −25) as determined by S1 nuclease mapping. Arrows pointing downward identify mutations that increase expression of the *ilvIH* operon. The positions of relevant restriction sites are italicized.

the *Hae*III$_{361}$ fragment only contains sequences that lie upstream of *ilvI*.

5′ endpoints of in vivo *ilvIH* mRNA established by S1 nuclease mapping. For S1 nuclease mapping (4, 16), separated strands of fragments labeled at their 5′ ends were employed as probes. Total cellular RNA was isolated from CV696 [PS1283(pCV35)] or CV670 [MI316(pCV7)], strains containing *ilvIH* on multicopy plasmids. When RNA isolated from cells grown in the absence of leucine was hybridized with the coding strand of fragment *Msp*I$_{455}$ (Fig. 1), the most prominent species observed after S1 nuclease treatment had sizes from 104 to 114 bases, representing 5′ RNA endpoints at positions −25 to −35 (Fig. 2; Fig. 3A, lane 2; and Fig. 3B, lane 3). Additional bands of length greater than 114 bases

TABLE 2. Specific activity of galactokinase and relative plasmid copy number in strains carrying pK01-derived plasmids

Strain[a]	Plasmids	Promoter	Orientation with respect to *galK*	Galactokinase activity[b]		Relative plasmid copy number[c]	
				−Leu	+Leu	−Leu	+Leu
CV781	pK01, pSE150	None		13 (1.5)	32 (3.6)	8.4	8.8
CV782	pK0110, pSE150	*lacUV5*	Toward	415 (45)	652 (51)	9.2	12.8
CV816	pCV52, pSE150	*ilvIH*	Toward	891 (72)	311 (28)	12.4	11.2
CV817	pCV66, pSE150	*ilvIH*	Away from	9 (1.5)	9 (1.0)	5.9	8.6
CV787	pCV53, pSE150	*ilvIH*	Toward	1,659 (157)	683 (54)	10.6	12.7
CV788	pCV65, pSE150	*ilvIH*	Away from	57 (9)	28 (4)	6.6	7.0
CV814	pCV101, pSE150	*ilvIH*	Toward	16 (2.9)	90 (12)	5.5	7.6

[a] The indicated plasmids are carried by host strain W3102 *galK*.

[b] Galactokinase-specific activity is expressed in nanomoles of galactose phosphate per minute per milligram of protein. −Leu and +Leu refer to growth in the absence or presence respectively, of 50 μg of L-leucine per ml. Specific activities are the average of two experiments, each done in duplicate. The numbers in parentheses represent the specific activity divided by the plasmid copy number.

[c] Relative plasmid copy number is the number of pK01-derived plasmid molecules per molecule of pSE150.

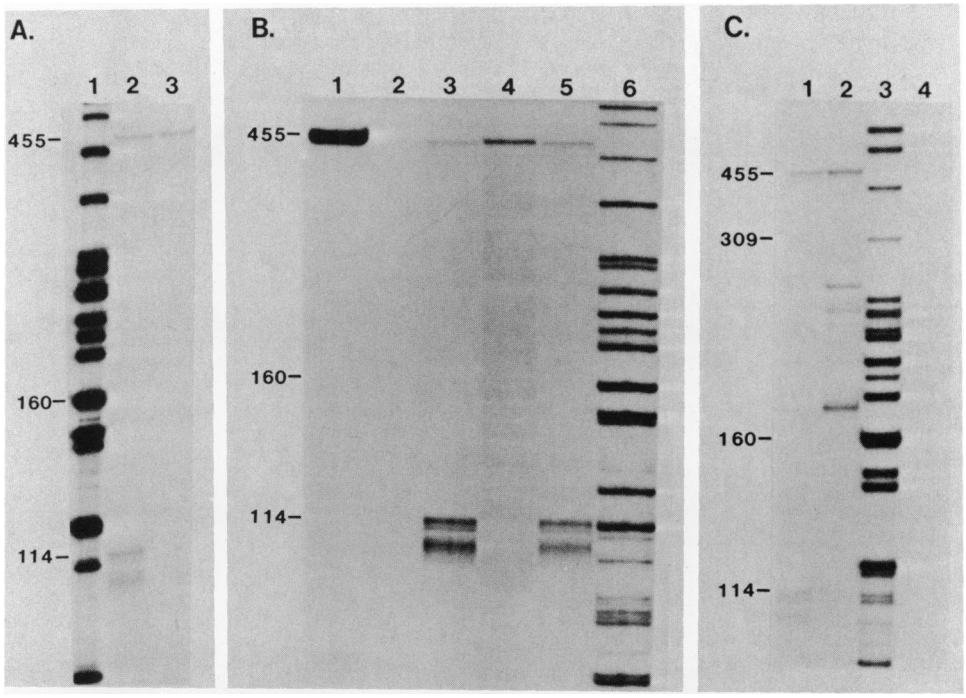

FIG. 3. S1 nuclease mapping of 5′ endpoints of *ilvIH* RNA. RNA was hybridized with the coding strand of 5′ end-labeled fragment $MspI_{455}$ (10,000 to 20,000 cpm; Fig. 1). (A) Lanes: 2 and 3, hybridization with 5 μg of total cellular RNA from strain CV670 [MI316(pCV7)] grown in the absence (lane 2) or presence (lane 3) of leucine; 1, 3′ end-labeled *Msp*I fragments from plasmid pBR322. (B) Lanes: 1, labeled coding strand of $MspI_{455}$; 2, hybridization with 25 μg of *E. coli* tRNA; 3 and 5, hybridization with 5 μg of total cellular RNA from strain CV696 [PS1283(pCV35)] grown in the absence (lane 3) or presence (lane 5) of leucine; 4, fragments of $MspI_{455}$ generated by A + G chemical cleavage reactions (26); 6, 3′ end-labeled *Msp*I fragments from pBR322. (C) Lanes: 1, hybridization with 5 μg of total cellular RNA from strain CV696 [PS1283(pCV35)] grown in the absence of leucine; 2 and 4, hybridization with one-third of the RNA transcribed from 1 pmol of plasmid pCV52 (lane 2) or pKO1 (lane 4); 3, 3′ end-labeled *Msp*I fragments from plasmid pBR322. Autoradiography was for 1 week (A), 3 days (B), or 2 days (C).

