

Activation from a Distance: Roles of Lrp and Integration Host Factor in Transcriptional Activation of *gltBDF*

LIGI PAUL,¹ ROBERT M. BLUMENTHAL,² AND ROWENA G. MATTHEWS^{1,3*}

Biophysics Research Division¹ and Department of Biological Chemistry,³ University of Michigan, Ann Arbor, Michigan 48109, and Department of Microbiology and Immunology, Medical College of Ohio, Toledo, Ohio 43614²

Received 23 January 2001/Accepted 13 April 2001

The leucine-responsive regulatory protein (Lrp) binds to three sites centered 252, 216, and 152 bp upstream of the transcription start site of the *Escherichia coli* glutamate synthase operon (*gltBDF*) and activates transcription. Activators of σ^{70} -dependent promoters usually bind closer to the -35 hexamer of the core promoter sequence. To study the mechanism by which Lrp-dependent activation occurs over this relatively large distance, the *gltBDF* upstream region was sequentially replaced with corresponding portions from the well-characterized σ^{70} -dependent promoter *lacZYAp*. The *glt-lac* promoter hybrids were placed upstream of *lacZ*, allowing transcriptional activity to be monitored via β -galactosidase assays. Even replacing all *gltBDF* sequences downstream of and including the -35 hexamer did not eliminate Lrp-dependent activation of transcription. When a 91-bp region between the -35 hexamer and the proximal Lrp binding site (-48 to -128) was replaced with heterologous DNA of the same length, transcription was reduced nearly 40-fold. Based on the presence of a consensus binding sequence, this region seemed likely to be a binding site for integration host factor (IHF). Experiments to study the effects of a *himD* mutant on expression of a *gltB::lacZ* transcriptional fusion, gel mobility shift analyses, and DNA footprinting assays were used to confirm the direct participation of IHF in *gltBDF* promoter regulation. Based on these results, we suggest that IHF plays a crucial architectural role, bringing the distant Lrp complex in close proximity to the promoter-bound RNA polymerase.

Escherichia coli synthesizes glutamate under ammonia-limiting conditions using the enzymes glutamate synthase, specified by the *gltBDF* operon, and glutamine synthetase, specified by *glnA*. The genes *gltB* and *gltD* specify the large and small subunits of glutamate synthase, respectively (7), and *gltF* specifies a protein of unknown function (8, 20). The *gltBDF* operon is positively regulated by the leucine-responsive regulatory protein (Lrp) (12, 13) and is transcribed at very low levels in the absence of Lrp (12). The coregulator for Lrp is leucine, and exogenous leucine reduces the expression of *gltBDF* about twofold, although this level is still substantially above baseline transcription in the absence of Lrp. Thus, relative to more responsive operons, *gltBDF* is a leucine-independent member of the Lrp regulon (12).

There are three Lrp binding sites centered 246, 215, and 152 bp upstream of the *gltBDF* transcription start site (35); the closest of these sites is ~ 110 bp from the -35 hexamer. The presence of such distal regulatory sites is unusual in σ^{70} -regulated promoters, although it has precedence in promoters that are recognized by σ^{54} (6, 9, 28). In σ^{54} -dependent promoters, upstream elements or activators binding far upstream of the core promoter sequence are brought closer to the RNA polymerase by a looping mechanism that is sometimes assisted by DNA-bending proteins (9), raising the question of whether additional regulatory proteins are required for activation of *gltBDF* transcription. Upon binding the *glt* DNA, the Lrp dimer bends the DNA towards itself; occupancy of all three

sites results in both phased hypersensitivity of the region from -126 to -264 (35) and compaction of the DNA in this region (D. E. Wiese II, M. Young, R. G. Matthews, and C. Bustamante, unpublished data), suggesting formation of a nucleosome-like structure. It is not known whether Lrp activates transcription by a mechanism involving direct contact with RNA polymerase.

In this study, we have examined the role of the core promoter and downstream sequences in Lrp-dependent activation of the *gltBDF* operon. We show that Lrp-dependent activation persists even when the entire region downstream of and including the -35 hexamer is replaced by the corresponding region from the *lac* operon. We also demonstrate that integration host factor (IHF) binds to the region between the proximal Lrp binding site and the -35 hexamer and positively regulates *gltBDF* transcription. We propose that IHF-induced bending of DNA positions bound Lrp close to RNA polymerase for transcriptional activation of the *gltBDF* operon.

MATERIALS AND METHODS

Strains and growth conditions. The strains used in this study are listed in Table 1. Cultures were grown in LB medium (30) or glucose minimal MOPS (morpholinopropanesulfonic acid) medium (25) at 37°C. The antibiotics ampicillin (80 to 100 $\mu\text{g/ml}$), tetracycline (20 $\mu\text{g/ml}$), and chloramphenicol (25 $\mu\text{g/ml}$), were added to the medium as indicated below.

Replacement of the *gltBDF* upstream region with the *lac* operon sequences and construction of λ lysogens. Fusions of the *glt* and *lac* operon regions (Fig. 1) were constructed from fragments of *glt* and *lac* operons obtained by PCR amplification using the primers listed in Table 1. The *glt* regions were amplified from plasmid pBE10 (13), and the *lac* regions were amplified from *E. coli* strain W3110 genomic DNA purchased from Sigma Aldrich (St. Louis, Mo.).

The PCR mixtures contained 2 ng of template plasmid DNA or 200 ng of *E. coli* genomic DNA, 1 U of Vent polymerase (New England Biolabs, Beverly, Mass.), 200 mM concentrations of each primer, and 200 μM concentrations of

* Corresponding author. Mailing address: Biophysics Research Division, 4028 Chemistry, 930 N. University Ave, University of Michigan, Ann Arbor, MI 48109-1055. Phone: (734) 764-5257. Fax: (734) 764-3323. E-mail rmatthew@umich.edu.

