JOURNAL OF BACTERIOLOGY, Sept. 1993, p. 5769–5777
0021-9193/93/185769-09$02.00/0

Vol. 175, No. 18

Integration Host Factor Is Required for Anaerobic Pyruvate Induction of *pfl* Operon Expression in *Escherichia coli*

AGNIESZKA SIRKO,[1,2] EVA ZEHELEIN,[3] MARTIN FREUNDLICH,[1] AND GARY SAWERS[3]*

Department of Biochemistry, State University of New York at Stony Brook, Stony Brook, New York 11794[1]; Institute of Biochemistry and Biophysics, Polish Academy of Sciences, 02532 Warsaw, Rakowiecka 36, Poland[2]; and Lehrstuhl für Mikrobiologie der Universität München, Maria-Ward-Strasse 1a, 80638 Munich, Germany[3]

Received 12 May 1993/Accepted 14 July 1993

The expression of the pyruvate formate-lyase gene (*pfl*) is induced by anaerobic growth, and this is increased further by growth in pyruvate. Previous work has shown that anaerobic induction is strongly dependent on the activator FNR and partially dependent on a second transcription factor, ArcA, while pyruvate induction only required FNR. Anaerobic and pyruvate regulation both require the presence of a 5′ nontranslated regulatory sequence which spans approximately 500 bp of DNA. A mobility shift assay was developed to identify proteins that bind to this regulatory region. Several binding activities were separated by heparin agarose chromatography, and one of these activities was characterized and shown to be integration host factor (IHF). Mobility shift and DNase I footprinting experiments defined a single IHF binding site in the *pfl* promoter-regulatory region. With *pfl-lacZ* fusions, it could be shown that introduction of a *himD* mutation abolished pyruvate-dependent induction of anaerobic expression in vivo. The same result was observed when the *pfl* IHF binding site was mutated. In addition, the partial anaerobic induction of expression found in an *fnr* strain was completely blocked in an *fnr himD* double mutant and in an *fnr* IHF binding site double mutant. Taken together, these data suggest that IHF is necessary for both pyruvate induction and the anaerobic induction mediated by ArcA.

Pyruvate formate-lyase (PFL) is the key enzyme of carbon catabolism in the anaerobic *Escherichia coli* cell. It is a free radical enzyme that catalyzes the nonoxidative cleavage of pyruvate to formate and acetyl-coenzyme A by a unique mechanism (19, 40). PFL is active only anaerobically (19), and it is the counterpart to the pyruvate dehydrogenase complex of aerobic metabolism.

The *pfl* gene forms an operon with an upstream gene with an unknown function that may encode a membrane protein (34). The operon is transcribed from multiple promoters, and expression is influenced by several factors including anaerobiosis and pyruvate (33). A 5′ nontranslated regulatory sequence encompassing approximately 500 bp of DNA (Fig. 1) has been shown to be essential for mediating the coordinate expression of all seven promoters in response to these regulatory phenomena (34). However, the mechanisms by which this regulatory sequence controls transcription of the downstream promoters (P1 to P5 in Fig. 1) are still unclear.

Shifting cultures from aerobic to anaerobic conditions results in a 12- to 15-fold induction in *pfl* expression (33, 34). Two transcriptional regulators, FNR and ArcA, control this anaerobic induction of gene expression (35). Both proteins function as activators, and strains in which the gene for either one of these proteins is deleted still can induce *pfl* expression anaerobically, albeit to a reduced level; however, anaerobic expression is abolished in an *arcA fnr* double null mutant (35).

The FNR protein probably binds to at least two related DNA sequences in the regulatory region of the *pfl* operon, both of which are proximal to promoters (Fig. 1A). It is currently assumed that ArcA also is a sequence-specific DNA-binding protein (17) and that it interacts with DNA in the *pfl* regulatory region, but its recognition sequence remains to be elucidated.

Expression of the *pfl* operon is also regulated by pyruvate, which stimulates transcription at least an additional twofold relative to the levels attained upon anaerobic induction (33). This increase in expression can be achieved either by supplementing the growth medium with pyruvate or by introduction of a *pfl* mutation, which causes an intracellular accumulation of pyruvate (33). Whether it is pyruvate per se or a metabolite of pyruvate which is responsible for the induction is unclear, as is the induction mechanism. In this regard, *pfl* is similar to a number of anaerobic operons whose maximal expression requires activation by substrate in addition to the major activation by general anaerobic activators (38).