(175, 255, and 314 bases, representing RNA endpoints at positions −96,–177, and −236, respectively [Fig. 2]) were seen after longer exposure (data not shown). Bands corresponding to a full-length probe may represent hybridization of readthrough RNAs whose synthesis was initiated at upstream plasmid promoters. When the noncoding strand of fragment $MspI_{455}$ was used as a probe (data not shown), or when *E. coli* tRNA replaced bacterial RNA (Fig. 3B, lane 2), no bands were observed. Similarly, when the experiments described above were repeated with fragment *Msp*I/*Sau*3A$_{325}$ as a probe (lacking 52 bp immediately upstream of *ilvI* [Fig. 1]), no major RNA endpoints were found, although the minor endpoints at positions −96, −177, and −236 were observed (data not shown). RNA from strain PS1035 (single chromosomal copy of *ilvIH*), when probed with the coding strand of fragment $MspI_{455}$, produced S1 nuclease mapping results indistinguishable from those observed from CV696 RNA except that the signal for the same amount of RNA was weaker (data not shown).

The effect of L-leucine on the frequency and distribution of 5′ endpoints of *ilvIH* mRNA was investigated by carrying out S1 nuclease mapping experiments with RNA isolated from strains grown with and without leucine. For plasmid-containing strains CV696 [PS1281(pCV35)] and CV670 [MI316(pCV7)], leucine caused 2-to 3-fold and 5- to 10-fold decreases, respectively, in the number of RNA molecules hybridizing to the coding strand of fragment $MspI_{455}$ (Fig. 3A, lanes 2 and 3, and 3B, lanes 3 and 5). On the other hand, leucine had no effect upon the distribution of endpoints, i.e., the major endpoints all map to the region −25 to −35, irrespective of whether cells were grown in the presence or absence of leucine (the relevant bands in Fig. 3A, lane 3, are difficult to see but are evident in the radioautogram). The same results were obtained from haploid strain PS1035 grown in the presence of leucine except that the corresponding bands were weaker (data not shown). These results complement those obtained previously by Squires et al. (42), who observed that the rate of synthesis of *ilvIH* mRNA was reduced 8- to 10-fold by growing strain CV670 [MI316(pCV7)] in the presence of leucine.

5′ endpoints of *ilvIH* mRNA transcribed in vitro. Several plasmid DNAs were transcribed in vitro with purified RNA polymerase, and the products were analyzed by S1 nuclease mapping. For RNA transcribed in vitro from plasmid pCV52 (a pKO1 derivative containing *Hin*cII$_{1733}$) and hybridized to the coding strand of fragment $MspI_{455}$ (Fig. 1), only a small proportion of the endpoints corresponded to the major RNA endpoints detected in vivo (Fig. 3C, compare lanes 1 and 2; band sizes 104 to 114 corresponded to positions −25 to −35 in Fig. 2). Most transcriptions initiated at positions upstream of −35. The four most prominent bands of approximately 175, 240, 255, and 288 bases represented RNA molecules initiating at positions −96, −160, −177, and −210 (Fig. 2). The two most abundant of these species (175 and 255 bases, representing initiation at positions −96 and −177) corresponded exactly in mobility with two of the minor RNA species detected by S1 mapping of in vivo RNA (these bands are too faint to be seen in Fig. 3C, lane 1). As expected, when RNA transcribed from plasmid pKO1 was hybridized to the coding strand of fragment $MspI_{455}$ (Fig. 3C, lane 4) or

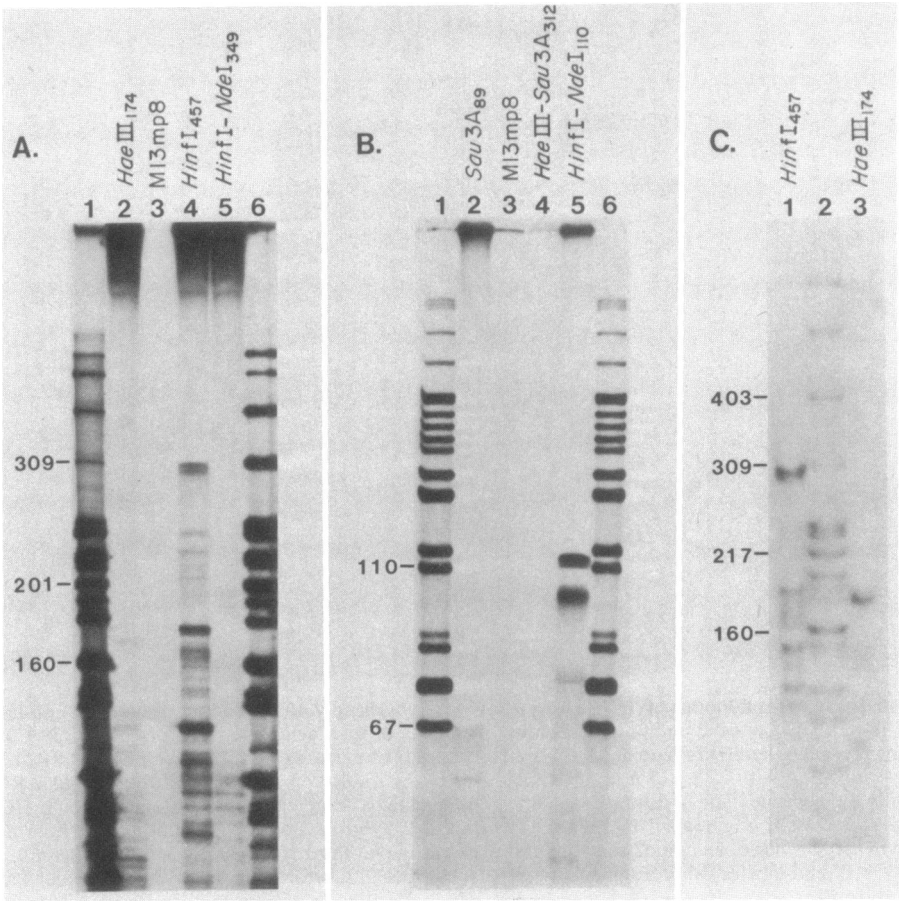

FIG. 4. Protection of pulse-labeled RNA from RNase degradation by hybridization to DNA probes. ^{32}P-pulse-labeled RNA was hybridized with the indicated probes (Fig. 1), and the products were treated with RNase and fractionated on a sucrose gradient as described in the text. Samples were denatured and fractionated on a 5% polyacrylamide–50% urea (A) or 10% polyacrylamide–50% urea (B) gel at 1,000 V for 4 h. Alternatively, samples were treated with S1 nuclease and fractionated at 1,000 V for 4 h on a nondenaturing 5% polyacrylamide gel (C). Autoradiography was for 30 h. RNA or RNA-DNA hybrid fragment lengths were estimated by comparison with fragments of pBR322 cut with *Msp*I and labeled at the 3′ ends with avian myeloblastosis virus reverse transcriptase (size standards). (A) Lanes: 1 and 6, size standards; 2 to 5, hybridizations employing probes M13GH4 (*Hae*III$_{174}$), M13mp8, M13GH1 (*Hin*fI$_{457}$), and M13GH2 (*Hin*fI-*Nde*I$_{349}$), respectively. (B) Lanes: 1 and 6, size standards; 2 to 5, hybridizations employing probes M13GH6 (*Sau*3A$_{96}$), M13mp8, M13GH5 (*Hae*III-*Sau*3A$_{312}$), and M13GH3 (*Hin*fI-*Nde*I$_{110}$), respectively. (C) Lanes: 2, size standards; 1 and 3, hybridizations employing M13GH1 (*Hin*fI$_{457}$) and M13GH4 (*Hae*III$_{174}$), respectively.