TABLE 1. Strains of *E. coli* and primers used in this work

Strain, lysogen, or primer	Description	Reference or source
Strains		
BE1	W3110 <i>lrp-201::Tn10</i>	12
DH5 α	<i>supE44</i> Δ (<i>lacZYA-argF</i>)U169 (ϕ 80 <i>dlacZ</i> Δ M15) <i>deoR hsdR17</i> ($r_K^- m_K^+$) <i>recA1 endA1 gyrA96 thi-1 relA1</i> $\lambda^- F'$	Life Technologies
JWD3	BE1/pJWD2	13
PS2209	W3110 Δ <i>lac-169</i>	F. C. Neidhardt
RJ1413	MC1000 <i>himD::cat</i>	R. Osuna (originally from B. Ely)
W3110	<i>E. coli</i> F ⁻ prototroph	F. C. Neidhardt
XL1-Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> [F' <i>proAB lacI^q Z</i> Δ M15 Tn10 (Tet ^r)]	Stratagene
Recombinant λ lysogens of strain PS2209		
LP1008	<i>glt-lacZ</i> transcriptional fusion with -406 to +8 of <i>glt</i> fused to +1 of the <i>lac</i> operon	This study
LP1023	<i>glt-lacZ</i> transcriptional fusion with -406 to -23 of <i>glt</i> fused to -24 of the <i>lac</i> operon	This study
LP1035	<i>glt-lacZ</i> transcriptional fusion with -406 to -35 of <i>glt</i> fused to -36 of the <i>lac</i> operon	This study
LP1048	LP1007 with the region from -48 to -128 replaced with a fragment of equal length from the <i>cat</i> gene	This study
LP1000	<i>gltB::lacZ</i> transcriptional fusion with -406 to +246 of the <i>gltBDF</i> operon	This study
LP2000	LP1000 with the <i>himD::cat</i> allele from strain RJ1413	This study
Primers^a		
GLT1	<u>ggatccgAATTCGGCGATGACCTGGATCAA</u> (-406 to -384)	
GLT2	<u>ACTGCaGATTTCCAACCTATCGGG</u> (complementary to -15 to +8)	
GLT3	<u>cTgCAGGTCAAATGGCAAGCTTAT</u> (complementary to -47 to -23)	
GLT4	<u>AAGCTTATTGGTACAGAAATGTGC</u> (complementary to -62 to -39)	
GLT5	<u>ggatcccccggATCCCTCTCAAGGGATTTATCGTA</u> (complementary to +217 to +246)	
GLT6	<u>ctgcGAAGTCGTTAGAGAAACAGTC</u> (complementary to -148 to -128)	
GLT7	<u>ctgcagCAATAAGCTTGCCATTTGACC</u> (-54 to -28)	
GLTFP	<u>GTCGTCAGTTCAAGGCAGGATAAAGG</u> (-203 to -179)	
LAC1	<u>CAGCAGGATATCCTGCACCATCGTCTGCTC</u> (complementary to +1175 to +1146)	
LAC2	<u>ctgcagAATTGTGAGCGGATAACAAT</u> (+1 to +20)	
LAC3	<u>ctgcagGCTTCCGGCTCGTATGTTGTGTGGA</u> (-24 to +1)	
LAC4	<u>aaGCTTGCCATTTACACTTTATGCTTCCGG</u> (-40 to -17)	
CAT1	<u>GTCGcgaTATCACTCCAGAGCGATG</u> (complementary to bp 576 to 601 of pKK-232)	
CAT2	<u>ctgCagTAGTGTTACCCCTTGTTACACC</u> (bp 526 to 550 of pKK-232)	

^a The restriction enzyme sites present in the primers are underlined, and sequences different from the template are given in lowercase letters. GLT primers correspond to *gltBDF* operon sequences, LAC primers correspond to the *lac* operon sequences, and CAT primers correspond to the *cat* gene sequences. The locations of the GLT and LAC primers are given with respect to the transcription start sites of the *glt* (26) and *lac* (23) operons, respectively.

each deoxynucleoside triphosphate (Life Technologies, Rockville, Md.) in thermopol reaction buffer [10 mM KCl, 10 mM (NH₄)₂SO₄, 20 mM Tris-HCl (pH 8.8), 2 mM MgSO₄, and 0.1% Triton X-100] in a total volume of 100 μ l. The reaction was carried out for 20 cycles of 96°C for 30 s, 60°C minus 0.5°C/cycle for 1 min, and 72°C for 1.5 min, followed by 10 cycles of 96°C for 30 s, 50°C for 1 min, and 72°C for 1.5 min. The PCR products were purified using a Qiaquick PCR purification kit (Qiagen Inc., Valencia, Calif.). A single deoxyadenosine was added to the 3' end of the products by incubating them at 72°C for 20 min with 5 U of *Taq* DNA polymerase (Life Technologies) and 200 μ M dATP in 20 mM Tris-HCl (pH 8.4) containing 50 mM KCl and 2 mM MgCl₂. The products were resolved on a 1.0% agarose gel, purified using a Qiaquick gel extraction kit (Qiagen), and cloned into *E. coli* strains XL-1 Blue or DH5 α using the pGEM-T vector (Promega, Madison, Wis.). The sequences of the cloned fragments were verified using the ABI PRIZM dye terminator cycle sequencing method (Applied Biosystems, Foster City, Calif.) at the University of Michigan DNA Sequencing Core Facility.

Restriction enzyme sites were incorporated into the primers used for PCR (Table 1), allowing ligation of the *glt* and *lac* operon sequences (Fig. 1). All the *glt* fragments had an *EcoRI* site at the 5' end introduced from the primer *glt1*, and all the *lac* fragments had an *EcoRV* site at the 3' end introduced from the primer *lac1*. The ligated *glt-lac* fragments were introduced into plasmid pRS528 (31) cut with *EcoRI* and *EcoRV*.

The *glt-lac* chimera in strain LP1008 was generated using fragments derived by amplification of *glt* DNA using primers *glt1* and *glt2* (-406 to +8) and amplification of *lac* DNA using primers *lac2* and *lac1* (+1 to +1166); primers *glt2* and *lac2* each contain a *PstI* site, allowing ligation of the fragments together. The *glt* and *lac* fragments are numbered with respect to their transcription start sites (23, 26).

The *glt-lac* chimera in strain LP1023 was obtained using the primer pairs *glt1* and *glt3* (-406 to -23) and *lac3* and *lac1* (-24 to +1166). Primers *glt3* and *lac3* each contain a *PstI* site, which was used to ligate the two fragments together.

The chimera in strain LP1035 was generated using primers *glt1* and *glt4* (-406 to -35) and *lac4* and *lac1* (-36 to +1166). Primer *glt4* has a *HindIII* site that is present in the *glt* sequence, and primer *lac4* has a *HindIII* site upstream of the *lac* sequence.

The chimera in strain LP1048 was constructed by replacing the *glt* sequence between -48 and -128 in strain LP1008 with a fragment of the chloramphenicol acetyltransferase gene (*cat*). The *cat* fragment was amplified from the plasmid vector pKK232-8 (Pharmacia) using the primers *cat1* and *cat2*. The flanking regions were amplified from pLP1008, which contains the *glt-lac* chimera introduced into strain LP1008, with primers *glt1* and *glt6* for the upstream flanking region and *glt7* and *lac1* for the downstream flanking region. Primer *glt6* contains an *NruI* site, and primer *cat1* contains complementary sequence, allowing ligation of the upstream fragment with the *cat* insert. Primer *glt7* contains a *PstI* site, as does primer *cat2*, allowing ligation of the downstream fragment to the *cat* insert.

The *glt* region of strain LP1000 (Table 1) was generated using primers *glt1* and *glt5* (-406 to +246), resulting in a fragment with an *EcoRI* site at the 5' end and a *BamHI* site at the 3' end which was ligated into plasmid pRS415 (31) cut with *EcoRI* and *BamHI*.

The *glt-lac* fusions from the pRS vectors were transferred into the chromosome of *E. coli* strain PS2209 (W3110 Δ *lac*) at the *attB* site using λ RZ5 (2) according to the procedure described in the work of Weise et al. (35). Single-copy lysogens were identified using the PCR method described by Powell et al. (27).