In order to understand the mechanism of transcriptional regulation of the *pfl* operon at the biochemical level, all the factors involved in controlling this system must be identified and characterized. This report describes a method for identifying proteins which bind to the *pfl* operon regulatory region. Using this method, we have purified and identified one of these proteins as being integration host factor (IHF). IHF is a sequence-specific DNA-binding and -bending protein encoded by the *himA* and *himD* (*hip*) genes of *E. coli* (29). IHF participates in a number of processes involving DNA including site-specific recombination, DNA replication, and positive or negative effects on gene expression (10, 12). In many systems, IHF acts as an accessory factor to influence the activity of other DNA-binding proteins. For example, in some operons involved in nitrogen metabolism IHF is thought to enhance the effectiveness of protein-protein interactions between a transcriptional activator bound to a remote promoter-distal site and either RNA polymerase or a second activator bound at a proximal site (4, 16, 32, 36).

* Corresponding author.

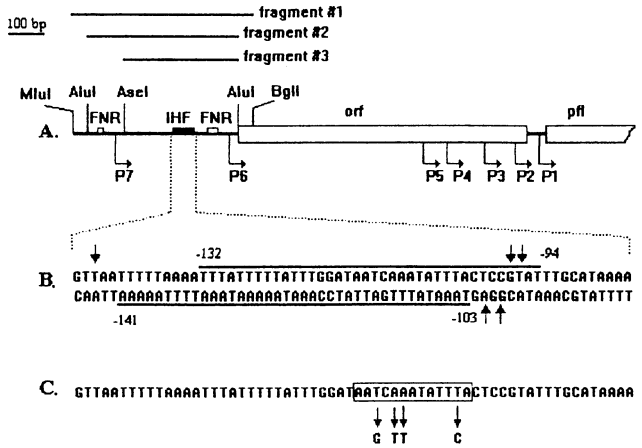

FIG. 1. Organization of the *pfl* operon regulatory region. (A) The *orf* gene and a portion of the *pfl* gene are shown. The arrows signify the 5′ ends of the mapped transcripts, derived from promoters P1 to P7 (34). The small open boxes indicate the positions of the putative FNR binding sites, and the small filled box indicates the binding site for IHF. The *Mlu*I-*Bgl*I and *Alu*I fragments were used for the gel mobility shift, and the *Alu*I-*Ase*I fragment was used for the DNase I protection experiments. (B) The DNA sequences protected by IHF from DNase I are shown as bars above and below the sequence for the noncoding and coding strands, respectively. The arrows signify hypersensitive sites. Numbering is with respect to the major transcription initiation site from promoter 6, which is taken as +1 (34). (C) The nucleotide sequence of the noncoding strand around the IHF binding site in the *pfl* regulatory region. The nucleotides containing the 13-bp consensus sequence (23) are overlined. The nucleotides changed by mutagenesis are shown below the sequence. This change reduces the predicted affinity of IHF from 60.2 to 34.5 as expressed with the similarity score of Goodrich et al. (15).

In the present report, we show that IHF binds to a single site in the *pfl* promoter-regulatory region. In vivo experiments with a *himD* mutant or with a strain containing an altered *pfl* IHF binding site show that IHF is required for maximal *pfl* expression. IHF is absolutely necessary for pyruvate-mediated induction of the *pfl* operon. In addition, the residual ArcA-dependent anaerobic induction of *pfl* found in *fnr* mutants is abolished in strains that lack IHF or the *pfl* IHF binding site.

MATERIALS AND METHODS

Bacterial strains, vectors, and growth conditions. All bacterial strains, plasmids, and phages used in this study are listed in Table 1. Strains were grown in buffered, rich medium (TYEP, pH 6.5) as described by Begg et al. (1). Glucose (20 mM) was used as the carbon source in all growth experiments except when cells were grown aerobically in TYEP medium, in which case carbon source supplementation was omitted. Pyruvate was added to 70 mM when required. Antibiotics were used at the following final concentrations: ampicillin, 100 μg/ml; chloramphenicol, 15 μg/ml; kanamycin, 50 μg/ml; tetracycline, 15 μg/ml. Media and buffers for lambda work were prepared as described previously (34). Amino acid analogs (Sigma Chemical Co.) were added to a final concentration of 20 μg/ml.