when the noncoding strand of fragment *Msp*I$_{455}$ was hybridized with any of the RNA samples described above (data not shown), no major bands were observed. All of the RNA endpoints detected upstream of position −51 (Fig. 2 and Fig. 3C, lane 2) were also detected when the coding strand of fragment *Msp*I-*Sau*3A$_{325}$ (Fig. 1) was hybridized to RNA transcribed in vitro from plasmid pCV52 (data not shown).

In summary, most *ilvIH* mRNAs formed in vivo have endpoints 25 to 35 bp upstream of *ilvI* and the amounts of these RNAs decrease when cells are grown in the presence of leucine. By contrast, most synthesis of RNA from this region in vitro is initiated at positions upstream of −35 at sites corresponding to endpoints of minor RNA species detected in vivo.

Analysis of 5′ endpoints of pulse-labeled *ilvIH* mRNA. S1 nuclease mapping experiments of the type described by Berk and Sharp (4) employ total unlabeled RNA and, as such, give information on endpoints of a steady-state population of molecules. The observed in vivo 5′ endpoints of *ilvIH* mRNA could be due to transcription initiation or to processing of transcripts whose synthesis was initiated upstream. To further investigate these two possibilities, the position of 5′ endpoints of *ilvIH* mRNA pulse-labeled in vivo was determined, following a strategy employed by Salditt-Georgieff and Darnell (35). For short pulse times, this method offers a better chance of detecting unprocessed precursor molecules, if such molecules exist.

Total cellular RNA of strain ID86 was pulse-labeled with ^{32}P for 3 min, a short period considering the time required for uptake and incorporation into the α-phosphate position of ATP (6). The pulse-labeled total cellular RNA was hybridized to single-stranded phage M13mp8 DNA containing sequences from *ilvIH* (Fig. 1), and the products were digested with RNase. M13 DNA (including DNA-RNA hybrids) was purified on sucrose gradients, and the DNA-RNA hybrids were denatured and fractionated on polyacrylamide gels (Fig. 4A and B). No RNA species were protected when M13mp8 alone was used as a probe (Fig. 4A, lane 3 and 4B, lane 3) or when the probe carried the noncoding strand of fragment *Hin*fI$_{457}$ (Fig. 1) (data not shown). To further test the reliability of these procedures, probes containing bacterial DNA entirely from *ilvI* were employed (Fig. 1, *Hae*III$_{174}$

and *Nde*I-*Hin*fI$_{110}$). The highest-molecular-weight species of RNA protected by these probes corresponded to the full-length probes, i.e., 174 and 110 bases, respectively (Fig. 4A, lane 2, and 4B, lane 5).

The highest-molecular-weight species of RNA protected by probes carrying the *ilvIH* coding strands of fragments *Hin*fI$_{457}$ and *Sau*3A$_{89}$ (Fig. 1) were RNA species smaller than the full length of the *ilvIH* fragment (306 and 67 bases, respectively [Fig. 4A, lane 4, and 4B, lane 2]), suggesting that the 5′ endpoints of the mRNA lie within these fragments. Fragment *Hin*fI-*Nde*I$_{349}$ differs from *Hin*fI$_{457}$ by lacking the 108 bp downstream of the *Nde*I site (Fig. 1). The highest-molecular-weight species of RNA protected by probes carrying fragments *Hin*fI$_{457}$ and *Hin*fI-*Nde*I$_{349}$ also differ by 108 bases (306 versus 198 bases, respectively [Fig. 4A, compare lanes 4 and 5]). This result indicates that a contiguous stretch of probe is protected, beginning at the 5′ end of the probe. From this information we estimate that the 5′ endpoints of pulse-labeled RNA lie at approximately position −30 (Fig. 2).

The same analysis was carried out with fragment *Hae*III-*Sau*3A$_{312}$ (Fig. 1) which contains upstream DNA except for 52 bp immediately adjacent to *ilvI*. This probe did not protect any species of RNA. Thus, under the conditions of these experiments, we were unable to detect transcription initiated upstream of position −50 (Fig. 4B, lane 4).

Beneath the higest-molecular-weight RNA species protected by the M13 probes are a number of obvious additional bands (Fig. 4A, lanes 2, 4, and 5, and 4B, lanes 2 and 5). Moreover, the highest-molelcular-weight species of RNA protected was not usually the predominant one. We suspected that these bands did not represent transcripts with initiation points further downstream, but rather were degradation products created either in vivo or in vitro. To test this hypothesis, the experiment described above was repeated except that the nucleic acids purified on sucrose gradients were treated with S1 nuclease to remove single-stranded M13 DNA and fractionated on polyacrylamide gels under nondenaturing conditions. Under such conditions, most nicked RNA molecules should be held together by hybridization to DNA. The results of this experiment with probes containing fragments *Hin*fI$_{457}$ and *Hae*III$_{174}$ are shown in Fig. 4C,lanes 1 and 3. Unlike the results shown in Fig. 4A, the highest-molecular-weight species of RNA protected were in the predominant ones, supporting the hypothesis given above (306 and 174 bases protected from *Hin*fI$_{457}$ and *Hae*III$_{174}$, respectively). Further, the results strongly suggest that the RNA degradation observed in Fig. 4A and B did not occur in the cell or at any time before hybridization. Had degradation occurred before hybridization, two RNA fragments would not likely have hybridized to the same DNA molecule, given that DNA was in great excess in these experiments.

The effect of L-leucine on the frequency and distribution of *ilvIH* mRNA 5′ endpoints was investigated by similar techniques. ^{32}P-pulse-labeled RNA isolated from strain ID86 grown with and without leucine was hybridized to M13 probes containing fragments *Hin*fI$_{457}$, *Hin*fI-*Nde*I$_{349}$, and *Hae*III$_{174}$. Leucine had no effect on the distribution of *ilvIH* mRNA endpoints, i.e., the endpoints mapped to position −30 for either condition of growth, but the extent of the signal was 5-10 fold lower for samples isolated from cells grown in the presence of leucine (data not shown). These results complement those obtained from S1 mapping studies using unlabeled RNA.