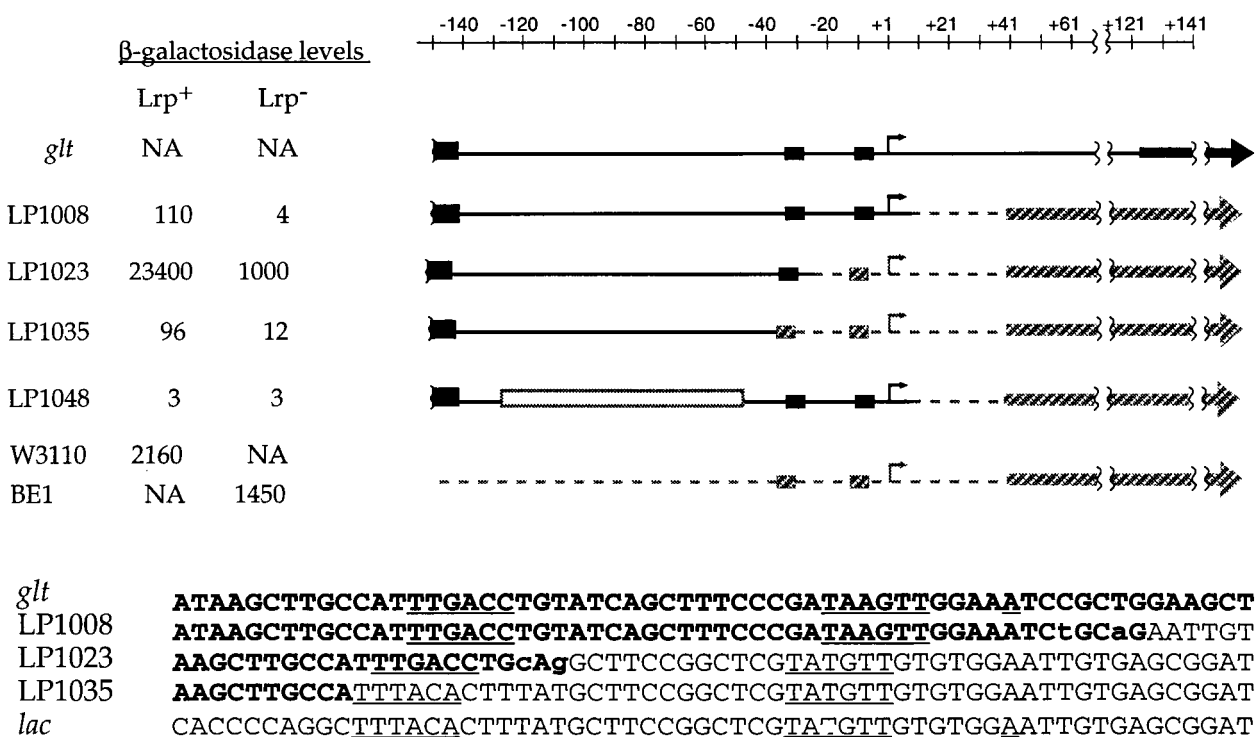


FIG. 1. Maps and β -galactosidase activities of *glt-lac* chimeric operon fusions. The chimeric promoters were introduced as a single copy into the *attB* site of *E. coli* strain PS2209 (W3110 Δ *lac*). The *gltBDF* sequences are shown with a black line, and the *lac* sequences are shown by a gray hatched line. The base pairs that were changed to introduce restriction sites in the chimeras are indicated in lowercase letters in the sequences shown at the bottom. The region from -48 to -128 in the *gltBDF* operon that was replaced with a fragment from the *cat* gene in strain LP1048 is shown as an open rectangular box. The box on the far left of each construct indicates the 3' portion of the proximal Lrp binding site that extends from -142 to -161 . The -35 and -10 hexamers are also represented by boxes and are shown underlined in the chimeric sequences at the bottom; arrows and the underlined A indicate the start sites of transcription determined for the native *gltBDF* (26) or *lacZYA* (23) promoters. The β -galactosidase activity from each chimeric strain in the presence or absence of Lrp is presented on the left side of the construct diagrams. The β -galactosidase activities from the chromosomal *lacZ* gene in *E. coli* strain W3110 (*lrp*⁺) and isogenic strain BE1 (*lrp*::Tn10) are also shown for comparison. NA, not applicable.

Transferring *lrp* and *himD* mutations into λ lysogens. An *lrp-201*::Tn10 allele from the *E. coli* strain BE1 (13) and a *himD*::*cat* allele from *E. coli* strain RJ1413 (from Robert Osuna, University at Albany, Albany, N.Y.) were transferred into the λ lysogens carrying *glt-lacZ* fusions by P1 *vir* transduction (24). The transductants were selected on LB medium-ampicillin plates containing tetracycline or chloramphenicol.

β -Galactosidase assays. The λ lysogens were grown in glucose minimal MOPS medium supplemented with 0.4 mM isoleucine and 0.6 mM valine (13). IPTG (isopropyl- β -D-thiogalactopyranoside; 0.5 mM) was added to the cultures to avoid any effect of the *lac* repressor on the expression of *glt-lac* fusion constructs. Samples for β -galactosidase assays were taken throughout the growth period, and the assays were carried out as described by Ernsting et al. (13). The absorbances of the cultures at 420 nm were plotted versus the β -galactosidase levels. The points were fitted by linear regression, and the slope of the line indicates the β -galactosidase activity of the culture (35). Only samples taken before the cultures reached an A_{420} of 1.0 were used for the slope determinations. The β -galactosidase levels from the wild-type chromosomal *lacZ* in Lrp⁺ (W3110) and Lrp⁻ (BE1) backgrounds were used as controls to account for the effect of Lrp on *lacZ* transcription.

Gel mobility shift assays. The DNA fragments used for gel mobility shift assays were amplified by PCR from pBE10 (13). The DNA was labeled at the 5' end using T4 polynucleotide kinase (Life Technologies) and [γ -³²P]ATP (7,000 Ci/mmol; ICN Biomedical Research Products, Costa Mesa, Calif.) at 37°C for 45 min. The unincorporated nucleotides were removed by passing the labeling reaction mixtures through a Biospin 6 gel filtration column (Bio-Rad Laboratories, Hercules, Calif.). Lrp was purified from an overexpressing *E. coli* strain (JWD3) using the procedure described earlier (13). Purified IHF was a gift from Steven D. Goodman (University of Southern California). The labeled DNA fragment (1.15 nM) was incubated with various concentrations of Lrp and/or IHF in a buffer containing 20 mM Tris-acetate (pH 8.0), 0.1 mM EDTA, 0.1 mM

dithiothreitol, 50 mM NaCl, 4 mM Mg acetate, 12.5% glycerol (vol/vol), and 200 ng of poly(dI-dC) · poly(dI-dC) (Amersham Pharmacia Biotech, Piscataway, N.J.) in a total volume of 20 μ l. The mixtures were left at room temperature for 5 min before being incubated at 28°C for 15 min. The samples were loaded directly onto a 6% polyacrylamide gel (8.3 cm wide by 7.3 cm long by 1.0 mm) in 0.5 \times TBE (45 mM Tris-borate [pH 8.3], 0.1 mM EDTA) and electrophoresed at 12 mA and 4°C. The electrophoresis buffer contained 0.5 \times TBE and 5 mM MgCl₂. The gel was fixed in a solution of 10% acetic acid and 10% methanol for 15 min and dried at 80°C. Biomax MS film (Kodak, Rochester, N.Y.) was used for autoradiography. The gels were scanned using a PhosphorImager (Molecular Dynamics, Sunnyvale, Calif.), and the intensities of the bands were determined using ImageQuant version 1.2 software. The data were fitted to the Michaelis-Menten equation using Kaleidagraph (Synergy Software, Reading, Pa.).