To obtain sufficient cell material for the purification of DNA-binding proteins from cell extracts, strain MC4100 was grown anaerobically in rich medium supplemented with 15 mM formate in a 50-liter fermentor (Biostat U50D; Braun, Melsungen, Germany) and harvested in the late exponential phase of growth.

Plasmid and strain constructions. The 447-bp *Alu*I restriction fragment and the 531-bp *Mlu*I-*Bgl*I restriction fragment were isolated from the *pfl* regulatory region (Fig. 1A), and *Bam*HI linkers were ligated after blunt ends were created with the Klenow fragment of DNA polymerase. These fragments were cloned into the *Bam*HI site of pUC19 to create pSN15 (447-bp *Alu*I fragment) and pSN20 (531-bp *Mlu*I-*Bgl*I fragment). M1323 was constructed by ligating the 1,790-bp *Bam*HI fragment from pRM23 into the *Bam*HI site of M13mp19. Genetic crosses were performed by P1 *kc*-mediated transduction as described previously (28). Strain RM312 was constructed by transducing the Δ3(*himD*)::Cam^r mutation from strain MC252 into MC4100 and selecting for chloramphenicol-resistant clones. The IHF$^-$ phenotype was verified by the inability of strains to maintain plasmid pSC101 (14). The *himD* mutation was introduced into strains containing *lacZ* fusions subsequent to infection with transducing phages. RM319 was constructed by first transducing MC4100 to leucine auxotrophy with BW7261 (*leu-63*::Tn*10*) as the donor. Subsequent transduction with a lysate derived from DC488 (*aroP-aceEF*) yielded prototrophic, tetracycline-sensitive clones. These were screened for resistance to the amino acid analogs DL-5-methyltryptophan and β-2-thienyl-DL-alanine, which are not taken up by cells lacking *aroP* (22).

Site-directed mutagenesis. The four base pair changes in the IHF binding site (Fig. 1C) were introduced by two different methods. Inverse polymerase chain reaction (31) was performed with pSN15 with 20 pmol each of primers 5′-GGAGTGATATAAGCTTATC-3′ and 5′-GTATTTGCA TAAAAACCATG-3′. Alternatively, the same mutations were introduced by the method of Kunkel et al. (20), in which single-stranded DNA from M1323 isolated from strain CJ236 (*dut ung*) was hybridized with an oligonucleotide with the sequence 5′-GGATAAGCTTATATTCACTCCG-3′. After second-strand synthesis and ligation, the DNA was used to transform JM109. Plasmid DNA was isolated from putative clones and tested for the presence of the *Hin*dIII restriction site created by the mutagenesis. The authenticity of the positive clones was determined by DNA sequencing.

β-Galactosidase assays. β-Galactosidase activity was assayed in cultures of exponentially growing cells, and the specific activity was calculated as described by Miller (28). Values reported are the average of at least three independent experiments performed in triplicate. The standard error of the values reported was not more than 15%.

Gel mobility shift assay of *pfl* regulatory region-binding activities. The *Bam*HI fragment isolated from pSN20 was labeled with [α-^{32}P]dATP with the Klenow fragment of DNA polymerase. Approximately 1 ng (30,000 cpm) of labeled fragment was restricted to completion with restriction enzyme *Dde*I, *Dra*I, or *Sau*96I (0.1 U) in a final volume of 10 μl. The binding reaction was performed in a final volume of 20 μl in the following buffer: 10 mM Tris-HCl (pH 7.5)–50 mM KCl–5 mM $MgCl_2$–0.5 mM EDTA–4 mM spermidine–1 mM dithiothreitol (DTT)–10% glycerol. A total of 1 μg of sonicated salmon sperm DNA, and either 5 to 50 μg of protein when S-100 fractions were tested or 1 to 10 μl when column fractions were tested was added to the reaction mixture as a competitor to remove nonspecific DNA binding activities. The reaction mixture was equilibrated at room temperature for 5 min, after which time 3 μl of labeled, restricted DNA fragments was added and the incubation was continued for a further 5 min. The reaction was stopped by