Mutations causing elevated expression of AHAS III. Strain CV670 [MI316(pCV7)] contains low AHAS activity because the *ilvIH* operon on plasmid pCV7 is missing about 100 base pairs of DNA from the promoter-distal end of *ilvH* (21, 42). The low AHAS activity is further reduced by leucine with the consequence that strain CV670 does not grow in a leucine-containing medium. This phenomenon was exploited to select leucine-resistant mutants from strain MI316 (lacking AHAS activity because of mutations in *ilvB* and *ilvI*) transformed with hydroxylamine-treated pCV7 DNA. AHAS levels in eight leucine-resistant mutants were significantly higher than in strain MI316 carrying unmutagenized plasmid pCV7 (Table 3). The increase in AHAS activity varied from about 2.5-fold for strain ID63 to 33-fold for strain ID64. For most strains, addition of leucine to the growth medium caused a 1.5- to 2.5-fold repression effect, but for two strains, ID64 and ID65, the repression effect was 6- to 8-fold. As expected, valine did not inhibit AHAS activity in extracts prepared from these strains (42).

TABLE 3. AHAS activity in leucine-resistant derivatives of strain CV670

Strain[a]	Plasmid	Mutation	Addition of minimal medium[b]	Sp act of AHAS[c]	
				−Valine	+Valine
CV670	pCV7	None	None	5.8	5.4
ID60	pCV93	*ilv-758*	None	32	32
			Leucine	19	17
ID61	pCV94	*ilv-753*	None	27	26
			Leucine	11	9.0
ID62	pCV95	*ilv-752*	None	20	22
			Leucine	8.2	8.0
ID63	pCV96	*ilv-754*	None	15	15
			Leucine	5.4	6.2
ID64	pCV97	*ilv-755*	None	191	176
			Leucine	32	27
ID65	pCV98	*ilv-756*	None	107	101
			Leucine	13	13
ID66	pCV99	*ilv-757*	None	28	29
			Leucine	13	12
ID69	pCV100	*ilv-759*	None	30	28
			Leucine	14	13

[a] All strains are plasmid-containing derivatives of strain MI316 (*ilvB619 ilvH612 ilvI614*). Strain MI316 and its derivatives are temperature sensitive and therefore were grown at 32°C.

[b] Where indicated, the medium contained L-leucine at 50 μg/l.

[c] Specific activity is micromoles of acetolactate per minute per milligram of protein. Assays were carried out in the absence of valine and, separately, with 0.5 mM valine.

For each of the strains identified in Table 3, the mutation leading to elevated AHAS activity was shown to be plasmid linked by transforming strain MI316 with isolated plasmid DNAs and observing elevated AHAS activity in the resulting strains. For six of the strains, plasmid DNAs were isolated, and sequences of the *Hae*III-*Nco*I$_{364}$ fragments derived from them (Fig. 1) were determined. The fragments from plasmids pCV97, pCV98, and pCV99 (strains ID64, ID65, and ID66, respectively) had the same sequence as that of the wild type. The positions of the mutations giving rise to elevated AHAS levels have not been located on these plasmids. Plasmids

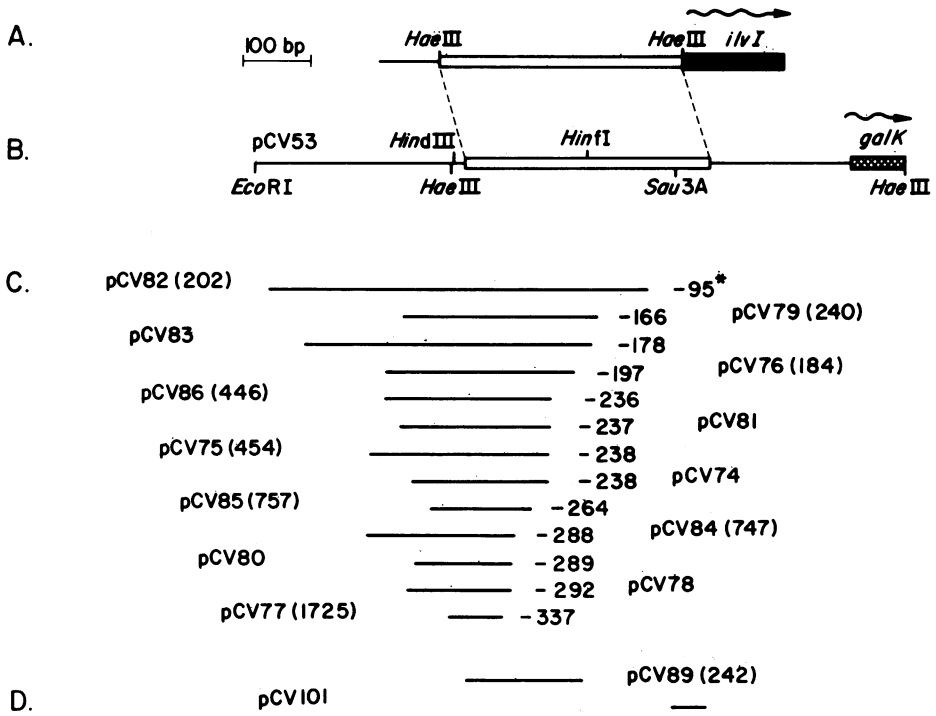

FIG. 5. Deletions of plasmid pCV53 constructed in vitro. (A) Promoter region of the *ilvIH* operon on plasmid pCV7. The shaded rectangle represents *ilvI* polypeptide coding sequences. (B) Region of plasmid pCV53 upstream of the *galK* gene. The cross-hatched rectangle represents *galK* polypeptide coding sequences. The open rectangle represents the *Hae*III fragment from plasmid pCV7 cloned into plasmid pKO1 upstream of *galK*. (C) Deletions created by BAL31 digestion of *Hin*dIII-cut pCV53. The lines represent the extent of DNA deleted. The numbers to the right of the deletions represent the positions in Fig. 2 of the last base deleted. Deletion endpoints were determined by sequencing *Eco*RI-*Hae*III fragments (the *Hae*III site used lies within *galK*; note that the *Hae*III site close to the *Hin*dIII site is removed by the deletions). The asterisk indicates that for plasmid pCV82, the exact deletion endpoint within *ilvIH* DNA is within 8 bp of the number shown. Numbers in parentheses are the averages of uncorrected galactokinase activities for the three experiments in Fig. 6A. (D) Deletions created by cloning fragments *Hae*III-*Sau*3A$_{312}$ and *Hin*fI-*Hae*III$_{178}$ (Fig. 1) into pKO1. The lines represent the part of pCV53 that was deleted. Panels are all drawn to the same scale. Wavy arrows indicate the direction of transcription.

pCV94 (strain ID61) and pCV95 (strain ID62) had the same mutation, a G→A transition at position −40, and plasmid pCV96 (strain ID63) had a G→A transition at position −50 (Fig. 2).