DNase I footprinting. Footprinting assays were carried out as described earlier (5). The *gltBDF* DNA fragment used for footprinting of the template strand was amplified by PCR from plasmid pBE10 (13) using primers *glt1* and *glt2* and digested with the restriction enzyme *HincII*, resulting in a fragment from -176 to $+8$. For footprinting of the nontemplate strand, DNA was amplified with primers *gltFP* and *glt5* and digested with *Bsp1286I*, resulting in a fragment from -203 to $+161$. The DNA fragments were labeled at the *glt2* and *gltFP* ends for the template and nontemplate strands, respectively, using T4 polynucleotide kinase and [γ -³²P]ATP as described above for gel mobility shift assays. Various concentrations of IHF were incubated with the labeled DNA in a buffer containing 10 mM Tris-HCl (pH 8.0), 5 mM MgCl₂, 1 mM CaCl₂, 2 mM dithiothreitol, 50 μ g of bovine serum albumin per ml, and 2 μ g of poly(dI-dC) · poly(dI-dC) per ml in a total volume of 20 μ l. The mixtures were incubated at 28°C for 30 min, after which 0.05 U of RNase-free DNase I (amplification grade from Life Technologies) was added to each tube. DNase I digestion was stopped after 2 min by adding 700 μ l of a stop solution (645 μ l of 100% ethanol, 5 μ l of saturated ammonium acetate, and 5 μ g of yeast tRNA) (5). The precipitated DNA frag-

ments were resolved on a 6% acrylamide gel containing 8 M urea (30). A dideoxy sequencing reaction of the *gltBDF* upstream region with the primer corresponding to the labeled end, carried out using a T7 Sequenase version 2.07-deaza-dGTP kit (Amersham Pharmacia Biotech), was used to generate the ladder.

RESULTS

Effect of replacing the *gltBDF* promoter regions with those of the *lac* promoter on Lrp-dependent activation. To study the role of the DNA sequences downstream of the proximal Lrp binding site centered at -152 in the regulation of the glutamate synthase (*gltBDF*) operon of *E. coli*, these downstream sequences were systematically replaced with the corresponding regions of the *lacZYA* operon. Single copies of the chimeric constructs were introduced into *E. coli* strain PS2209 (W3110 Δlac) at the *att λ* site by lambda integration. Transcription from these hybrid constructs was monitored using *lacZ* as the reporter gene. Figure 1 illustrates the sequence replacements and the β -galactosidase activities of each construct. As described below, replacing the *gltBDF* region downstream of and including the -35 hexamer with the corresponding regions of the *lac* operon did not abolish Lrp-dependent regulation. Thus, Lrp is clearly able to regulate the well-characterized σ^{70} -dependent promoter *lacZYAp* from its binding sites more than 110 bp upstream of the -35 hexamer.

In strain LP1008, the region downstream of $+8$ of the *gltBDF* operon was replaced by the *lac* operon sequence. Bases at $+4$ and $+7$ were changed (C to T and T to A, respectively) to introduce a *PstI* site, which was used for ligating the *glt* and *lac* fragments. In strain LP1008, *lacZ* was still under the control of Lrp and showed a nearly 30-fold reduction in β -galactosidase levels in an Lrp⁻ background relative to the wild-type strain.

Strain LP1023 has a chimeric promoter, with the -35 hexamer being from *gltBDF* and the -10 hexamer being from the *lac* operon. The sequence in the spacer region derived from *glt* was changed in two positions (T to G at -23 and T to C at -25) to generate a *PstI* site for preparing the construct, and the spacer sequence downstream of -23 was from the *lac* operon. The length of the spacer sequence is 17 bp as in *glt*; the wild-type *lac* spacer region is 18 bp. This construct showed an extremely high level of β -galactosidase activity: 200-fold higher at both the basal and activated levels compared to levels in strain LP1008. Nonetheless transcription from this chimeric promoter was activated over 20-fold by Lrp.

In strain LP1035 the -35 hexamer and the entire region downstream of it have been replaced with the corresponding sequences from the *lac* operon. The distance between the proximal Lrp binding site and the -35 hexamer was 1 bp shorter than in the other constructs (111 versus 110 bp). The β -galactosidase level from this construct was reduced sevenfold in an Lrp⁻ background, but the basal level of transcription was higher than in strain LP1008.

In the *gltBDF* operon, the proximal Lrp binding site centered at -152 and the RNA polymerase recognition sequence at -34 are farther apart (more than 110 bp) than is usually seen in other σ^{70} -dependent promoters. To see if the intervening region is involved in transcriptional regulation of the *glt* operon, the region from -48 to -128 was replaced with a portion of the chloramphenicol acetyltransferase (*cat*) gene of the same length (LP1048). This replacement had no effect on

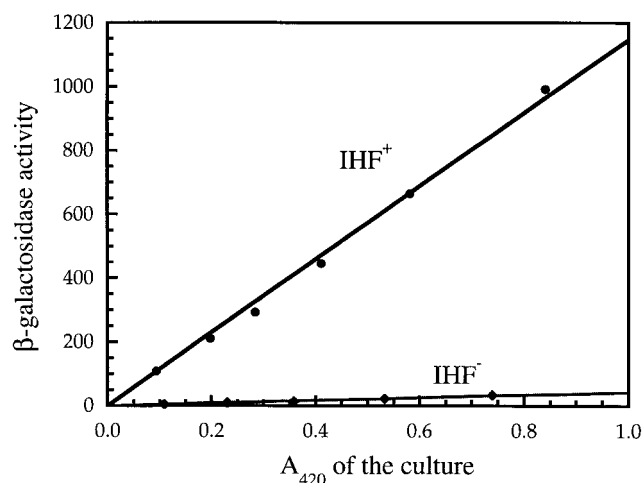


FIG. 2. Effect of IHF on *gltB-lacZ* expression. The β -galactosidase activity of the *gltB-lacZ* transcriptional fusion, integrated as a single copy into the *attB* site of isogenic IHF⁺ and IHF⁻ (*himD::cat*) strains (LP1000 and LP2000, respectively) constructed from *E. coli* strain PS2209 (W3110 Δlac), is shown. The cultures were grown in glucose minimal MOPS medium, and samples were taken at intervals during the growth. The optical density of the cultures was measured at 420 nm. The β -galactosidase activities for the cultures were calculated from the slopes of the lines and were 1,180 for strain LP1000 and 39 for strain LP2000.

basal transcription but completely abolished Lrp-mediated activation.

IHF is required for positive regulation of transcription from the *gltBDF* promoter. In preliminary experiments we used surface-enhanced laser desorption and ionization (SELDI) ProteinChip technology (Ciphergen Biosystems, Fremont, Calif.) to identify proteins that might bind to the upstream region of the *gltBDF* operon. In this approach, biotin-labeled DNA fragments of the *glt* region (from -406 to $+132$) were attached to streptavidin-coated chips and incubated with cell extracts. The masses of proteins that bound to the DNA fragments were determined using mass spectroscopy. This binding experiment was carried out under nonstringent conditions, and there were multiple peaks corresponding to various proteins that bound the DNA nonspecifically. Even though the results of these experiments were not conclusive, they suggested that two polypeptides with molecular masses of 10.6 and 11.2 kDa bound to the *gltBDF* upstream region. The molecular masses of these polypeptides matched those of the two subunits of IHF. These preliminary results led us to focus our study on the possible role of IHF in *gltBDF* transcription.