TABLE 1. Bacterial strains, plasmids, and phages used in this study

Strain, plasmid, or phage	Genotype	Reference or source
Strains		
MC4100	F⁻ *araD139* Δ(*argF-lac*)*U169 ptsF25 deoC1 relA1 flbB530 rpsL150* λ⁻	3
FM420	MC4100 *recA56*	45
RM101	MC4100 Δ*fnr*	35
RM102	MC4100 Δ*fnr* Δ(*srl-recA*)*306*::Tn*10*	2
RM312	MC4100 Δ*3*(*himD*)::Cam^r	This work
RM313	MC4100 *arcA zjj*::Tn*10*	35
RM202	FM420 Δ*pfl25* Ω(*pfl*::Cam^r)	33
RM319	MC4100 Δ(*aroP-aceEF*)	This work
MC252	*metB ara* Δ(*lac-pro*) *gyrA argEam thi supF* Δ*3*(*himD*)::Cam^r	M. Chandler
BW7261	Hfr Cavalli (KL226, PO2A) Δ(*argF-lac*)*205 U169 leu62*::Tn*10 relA1 tonA22 pit-10 spoT1 ompF627*	42
CJ236	*dut-1 ung-1 thi-1 relA1*/pCJ105	20
DC488	F⁻ *mel supE supF fadR* Δ(*aroP-aceEF*)*15*	D. Clark
JM109	*recA1 endA1 gyrA96 thi hsdR17 supE44 relA1* F(*lac-proAB*)	44
RM123	FM420 λRM123	33
RM135	RM102 λRM123	33
RM204	RM202 λRM123	33
RM320	RM313 λRM123	35
RM324	RM312 Δ*fnr* λRM123	This work
RM325	RM319 λRM123	This work
RM444	RM312 λRM123	This work
RM381	FM420 λRM12321	This work
RM382	RM202 λRM12321	This work
RM383	RM312 λRM12321	This work
RM384	RM102 λRM12321	This work
RM385	RM313 λRM12321	This work
Plasmids		
pRS552	Km^r Ap^r *lacZ' lacY*$^+$ *lacA*$^+$	37
pRM23	pRS552 Φ(*pfl'-'lacZ*)*1397*(Hyb)	33
pRM2321	pRM23 but mutated in the IHF binding site	This work
pSN15	Ap^r *Bam*HI −1353 to −906 of *pfl* in pUC19	This work
pSN15A1	pSN15 but mutated in the IHF binding site	This work
pSN20	Ap^r *Bam*HI −1392 to −861 of *pfl* in pUC19	This work
pUC19	Ap^r *lacIPOZ*	44
Phages		
M13mp19	*lacIPOZ*	44
M1323	*Bam*HI −1353 to +395 of *pfl* in M13mp19	This work
M132321	M1323 but mutated in the IHF binding site	This work
λRS45	*lacZ' lacY*$^+$ *imm-21 ind*$^+$	37
λRM123	λRS45 Φ(*pfl'-'lacZ*)*1397*(Hyb)	33
λRM12321	λRM123 but mutated in the IHF binding site	This work

placing the tubes on ice. A 5-μl aliquot of 10% glycerol containing bromophenol blue (0.001%, wt/vol) was added to the mixture, and 10 μl was loaded onto a 5% nondenaturing polyacrylamide gel prepared in TBE buffer (25) which had been preelectrophoresed for 15 min at 60 V prior to sample application. The presence of the restriction enzymes in the reaction mixture had no influence on the mobility shift assay.