Deletion analysis of the *ilvIH* control region. A nested set of deletions extending into the region upstream of *ilvI* was created by BAL31 treatment of linearized plasmid pCV53 (Fig. 2 and 5). The effects of these deletions on transcription of *galK* in vivo were evaluated by assaying galactokinase activity in plasmid-containing strains grown in the absence of leucine (Fig. 6A). A deletion removing 27 bp had no effect upon expression of *galK* (Fig. 6A and C; plasmid pCV77). Deletions removing more than 27 bp had progressively more severe effects upon *galK* expression (Fig. 6A and C; plasmids pCV84, pCV85, pCV75, pCV86, and pCV76). A reduction of more than 80% was observed for deletions of 165 or more base pairs of chromosomal sequences.

For all the cases described above, which represent deletions coming from the left, a level of promoter activity remained that was at least five times the background activity observed for promoterless plasmid pKO1 (Table 2). For the most extensive of these deletions, that in plasmid pCV82, only 94 bp of DNA upstream of *ilvI* remained. On the other hand, a deletion created by subcloning that removed 49 bp of DNA immediately upstream of *ilvI* (Fig. 5C; plasmid pCV101) reduced *galK* expression to background levels (Table 2).

For two reasons the changes in galactokinase levels associated with these deletions are not likely a result of changes in the vector sequences. First, variations in galactokinase levels were correlated with the number of base pairs of *ilvIH* sequences, and not of vector sequences deleted. For example, plasmid pCV84 with 140 bp of vector sequences deleted gave relatively high galactokinase activity, whereas plasmid pCV79 with 90 bp of vector sequences deleted gave low galactokinase activity (Fig. 5C). Second, the deletion in plasmid pCV89 (produced by subcloning rather than by BAL31 digestion; therefore, no vector sequences were deleted [Fig. 5C]) caused a drop in *galK* expression that was comparable to that caused by deletions from the nested set that had similar endpoints within *ilvIH* sequences (Fig. 5C and 6A; compare results for plasmids pCV76, pCV79, and pCV89). Furthermore, the results are not accounted for by variations in plasmid copy number because the shape of the curve in Fig. 6A is the same for galactokinase activity uncorrected or corrected for plasmid copy number.

The effect of growth in the presence of leucine upon *galK* expression is shown in Fig. 6B. For growth in the presence of leucine, *galK* was expressed at a low, near constant level, irrespective of the extent of the deletion. Note that the repressed level was about the same as that observed for strains grown in the absence of leucine but carrying plasmids with extensive deletions (Fig. 6A and C; plasmids pCV76, pCV89, pCV79, and pCV82).