Upstream of the *gltBDF* promoter, within the region replaced by the *cat* fragment in strain LP10048, there are two regions resembling the consensus IHF binding sequence WATCAANNNTTR, where W is A or T and R is A or G (10, 17). The region from -95 to -83 (TTTCAGTCATTTA) has two mismatches and the overlapping region from -91 to -79 (AGTCATTTAATAA) has three mismatches to the consensus sequence. An insertion in the *himD* gene encoding the β subunit of IHF (*himD::cat*) was transferred by P1 *vir* transduction into strain LP1000, which contains a *gltB::lacZ* transcriptional fusion. The IHF⁻ strain had a β -galactosidase level over 30-fold lower than that of the IHF⁺ strain (Fig. 2).

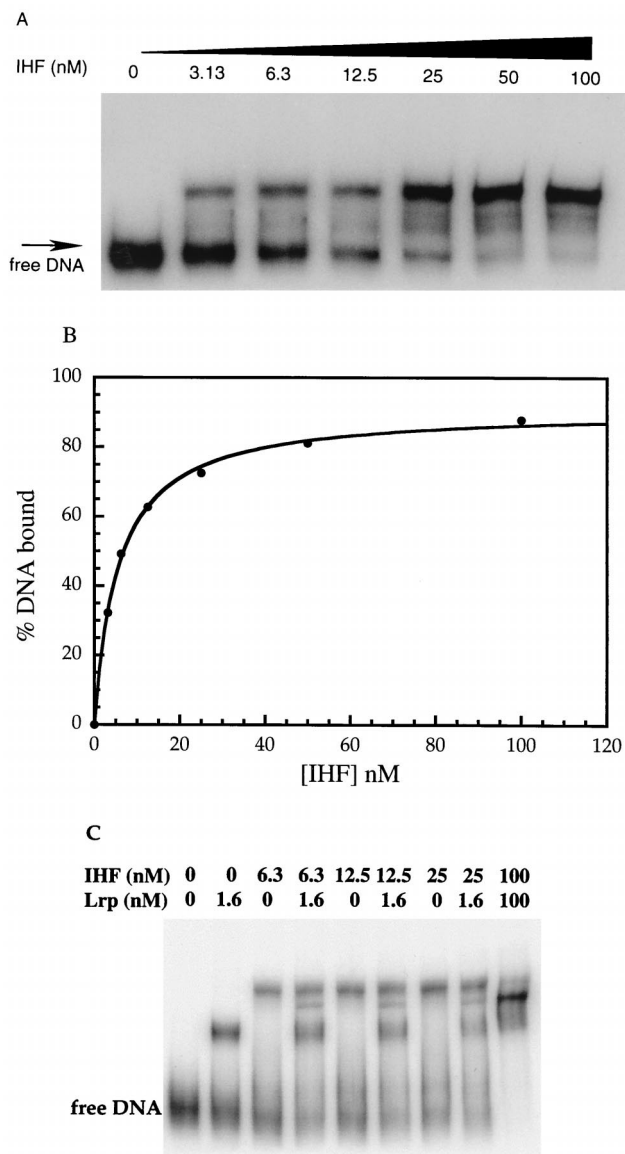


FIG. 3. Gel mobility shift analysis of IHF binding to *glt* DNA. (A) The assays used a DNA fragment containing the region from -406 to $+8$ of the *gltBDF* promoter region and various amounts of purified IHF. The reaction mixtures contained 1.15 nM DNA and 0 to 100 nM IHF. (B) The free DNA in the above gel mobility shift assay was quantitated using a phosphorImager and used to calculate the bound DNA. The data were fitted using the Michaelis-Menten equation. The K_d of IHF for the *gltBDF* upstream region was 5.7 ± 0.3 nM. (C) Lrp and IHF binding to *gltBDF* DNA (-324 to $+246$). The PCR fragment from -406 to $+246$ was digested with the restriction enzyme *Nsi*I to obtain the above-described fragment. The reaction mixtures contained 1.6 to 100 nM Lrp and 6.3 to 100 nM IHF as noted above the lanes. In the presence of both Lrp and IHF in the reaction mixture, we observed an additional band corresponding to the ternary complex, which was not present when either Lrp or IHF alone was used in the reactions; at saturating concentrations of Lrp and IHF, most of the DNA migrated in this band.

Purified IHF binds to the upstream region of the *gltBDF* promoter. To see if IHF affects *gltBDF* transcription directly, gel mobility shift assays were carried out using purified IHF and the region upstream of the *gltB* translational initiation

codon. Assays with both the region from -406 to $+246$ that included the first 30 bp of the coding region of *gltB* and a shorter fragment (-406 to $+8$) showed similar band shift patterns. A gel mobility shift assay with the shorter fragment is shown in Fig. 3A. Migration of the DNA fragment was retarded in the presence of IHF. IHF binds to the *gltBDF* upstream region with an apparent K_d of 5.7 ± 0.3 nM (Fig. 3B). A DNA fragment from -406 to $+246$ but with the region from -48 to -128 replaced with an equivalent length of the *cat* gene sequence did not bind IHF under identical binding conditions (not shown).

When both Lrp and IHF were present in the gel mobility shift assays, we observed an additional band that was not present in assays with either IHF or Lrp alone. The position of the shifted DNA fragment corresponding to the ternary complex depended on the distance of the IHF and Lrp binding sites from the ends of the DNA fragment used for the assay (37). When the *glt* DNA fragment from -324 to $+246$ was used, the band corresponding to the ternary complex was seen between those corresponding to IHF or Lrp alone (Fig. 3C). In the presence of saturating amounts of Lrp and IHF, most of the DNA was bound to both IHF and Lrp.

The binding site of IHF on the *gltBDF* upstream DNA was determined using DNase I footprinting assays of both strands of the DNA (Fig. 4). IHF protected a 41-bp region from -115 to -75 on the nontemplate strand and a 38-bp region from -113 to -76 on the template strand (Fig. 5). The bases -78 on the nontemplate strand and -82 on the template strand were made hypersensitive to DNase I digestion by IHF binding. Upstream of the IHF binding consensus site there are two short tracts of adenine (-106 to -109 and -113 to -115), both of which were protected in the DNase I footprinting assays. High-affinity IHF binding sites are often preceded by a poly (A) tract (21), whose narrow minor groove helps optimize IHF-DNA contacts (29). In Fig. 6, the sequences upstream of the *gltBDF* operon of *E. coli* are compared with those in other genera of *Enterobacteriaceae*: *Salmonella enterica* serovar *Typhimurium* LT2, *Klebsiella pneumoniae*, and *Yersinia pestis*. The approximate distance between the proximal Lrp binding site and the IHF site is conserved among the different genera, but the sequence of the intervening region is not highly conserved except for the poly (A) region. Though the sequence of the IHF binding site is not conserved among the four genera, the cytosine that is present in all known natural IHF binding sites (16) is conserved. An adenine following the cytosine is also generally conserved (17), but this residue is a T rather than an A in *K. pneumoniae*.

DISCUSSION

***glt:lac* hybrid promoters are activated by Lrp.** To study the minimum DNA sequence required for Lrp-dependent regulation, the *gltBDF* operon sequence downstream of and including the -35 hexamer was sequentially replaced with the corresponding regions of the *lacZYA* operon. The effects of these substitutions on transcription were studied by monitoring the activity of the reporter gene *lacZ* (Fig. 1).