Partial purification of DNA binding activities. All steps were performed at 4°C. A total of 25 g (wet weight) of *E. coli* MC4100 cells was resuspended in a final volume of 80 ml in potassium phosphate buffer, pH 7.0, containing 2 mM DTT and 5 mM benzamidine (Sigma). After disruption by sonification, the extract was clarified by centrifugation at 15,000 × *g* for 20 min. The pellet was discarded, and the centrifugation step was repeated. The crude extract was then centrifuged at 100,000 × *g* for 2 h. The S-100 supernatant (60 ml) was brought to 30% saturation in ammonium sulfate by slowly adding 100% saturated ammonium sulfate, stirred slowly for 30 min, and centrifuged at 15,000 × *g* for 30 min. The pellet was discarded, the supernatant was adjusted to 70% saturation in ammonium sulfate and stirred for a further 30 min at 4°C, and the precipitate was collected by centrifugation at 15,000 × *g* for 30 min. The pellet was resuspended to a final volume of 15 ml in 20 mM Tris-HCl–50 mM KCl–2 mM DTT, pH 7.5 (buffer A), and desalted by passage over a G-25 Sephadex column equilibrated in buffer A.

The desalted material was adjusted to a protein concentration of approximately 30 mg/ml (final volume, 50 ml) and was applied directly to a column (2 by 16 cm) of heparin-Affi-Gel agarose (Bio-Rad) equilibrated in buffer A. Approximately 80% of the protein did not bind to the matrix, and this included only a minor portion of the *pfl* regulatory region binding activity. The column was developed with a linear gradient of 50 to 1,000 mM KCl (500 ml) in buffer A. Fractions (7 ml) were collected, and 5 to 10 μl of each was tested for DNA binding activity as described above.

Active fractions were collected in three pools (pool I, 0.3 to 0.45 M KCl; pool II, 0.5 to 0.6 M KCl; pool III, 0.7 to 0.8 M KCl) which were concentrated (approximately sixfold) to a final volume of 10 ml by pressure dialysis (Amicon cell using a PM-10 membrane), and these were stored at −20°C.

IHF purification. IHF protein was purified by the method

of Filutowicz and Appelt (7) except that bacteria were disrupted by sonication. The pp_L *hip himA-5* plasmid (30) containing the *himA* and *himD* genes under the control of the p_L promoter of lambda was kindly supplied by H. Nash.

Gel mobility shift assay with purified IHF. IHF binding was performed in a total volume of 25 μl in the following buffer: 40 mM Tris-HCl (pH 8.0)–70 mM KCl–10 mM $MgCl_2$–1 mM DTT–0.1 mM ATP–100 μg of bovine serum albumin per ml–6.6 μg of poly(dI-dC) per ml (Boehringer Mannheim)–10% glycerol. The 447-bp *Alu*I-*Alu*I fragment containing the *pfl* regulatory region (Fig. 1A) was isolated from pSN15 or pSN15A1 by digestion with *Sph*I and *Eco*RI (both enzymes cut only in the polylinker of pUC19) and 3′ labeled with the Klenow fragment of DNA polymerase I. Approximately 0.25 ng of DNA and IHF (0 to 40 nM) were used in the reactions. An IHF-specific 48-bp synthetic DNA fragment (8 ng) was used as a competitor. This fragment was designed to give the maximum score for IHF binding as calculated by Goodrich et al. (15). After binding (20 min, room temperature), the samples were loaded on the gel without loading buffer and separated on a 4% polyacrylamide gel (29:1) in LIS buffer (7 mM Tris-HCl [pH 8.0], 33 mM NaOAc, 1 mM EDTA).

DNase I protection. Binding of IHF to the 322-bp *Ase*I-*Bam*HI DNA fragment (Fig. 1A) isolated from pSN15 was performed in the following buffer: 50 mM Tris-HCl (pH 6.5)–70 mM KCl–1 mM EDTA–7 mM $MgCl_2$–2 mM $CaCl_2$–5% glycerol–1 mM β-mercaptoethanol. The Klenow fragment of DNA polymerase I and either [α-^{32}P]dATP for labeling at the *Ase*I site or [α-^{32}P]dGTP for labeling at the *Bam*HI site were used to radiolabel only one 3′ end of the DNA fragment. Approximately 2.5 ng of DNA and various amounts of IHF were used in a total reaction volume of 25 μl. After incubation for 20 min at room temperature, 2 μl of DNase I (0.2 μg/ml; GIBCO-BRL) was added. After 4 min at room temperature, the reaction was stopped by adding 3.5 μl of 0.25 M EDTA (pH 8.0). The products were analyzed by electrophoresis on an 8% polyacrylamide gel containing 8 M urea.