Three conclusions can be drawn from these data. (i) A

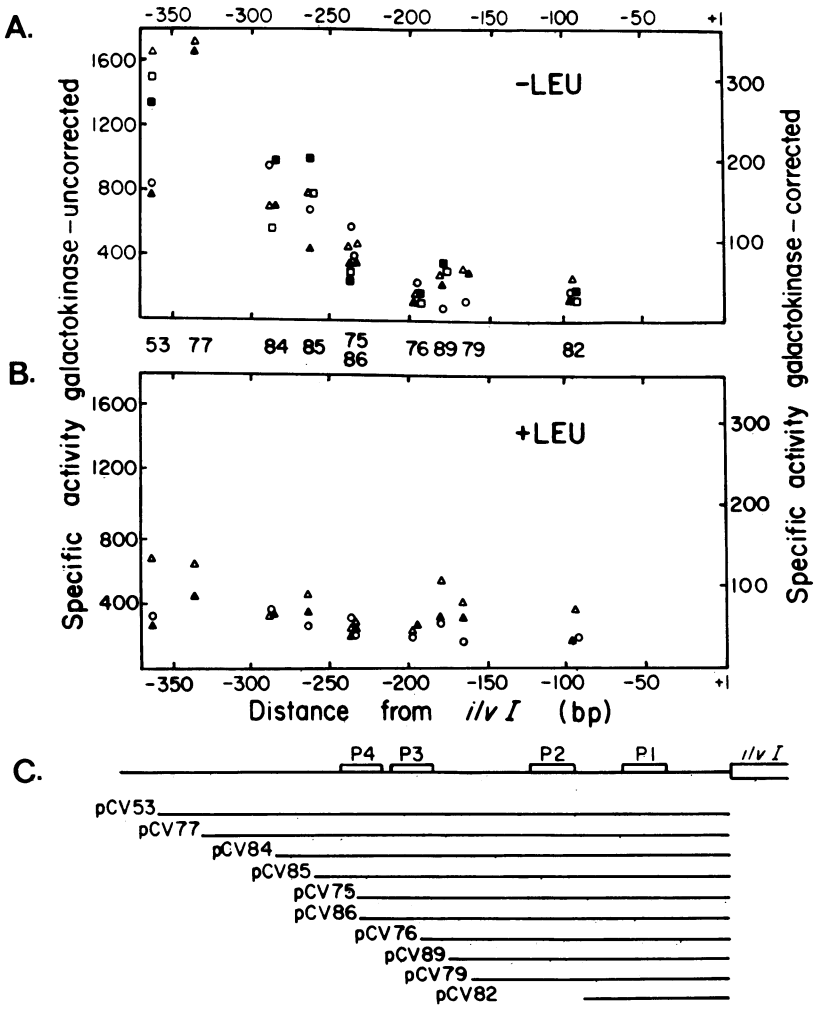

FIG. 6. Specific activity of galactokinase in strains containing deletion derivatives of plasmid pCV53. Numbers on the abscissa in panels A and B represent the distance upstream from *ilvI* and correspond to the numbering in Fig. 2. The numbers between panels A and B identify deletion endpoints that are represented more completely in panel C and in Fig. 5. (A) Strains grown in the absence of leucine. (B) Strains grown in the presence of 50 μg of L-leucine per ml. (C) The top line shows the positions of putative promoters (Fig. 2) relative to *ilvI*. The amount of this region remaining in the indicated plasmids after deletion is shown underneath. Assays were done in duplicate or quadruplicate. Data are included for three separate experiments indicated by the following symbols: ○ and ●, strains CV709, CV794, CV795, CV798, CV801, CV803, CV804, CV805, and CV806 (strain W3102 containing plasmids pCV53, pCV75, pCV76, pCV79, pCV82, pCV84, pCV85, pCV86, and pCV89, respectively); □ and ■, strains CV787, CV808, CV809, CV810, CV811, CV812, and CV813 [strain W3102(pSE150) containing plasmids pCV53, pCV75, pCV76, pCV82, pCV84, pCV85, and pCV89, respectively]; △ and ▲, strains CV787, CV808, CV809, CV810, CV811, CV812, CV813, CV818, CV819, and CV820 [strain W3102(pSE150) containing plasmids pCV53, pCV75, pCV76, pCV82, pCV84, pCV85, pCV89, pCV77, pCV79, and pCV86, respectively]. Uncorrected galactokinase activity (open symbols) is nanomoles of galactose phosphate per minute per milligram of protein; specific activity was corrected for plasmid copy number (closed symbols) by dividing by the relative plasmid copy number.

region of DNA immediately upstream of *ilvI* of 50 bp or fewer is required for transcription of the *ilvIH* operon. (ii) Sequences more than 195 bp upstream of *ilvI* are necessary for optimal transcription. (iii) Leucine seems to negate the stimulatory effect of upstream sequences.

DISCUSSION

The *ilvIH* promoter and sequences causing leucine repression of *ilvIH* were within 363 bp of the *ilvI* gene. Major 5′ endpoints of *ilvIH* mRNA transcribed in vivo map within this region between positions −25 and −35 (Fig. 2). These endpoints are not an artifact of S1 nuclease cleavage of longer molecules at sites of a nucleic acid secondary structure because S1 mapping experiments with a probe that did not include sequences between −1 and −49 did not identify any major upstream RNA endpoints. Furthermore, it should be noted that endpoints lying upstream of −50 were detected by the S1 mapping technique used here when RNA synthesized in vitro was analyzed.

The observed 5′ endpoints of RNA transcribed in vivo may represent points of transcription initiation. Alternatively, transcription may begin at one or more upstream promoters, in which case the observed 5′ endpoints correspond to the endpoints of processed RNA molecules. The results of our experiments, summarized below, support the first of these two hypotheses. (i) An *Hae*III-*Sau*3A fragment containing 312 bp of upstream DNA but lacking 49 bp immediately upstream of *ilvI* had no promoter activity when cloned into plasmid pKO1 (Fig. 1 and Table 2). Thus, sequences immediately upstream of *ilvI* were required for

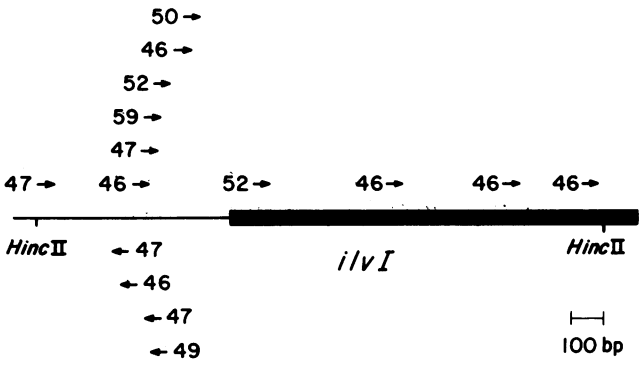

FIG. 7. Potential promoters within fragment *Hin*cII identified by a computer search for homology to a consensus sequence (31). Arrows indicate both the position of the potential promoter and the direction of transcription. Numbers in front of arrows represent the calculated strength (31) of the indicated potential promoter.

transcription. If transcription is initiated further upstream, as proposed in the second hypothesis, then it is difficult to understand why a small downstream deletion should affect promoter activity. (ii) 5′ endpoints corresponding approximately to position −30 were the only endpoints found among pulse-labeled RNA molecules. (iii) 5′ RNA endpoints clustered around position −30 were detected in RNA transcribed in vitro, suggesting that transcription initiation can occur at this point. (iv) When upstream sequences were deleted, including sequences to within 100 bp of *ilvI*, promoter expression was weakened but not eliminated (Fig. 5). These results indicate that there is at least some promoter activity immediately upstream of *ilvI*. The simplest interpretation of these data is that in vivo, transcription of the *ilvIH* operon is initiated near position −30 from a promoter termed hereafter P1. This is assumed to be true for the remainder of this discussion.

The −10 and −35 regions of promoter P1 are identified in Fig. 2 by bars. Mutation *ilv-753* (and an identical mutation, *ilv-752*), which increased expression of *ilvIH* threefold (Table 3), is a G→A transition that brings the −10 region closer to the consensus for *E. coli* promoters (position 40 [Fig. 2]) (19). Mutation *ilv-754*, on the other hand, lies between −10 and −35, in a region not normally considered important for promoter strength. However, a mutation in an analogous position in the *lacP1* promoter (13, 19) results in increased expression of the *lac* operon.

The *ilvIH* promoter activity associated with fragment *Hae*III$_{361}$ is equivalent to that of a moderate-strength promoter, with two- to threefold greater strength than that of the *lacUV5* promoter (Table 2). However, promoter P1 by itself was relatively weak both in vivo (for example, in plasmid pCV82; Fig. 6) and in vitro (Fig. 3). Both of these experimental estimates of promoter strength are consistent with the low calculated estimate based upon comparisons to a consensus sequence (31). The calculated strengths of the *lacUV5* and *ilvIH* P1 promoters are 59 and 40, respectively (a difference of 19 corresponds to an almost 100-fold difference in estimated strength) (31). This analysis suggests that additional elements are needed to activate promoter P1 in vivo.

A number of other bacterial promoters with low homology ratings are known to require activators in vivo, including the *cat* promoter (24), λPRM (27), the *araBAD* promoter (23, 33), *ilvGEDA* promoter P2 (C. W. Adams and G. W. Hatfield, personal communication), *galP1* (2), and *lacP1* (13). For example, promoters λPRM (34) and *lacP1* (13) are activated by the binding of λ cI repressor and catabolite activator proteins, respectively, to upstream sequences. The *araBAD* promoter requires the binding upstream of both the catabolite activator protein and the *araC* gene product for activation (23, 33). These upstream sequences are all within 110 bp of the site of transcription initiation. The striking finding reported here is that sequences more than 150 bp upstream of the presumed site of transcription initiation appear to be necessary for optimal expression of *ilvIH*.