The *gltBDF* sequences downstream of -23 , including the -10 hexamer, are not required for Lrp-dependent activation of transcription. Strain LP1023 has the -35 hexamer of *glt*-

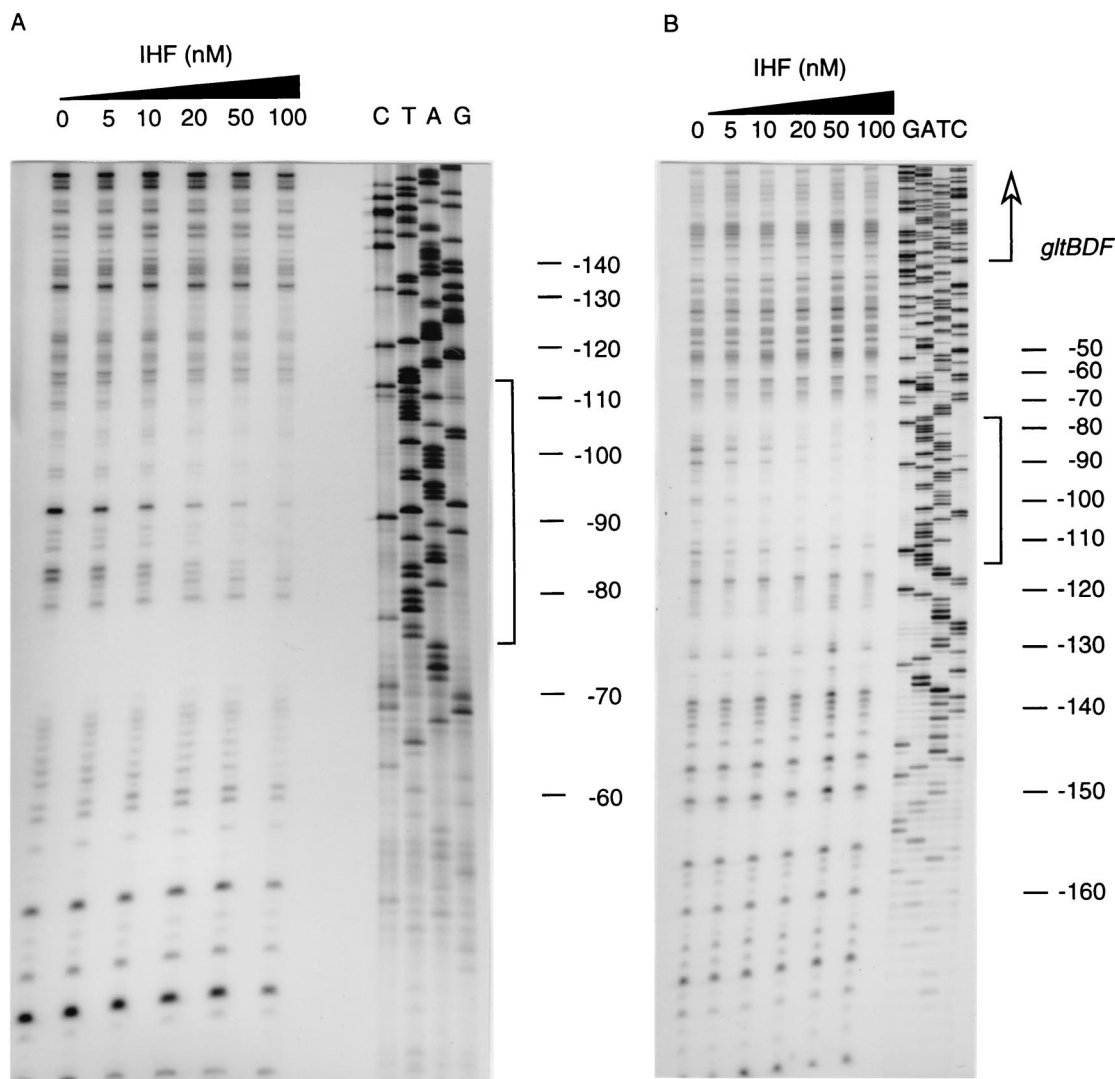


FIG. 4. DNase I footprinting analysis of IHF binding to the *glt* DNA. DNA fragments extending from -176 to +8 of the template strand (the strand complementary to the mRNA sequence) (A) or from -203 to +161 of the nontemplate strand (B) of the *glbBDF* operon were footprinted in the presence of 0 to 100 nM IHF. Dideoxy sequencing reactions of the *glbBDF* promoter region with the primer *glt2* (for the template strand) and *gltFP* (for the nontemplate strand) were used to generate the ladders. IHF protected the region (shown in brackets) from -113 to -76 on the template strand and the region from -115 to -75 on the nontemplate strand against DNase I digestion. The positions of the sequences are given with respect to the transcription start site.

BDF, the -10 hexamer of the *lac* operon, and a hybrid spacer region, yet Lrp-dependent activation (23.4-fold) was similar to that of strain LP1008 (27.5-fold), in which the *glbBDF* core promoter regions were intact. β -Galactosidase expression in

strain LP1023 was >200-fold higher than in strain LP1008 and over 10-fold higher than the IPTG-induced level from the native chromosomal *lacZ* gene in W3110. Apparently, the combination of the -35 and the -10 hexamers or the hybrid

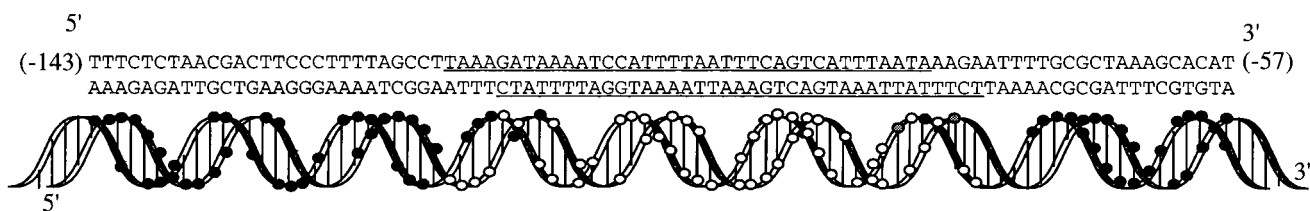


FIG. 5. Diagrammatic representation of the region upstream of the *glbBDF* promoter that was protected against DNase I digestion by IHF binding. Open circles indicate protection against cleavage in the presence of IHF, and the filled circles indicate unchanged cleavage. Where no circles are shown, the DNA was not cleaved by DNase I in the absence of IHF. The two gray circles indicate hypersensitivity to DNase I digestion due to IHF binding. The underlined sequences indicate the region protected from DNase I digestion.



FIG. 6. Comparison of sequences upstream of the *gltBDF* operon from *E. coli* (4) with corresponding sequences from related genera: *S. enterica* serovar *Typhimurium* LT2 (WUGSC 99287/stmlt2-contig1496), *K. pneumoniae* (WUGSC 573/kpneumo B KPN.contig272), and *Y. pestis* (Sanger 632/Y pestis Contig850). The sequences of the related genera were obtained from unfinished genomes using a BLAST search (1), and the alignment was generated using CLUSTAL W (33). Bases conserved in all four genera are shown in bold, and the shaded regions indicate the proximal Lrp binding site, the poly(A) region upstream of the IHF binding site, the consensus IHF binding site (−95 to −83), and the −35 and −10 hexamers. The non-*E. coli* sequences are prepublication communications from the Genome Sequencing Center at Washington University and The Sanger Centre.

spacer region affected the strength of the promoter. A study using random sequences in the spacer region of *Lactococcus lactis* promoters in *L. lactis* and *E. coli* backgrounds has shown that the sequence of the spacer region can affect promoter strength up to 400-fold (22).