Recently, deletion analyses similar to the one presented here were reported for the *ilvGEDA* promoter region in *E. coli* (Adams and Hatfield, personal communication) and the *nifLA* promoter region of *Klebsiella pneumoniae* (14). For both of these operons, sequences more than 100 bp upstream of the respective transcription initiation sites are needed for optimal transcription. These results together with ours suggest that there may be a class of promoters in enteric bacteria that differs from known classes. The distinguishing feature of this putative new class is that sequences relatively far upstream of the transcription initiation site are required for optimal transcription initiation.

How might upstream sequences play a role in *ilvIH* expression? Conceivably, optimal expression requires a positive regulatory protein that binds to sites more than 150 bp upstream of the site of *ilvIH* transcription initiation. Such a protein bound to upstream sequences might interact directly with RNA polymerase bound at the *ilvIH* promoter, assuming sufficient flexibility of the intervening DNA (15), or activation might require binding of a regulatory protein to multiple sites extending upstream from the *ilvIH* promoter for 150 bp. It may be noted that Ursini et al. (45) have identified transacting mutations that increase *ilvIH* expression.

Another possibility is that RNA polymerase itself is a positive regulatory factor. The sequence of fragment *Hin*cII$_{1733}$ (Fig. 2) (41; C. Squires and J. Calvo, unpublished data) was searched for sequences having homology to an *E. coli* promoter consensus sequence, using the computer homology search of Mulligan et al. (31). A total of 15 potential promoter sequences of calculated strength greater than 45 were found, 10 of them within 300 bp immediately upstream of *ilvI* (Fig. 7; note that *ilvIH* promoter P1, of calculated strength 40, is not included among the 15 sequences). The high frequency of putative promoter sequences within this latter region as compared with surrounding sequences may be in part due to its high A+T content (71% A+T for the region from −123 to −273 [Fig. 2]). The positions of the three putative promoters within *Hae*III$_{361}$ ranking above 50 and oriented towards *ilvIH* are shown in Fig. 2 (labeled P4, P3, and P2). These three putative promoters could give rise to three of the four major species of RNA transcribed in vitro from this region (5′ endpoints at −96, −177, and −210 [Fig. 2 and 3]). Moreover, P4 and P3 lie within the region identified by deletion analysis as being required for optimal transcription of the *ilvIH* operon (Fig. 6). It is interesting to note that the two strongest promoters in vitro, P3 and P2, are arranged with respect to P1 such that three molecules of RNA polymerase could bind to this region simultaneously. Perhaps several molecules of RNA polymerase bound in tandem are needed to activate a downstream promoter. This idea has been discussed by Travers et al. (44), Lamond and Travers (22), and Adams and Hatfield (personal communication).

The mechanism by which leucine represses transcription of *ilvIH* remains unclear. Transcription initiated near posi-

tion −30 was repressed by leucine (Fig. 3). However, transcription from promoter P1 itself (without upstream sequences) was not repressed by leucine (Fig. 6). Thus, leucine repression and promoter P1 activation by upstream sequences may represent different aspects of the same regulatory mechanism.

ACKNOWLEDGMENTS

This work was supported by joint United States-Italy program grant INT8003321 from the National Science Foundation (to J.M.C.) and from the CNR (to M.D.), by grants PCM8003480 and PCM8315842 from the National Science Foundation (to J.M.C.), and by Progetto Finalizzato Ingegneria Genetica e Basi Molecolari delle Malattie Ereditarie of the CNR (to M.D.).

We thank K. McKenney, H. Weiher, G. Walker, B. Buchmann, M. Levinthal, and L. Lau for gifts of strains, plasmids, and enzymes and M. Stucky for typing the manuscript. We also thank E. Demarco and E. Lin for help in constructing strains and assaying enzymes and E. Rosenthal and B. Tyler for a careful reading of the manuscript.

LITERATURE CITED

1. **Adams, C. W., and G. W. Hatfield.** 1984. Effects of promoter strengths and growth conditions on copy number of transcription fusion vectors. J. Biol. Chem. **259**:7399–7403.
2. **Adhya, S., and W. Miller.** 1979. Modulation of the two promoters of the galactose operon of *Escherichia coli*. Nature (London) **279**:492–494.
3. **Bachmann, B. J.** 1972. Pedigrees of some mutant strains of *Escherichia coli* K-12. Bacteriol. Rev. **36**:525–557.
4. **Berk, A. J., and P. A. Sharp.** 1978. Spliced early mRNAs of simian virus 40. Proc. Natl. Acad. Sci. U.S.A. **75**:1274–1278.
5. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
6. **Bolton, E. T., and R. B. Roberts.** 1964. *E. coli* nucleotide pools: phosphorus incorporation, p. 61–70. *In* R. B. Roberts (ed.), Studies of macromolecular biosynthesis. Carnegie Institute of Washington, Washington, D.C.
7. **Calvo, J. M., M. Freundlich, and H. E. Umbarger.** 1969. Regulation of branched-chain amino acid biosynthesis on *Salmonella typhimurium*: isolation of regulatory mutants. J. Bacteriol. **97**:1277–1282.
8. **Cohen, S. N., A. C. Y. Chang, and L. Hsu.** 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. Proc. Natl. Acad. Sci. U.S.A. **69**:2110–2114.
9. **Cooper, T. G., P. Whitney, and B. Magasanik.** 1974. Reaction of *lac*-specific ribonucleic acid from *Escherichia coli* with *lac* deoxyribonucleic acid. J. Biol. Chem. **249**:6548–6555.
10. **DeFelice, M., J. Guardiola, B. Esposito, and M. Iaccarino.** 1974. Structural genes for a newly recognized acetolactate synthase in *Escherichia coli* K-12. J. Bacteriol. **120**:1068–1077.
11. **DeFelice, M., C. T. Lago, C. H. Squires, and J. M. Calvo.** 1982. Acetohydroxyacid synthase isoenzymes of *Escherichia coli* K12 and *Salmonella typhimurium*. Ann. Microbiol. Inst. Pasteur **133A**:251–256.
12. **DeFelice, M., and M. Levinthal.** 1977. The acetohydroxy acid synthase III isoenzyme of *Escherichia coli* K-12: regulation of synthesis by leucine. Biochem. Biophys. Res. Commun. **79**:82–87.
13. **Dickson, R. C., J. Abelson, W. M. Barnes, and W. S. Reznikoff.** 1975. Genetic regulation: the *lac* control region. Science **187**:27–35.
14. **Drummond, M., J. Clements, M. Merrick, and R. Dixon.** 1983. Positive control and autogenous regulation of the *nifLA* promoter in *Klebsiella pneumoniae*. Nature (London) **301**:302–307.
15. **Dunn, T. M., S. Hahn, S. Ogden, and R. F. Schleif.** 1984. An operator at −280 base pairs that is required for repression of *araBAD* operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression. Proc. Natl. Acad. Sci. U.S.A. **81**:5017–5020.
16. **Favaloro, J., R. Treisman, and R. Kamen.** 1980. Transcription maps of polyoma virus-specific RNA: analysis by two-dimensional nuclease S1 gel mapping. Methods Enzymol. **65**:718–749.
17. **Groves, W., F. Davis, and B. Sells.** 1968. Spectrophotometric determination of microgram quantities of protein without nucleic acid interference. Anal. Biochem. **22**:195–210.
18. **Hashimoto, T., and M. Sekiguchi.** 1976. Isolation of temperature-sensitive mutants of R plasmid by *in vitro* mutagenesis with hydroxylamine. J. Bacteriol. **127**:1561–1563.
19. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11**:2237–2255.
20. **LaCara, F., and M. DeFelice.** 1979. Chromatographic detection of the acetohydroxy acid synthase isoenzymes of *Escherichia coli* K-12. Biochem. Biophys. Res. Commun. **91**:319–326.
21. **Lago, C. T., G. Sannia, G. Marino, C. H. Squires, J. M. Calvo, and M. DeFelice.** 1985. The *ilvIH* operon of *Escherichia coli* K-12: identification of the gene products and recognition of the translational start by polypeptide microsequencing. Biochim. Biophys. Acta **824**:74–79.
22. **Lamond, A. I., and A. A. Travers.** 1983. Requirement for an upstream element for optimal transcription of a bacterial tRNA gene. Nature (London) **305**:248–250.
23. **Lee, N. L., W. O. Gielow, and R. G. Wallace.** 1981. Mechanism of *ara*C autoregulation and the domains of two overlapping promoters P_C and P_{BAD} in the L-arabinose regulatory region of *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. **78**:752–756.
24. **LeGrice, S. F. J., and H. Matzura.** 1981. Binding of RNA polymerase and the catabolite gene activator protein within the *cat* promoter in *Escherichia coli*. J. Mol. Biol. **150**:185–196.
25. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular Cloning. A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
27. **McClure, W. R., D. K. Hawley, and T. P. Malan.** 1982. The mechanism of RNA polymerase activation in the λP_{RM} and *lac* P^+ promoters, p. 111–120. *In* R. L. Rodriguez and M. J. Chamberlin (ed.), Promoters: structure and function. Praeger Scientific Publishing Co., New York.
28. **McKenney, K., H. Shimatake, D. Court, U. Schmeissner, and M. Rosenberg.** 1981. A system to study promoter and terminator signals recognized by *E. coli* RNA polymerase, p. 383–415. *In* J. Chirikjian and T. Papas (ed.), Gene amplification and analysis, vol. II. Elsevier-North Holland Publishing Co., New York.
29. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101C**:20–78.
30. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 433. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
31. **Mulligan, M. E., D. K. Hawley, R. Entriken, and W. R. McClure.** 1984. *Escherichia coli* promoter sequences predict *in vitro* RNA polymerase selectivity. Nucleic Acids Res. **12**:789–800.
32. **Neidhardt, F. C., P. L. Bloch, and D. F. Smith.** 1974. Culture medium for enterobacteria. J. Bacteriol. **119**:736–747.
33. **Ogden, S., D. Haggerty, C. M. Storer, D. Kolodrubetz, and R. Schleif.** 1980. The *Escherichia coli* L-arabinose operon: binding sites of the regulatory proteins and a mechanism of positive and negative regulation. Proc. Natl. Acad. Sci. U.S.A. **77**: 3346–3350.
34. **Ptashne, M., A. Jeffrey, A. D. Johnson, R. Maurer, B. J. Moyer, C. O. Pabo, J. M. Roberts, and R. T. Sauer.** 1980. How the λ repressor and *cro* work. Cell **19**:1–11.
35. **Salditt-Georgieff, M., and J. E. Darnell, Jr.** 1983. A precise termination site in the mouse β^{major}-globin transcription unit. Proc. Natl. Acad. Sci. U.S.A. **80**:4694–4698.
36. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. **74**:5463–5467.

37. **Schreier, P. H., and R. Cortese.** 1979. A fast and simple method for sequencing DNA cloned in the single-stranded bacteriophage M13. J. Mol. Biol. **129**:169–172.
38. **Smith, D. R., and J. M. Calvo.** 1979. Regulation of dihydrofolate reductase synthesis in *Escherichia coli*. Mol. Gen. Genet. **175**:31–38.
39. **Smith, D. R., and J. M. Calvo.** 1980. Nucleotide sequence of the *E. coli* gene coding for dihydrofolate reductase. Nucleic Acids Res. **8**:2255–2274.
40. **Sninsky, J. J., B. E. Uhlin, P. Gustafsson, and S. N. Cohen.** 1981. Construction and characterization of a novel two-plasmid system for accomplishing temperature-regulated, amplified expression of cloned adventitious genes in *Escherichia coli*. Gene **16**:275–286.
41. **Squires, C. H., M. DeFelice, J. Devereux, and J. M. Calvo.** 1983. Molecular structure of *ilvIH* and its evolutionary relationship to *ilvG* in *Escherichia coli* K-12. Nucleic Acids Res. **11**:5299–5313.
42. **Squires, C. H., M. DeFelice, S. R. Wessler, and J. M. Calvo.** 1981. Physical characterization of the *ilvHI* operon of *Escherichia coli* K-12. J. Bacteriol. **147**:797–804.
43. **Stueber, D., and H. Bujard.** 1982. Transcription from efficient promoters can interfere with plasmid replication and diminish expression of plasmid specified genes. EMBO J. **1**:1399–1404.
44. **Travers, A. A., A. I. Lamond, H. A. F. Mace, and M. L. Berman.** 1983. RNA polymerase interactions with the upstream region of the *E. coli tyrT* promoter. Cell **35**:265–273.
45. **Ursini, M. U., P. Arcari, and M. DeFelice.** 1981. Acetohydroxy acid synthase isoenzymes of *Escherichia coli* K-12: a *trans*-acting regulatory locus for *ilvHI* gene expression. Mol. Gen. Genet. **181**:491–496.
46. **Weiher, H., and H. Schaller.** 1982. Segment-specific mutagenesis: extensive mutagenesis of a *lac* promoter/operator element. Proc. Natl. Acad. Sci. U.S.A. **79**:1408–1412.