Lrp was able to activate transcription even when the *gltBDF* core promoter sequence (the −35 hexamer and downstream sequence) was completely replaced with that of the σ^{70} -dependent *lac* promoter in strain LP1035. The *lac* promoter normally requires catabolite activator protein for activation. Dove et al. (11) have shown that any DNA binding protein that contacts RNA polymerase can activate transcription by strengthening the interaction of RNA polymerase with the promoter region. The extent of Lrp-dependent activation was lower in LP1035 than in strain LP1023 with the chimeric promoter or in strain LP1008 with the *gltBDF* core promoter. It is possible that the longer spacer region between the −35 and −10 hexamers or the shorter region between the proximal Lrp binding site and the −35 hexamer in strain LP1035 affected transcriptional activation. Alternatively, there may be something about the *gltBDF* −35 hexamer that makes it more amenable to Lrp-dependent activation. In the *E. coli fis* promoter, changes in the −35 hexamer, but not in the −10 hexamer, reduce the response to stringent control (34).

The native *gltBDF* promoter is presumed to be recognized by σ^{70} , but evidence for this is not definitive. The promoter is recognized in vitro by purified σ^{70} and core polymerase (B. R. Ernstring, unpublished data), and the sequence and spacing of the −35 and −10 hexamers relative to the transcription start site defined by primer extension are consistent with those of moderately expressed σ^{70} -dependent promoters. The experiments reported here provide definitive evidence that Lrp

dimers, bound more than 110 bp upstream of the −35 hexamer, can activate transcription of a σ^{70} -dependent promoter, since the *lac* operon is controlled by what is probably the best-categorized such promoter.

IHF binding is required for Lrp-dependent activation of the *gltBDF* operon. The mechanism by which the leucine-responsive regulatory protein (Lrp) regulates the genes under its control is not completely understood. In some operons, e.g., *gcv*, Lrp exerts its effect in tandem with other regulatory proteins (32); in others, such as *ilvIH*, Lrp is sufficient for regulation (36). Lrp has been reported to play an architectural role in the transcription of the *gcv* operon (32). Although the *gltBDF* operon of *E. coli* is positively regulated by Lrp (13), in vitro transcription with purified Lrp produced minimal activation of transcription from the *gltBDF* promoter (B. R. Ernstring, unpublished results). This observation suggested that additional factors might be involved in the transcriptional activation of the *gltBDF* operon of *E. coli*.

The 110-bp distance between the proximal Lrp binding site (centered at −152) and the −35 hexamer of the *gltBDF* operon is unusual for a σ^{70} -dependent promoter. When a 91-bp sequence in this region (−128 to −48) was replaced with a fragment of the *cat* gene (strain LP1048), transcription was reduced almost 40-fold. Results of SELDI ProteinChip experiments suggested that IHF might bind to the *gltBDF* upstream sequence. The presence of a consensus IHF binding sequence in this region further supported this idea. In vivo assays using a *himD* mutant (Fig. 2) confirmed that IHF is required for activation of transcription from the *gltBDF* promoter.

While this work was in progress, a global transcript analysis of a *himA* mutant was published (3). Mutation in *himA* led to a 7.1-fold reduction in *gltD* transcript levels in microarrays.

This effect was not as drastic as the approximately 30-fold effect seen in our assays with the *himD* mutant or with a strain where the IHF binding region was replaced by heterologous DNA (strain LP1048). It is possible that functional *himD* homodimers were formed in the *himA* mutant used for the microarray assay, resulting in activation of the *gltBDF* operon transcription to a certain extent (3, 38).

We confirmed that IHF binds to the *gltBDF* upstream region using gel mobility shift and footprinting assays (Fig. 3 to 5). The two overlapping matches to the consensus IHF binding sequence in the *glt* upstream region (−95 to −83 and −91 to −79) are protected in footprinting assays. Site-specific mutational studies have to be conducted to identify the site actually recognized by IHF. The activation of the *gltBDF* operon by Lrp is totally dependent on the presence of IHF, since introduction of an Lrp mutation in strain LP1048 did not lead to a further decrease in β -galactosidase level.

IHF has been shown to be involved in transcriptional regulation of many operons on its own or in concert with additional activator proteins (14, 18). The crystal structure of IHF bound to DNA shows that this protein bends DNA at an $\sim 160^\circ$ angle (29). This facilitates looping of DNA, allowing an upstream element or upstream-bound activator protein to come in close contact with RNA polymerase. Transcriptional activation of the *gltBDF* operon depends completely on the presence of both proteins, suggesting that a DNA looping mechanism may be involved in *gltBDF* regulation.

The proposed IHF-induced bend in the *gltBDF* upstream region would bring the proximal Lrp binding site close to the region just upstream of the −35 hexamer, a position where activators of σ^{70} -dependent promoters usually bind (19). So far, activation of transcription by direct contacts between Lrp and RNA polymerase has not been documented for any promoter. The effect of IHF on transcription from the *gltBDF* promoter is phase dependent: transcription from the *gltBDF* promoter was reduced fivefold when the IHF binding site was taken out of phase with respect to the −35 hexamer and the Lrp binding sites by 5-bp insertions at −40 and −120 (D. E. Weise II, unpublished results).

It is not known at present whether IHF acts as a direct regulator of *gltBDF* transcription or whether it plays a structural role in bringing the transcriptional machinery together. When IHF plays an architectural role, it can be replaced by an intrinsically bent region of DNA (15). Such an experiment remains to be done for the *gltBDF* operon. Since the IHF binding site and the Lrp binding sites upstream of the *glt* promoter are conserved in *S. enterica*, *K. pneumoniae* and *Y. pestis*, the glutamate synthase operons of enteric bacteria appear to be regulated by a common mechanism.

Our results suggest an exception to the general rule that transcriptional activators of σ^{70} -dependent promoters bind close to the −35 hexamer (19). When a confirmed binding site for IHF is present, activators may bind farther upstream and retain function. Since the proximal Lrp binding site in *gltBDF* operon is positioned 110 bp upstream of the −35 hexamer, an intermediate IHF binding site appears to be essential to Lrp-mediated activation. Future experiments will probe for evidence of direct contact between Lrp and/or IHF and RNA polymerase.

ACKNOWLEDGMENTS

We thank Steven D. Goodman (University of Southern California) for providing us with purified IHF, William S. Brusilow (Wayne State University) for λ RZ5, Robert Osuna (University at Albany) for strain RJ1413, and Ruth Van Bogelen (Pfizer, Ann Arbor, Mich.) and Lisa Bradbury (Ciphergen Biosystems) for their assistance with the SELDI ProteinChip experiment. We thank the Genome Sequencing Center, Washington University, St. Louis, Mo., and The Sanger Centre for communication of DNA sequence data prior to publication.

This work was supported by a grant (MCB 9807237) from the National Science Foundation to R.G.M. and R.M.B.

REFERENCES

- Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**:3389–3402.
- Angov, E., and W. S. A. Brusilow. 1988. Use of *lac* fusions to measure *in vivo* regulation of expression of *Escherichia coli* proton-translocating ATPase (*unc*) genes. *J. Bacteriol.* **170**:459–462.
- Arfin, S. M., A. D. Long, E. T. Ito, L. Toller, M. M. Riehle, E. S. Paegle, and G. W. Hatfield. 2000. Global gene expression profiling in *Escherichia coli* K12. The effects of integration host factor. *J. Biol. Chem.* **275**:29672–29684.
- Blattner, F. R., G. I. Plunkett III, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. Davis, W. H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* **277**:1453–1474.
- Brenowitz, M., D. F. Seneor, and R. E. Kingston. 1994. DNA-protein interactions, p. 12.4.1–12.4.11. In V. B. Chanda (ed.), *Current protocols in molecular biology*, vol. 2. John Wiley & Sons, Inc., New York, N.Y.
- Buck, M., S. Miller, M. Drummond, and R. Dixon. 1986. Upstream activator sequences are present in the promoters of nitrogen fixation genes. *Nature* **320**:374–378.
- Castano, I., F. Bastarrachea, and A. A. Covarrubias. 1988. *gltBDF* operon of *Escherichia coli*. *J. Bacteriol.* **170**:821–827.
- Castano, I., N. Flores, F. Valle, A. A. Covarrubias, and F. Bolivar. 1992. *gltF*, a member of the *gltBDF* operon of *Escherichia coli*, is involved in nitrogen-regulated gene expression. *Mol. Microbiol.* **6**:2733–2741.
- Collado-Vides, J., B. Magasanik, and J. D. Gralla. 1991. Control site location and transcriptional regulation in *Escherichia coli*. *Microbiol. Rev.* **55**:371–394.
- Craig, N. L., and H. A. Nash. 1984. *E. coli* integration host factor binds to specific sites in DNA. *Cell* **39**:707–716.
- Dove, S. L., J. K. Joung, and A. Hochschild. 1997. Activation of prokaryotic transcription through arbitrary protein-protein contacts. *Nature* **386**:627–630.
- Ernsting, B. R., M. R. Atkinson, A. J. Ninfa, and R. G. Matthews. 1992. Characterization of the regulon controlled by the leucine-responsive regulatory protein in *Escherichia coli*. *J. Bacteriol.* **174**:1109–1118.
- Ernsting, B. R., J. W. Denninger, R. M. Blumenthal, and R. G. Matthews. 1993. Regulation of the *gltBDF* operon of *Escherichia coli*: how is a leucine-insensitive operon regulated by the leucine-responsive regulatory protein? *J. Bacteriol.* **175**:7160–7169.
- Friedman, D. I. 1988. Integration host factor: a protein for all reasons. *Cell* **55**:545–554.
- Goodman, S. D., S. C. Nicholson, and H. A. Nash. 1992. Deformation of DNA during site-specific recombination of bacteriophage λ : replacement of IHF protein by HU protein or sequence directed bends. *Proc. Natl. Acad. Sci. USA* **89**:11910–11914.
- Goodman, S. D., N. J. Velten, Q. Gao, S. Robinson, and A. M. Segall. 1999. *In vitro* selection of integration host factor binding sites. *J. Bacteriol.* **181**:3246–3255.
- Goodrich, J. A., M. L. Schwartz, and W. R. McClure. 1990. Searching for and predicting the activity of sites for DNA binding proteins: compilation and analysis of the binding sites for *Escherichia coli* integration host factor (IHF). *Nucleic Acids Res.* **18**:4993–5000.
- Goosen, N., and P. van de Putte. 1995. The regulation of transcription initiation by integration host factor. *Mol. Microbiol.* **16**:1–7.
- Gralla, J. D., and J. Collado-Vides. 1996. Organization and function of transcription regulatory elements, p. 1232–1245. In F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umberger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed., vol. 1. ASM Press, Washington, D.C.
- Grassl, G., B. Bufer, B. Muller, M. Rosel, and D. Kleiner. 1999. Characterization of the *gltF* gene product of *Escherichia coli*. *FEMS Microbiol. Lett.* **179**:79–84.
- Hales, L. M., R. I. Gumport, and J. F. Gardner. 1996. Examining the contribution of a dA+dT element to the conformation of *Escherichia coli* integration host factor-DNA complexes. *Nucleic Acids Res.* **24**:1780–1786.

22. **Jensen, P. R., and K. Hammer.** 1998. The sequence of the spacers between the consensus sequences modulates the strength of prokaryotic promoters. *Appl. Environ. Microbiol.* **64**:82–87.
23. **Maizels, N. M.** 1973. The nucleotide sequence of the lactose messenger ribonucleic acid transcribed from the UV5 promoter mutant of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **70**:3585–3589.
24. **Miller, J. H.** 1972. *Experiments in molecular genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Neidhardt, F. C., P. L. Bloch, and D. F. Smith.** 1974. Culture medium for enterobacteria. *J. Bacteriol.* **119**:736–747.
26. **Oliver, G., G. Gosset, R. Sanchez-Pescador, E. Lozoya, L. M. Ku, N. Flores, B. Becerril, F. Valle, and F. Bolivar.** 1987. Determination of nucleotide sequence of glutamate synthase structural genes of *Escherichia coli* K-12. *Gene* **60**:1–11.
27. **Powell, B. S., D. L. Court, Y. Nakamura, M. P. Rivas, and C. L. Turnbough, Jr.** 1994. Rapid confirmation of single copy lambda prophage integration by PCR. *Nucleic Acids Res.* **22**:5765–5766.
28. **Reitzer, L. J., and B. Magasanik.** 1986. Transcription of *glnA* in *E. coli* is stimulated by activator bound to sites far from the promoter. *Cell* **45**:785–792.
29. **Rice, P. A., S. Yang, K. Mizuuchi, and H. A. Nash.** 1996. Crystal structure of an IHF-DNA complex: a protein induced DNA U-turn. *Cell* **87**:1295–1306.
30. **Sambrook, J., E. F. Fritsch, and T. Maniatis (ed.).** 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
31. **Simons, R. W., F. Houtman, and N. Kleckner.** 1987. Improved single and multicopy *lac* based cloning vectors for protein and operon fusions. *Gene* **53**:85–96.
32. **Stauffer, L. T., and G. V. Stauffer.** 1999. Role for the leucine-responsive regulatory protein (Lrp) as a structural protein in regulating the *Escherichia coli gcvTHP* operon. *Microbiology* **145**:569–576.
33. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673–4680.
34. **Walker, K. A., C. L. Atkins, and R. Osuna.** 1999. Functional determinants of the *Escherichia coli fis* promoter: roles of –35, –10, and transcription initiation regions in the response to stringent control and growth phase-dependent regulation. *J. Bacteriol.* **181**:1269–1280.
35. **Weise, D. E., II, B. R. Ernstring, R. M. Blumenthal, and R. G. Matthews.** 1997. A nucleoprotein activation complex between the leucine-responsive regulatory protein and DNA upstream of the *gtBDF* operon in *Escherichia coli*. *J. Mol. Biol.* **270**:152–168.
36. **Willins, D. A., and J. M. Calvo.** 1992. In vitro transcription from the *Escherichia coli ilvIH* promoter. *J. Bacteriol.* **174**:7648–7655.
37. **Wu, H.-M., and D. M. Crothers.** 1984. The locus of sequence-directed and protein-induced DNA bending. *Nature* **308**:509–513.
38. **Zulianello, L., E. de la Gorgue de Rosny, P. van Ulsen, P. van de Putte, and N. Goosen.** 1994. The HimA and HimD subunits of integration host factor can specifically bind to DNA as homodimers. *EMBO J.* **13**:1534–1540.