Other methods. Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis was carried out in 15% (wt/vol) polyacrylamide gels generally following the procedure of Laemmli (21). Protein concentration was determined by the method of Lowry et al. (24) with bovine serum albumin as the standard. N-terminal amino acid sequencing was performed with an Applied Biosystems gas-phase sequencer.

RESULTS

Identification of proteins which bind to the *pfl* promoter-regulatory region. To identify proteins which bind to the regulatory region of the *pfl* operon, we developed a gel mobility shift assay (11) in which the *pfl* promoter-regulatory region was used as the substrate. *Bam*HI linkers were added to the *Mlu*I-*Bgl*I fragment from the *pfl* promoter-regulatory region (Fig. 1), which both facilitated cloning of the fragment and ensured identical labeling of both ends of the fragment for the mobility shift assay. Digestion of the labeled fragment at one of the unique internal restriction sites (Fig. 2A) prior to addition to the assay enabled the localization of any DNA binding activity as well as delivering an internal control for the specificity of binding.

A mobility shift using a *Dde*I-restricted fragment demonstrated that several distinct binding activities could be identified in an S-100 extract of MC4100 (data not shown). Fractionated ammonium sulfate precipitation followed by chromatography on heparin agarose (see Materials and Methods for details) identified a specific binding activity that eluted from the heparin agarose column between 0.7 and 0.8 M KCl (pool III), bound to the larger of the *Dde*I fragments, and is probably composed of several distinct activities (Fig. 2B). The location in the promoter-regulatory region to which the activity in pool III binds could be delimited by digesting the labeled DNA fragment with different restriction enzymes prior to its addition to the mobility shift assay. The autoradiogram in Fig. 2B shows that the binding activity in pool III could be localized to the 90-bp region of DNA bounded by the *Dra*I and *Sau*96I restriction sites (Fig. 2A). Treatment of an aliquot of pool III with proteinase K before addition to the mobility shift assay clearly showed that the activity was due to a protein and not a nucleic acid (Fig. 2B).

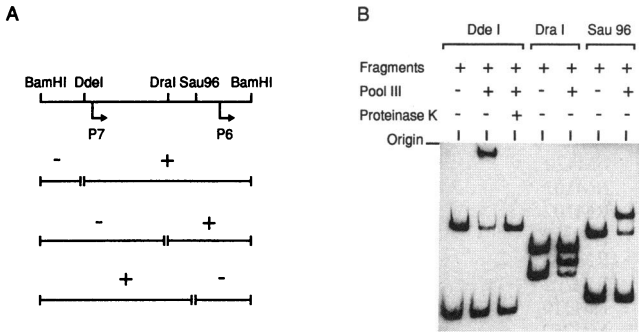

FIG. 2. Localization of the binding activity in pool III. Panel A shows the *Bam*HI fragment from pSN20 which includes the *pfl* operon regulatory region and promoters 6 and 7 (P6 and P7). The arrows indicate the direction and sites of transcription initiation. The unique restriction sites for *Dde*I, *Dra*I, and *Sau*96I are shown. A plus sign indicates that the protein in pool III binds the restriction fragment, and a minus sign indicates that it does not bind the fragment. Panel B shows an autoradiogram of a mobility shift experiment in which the regulatory region was restricted with either *Dde*I, *Dra*I, or *Sau*96I prior to incubation with 20 ng of pool III. Proteinase K treatment involved incubating the aliquot of pool III with 10 μg of proteinase K at 37°C for 10 min before addition to the mobility shift assay.

IHF is the major protein in pool III. SDS-polyacrylamide gel electrophoretic analysis of an aliquot from pool III showed that two polypeptides (constituting more than 90% of the protein in pool III) with molecular masses of approximately 8 and 10 kDa could be identified (data not shown). The elution profile of the polypeptides from the heparin agarose column and their sizes on SDS-polyacrylamide gels are characteristic of IHF (7, 29). Determination of the N-terminal amino acid sequences of both polypeptides gave the sequences ALTKAE for the larger polypeptide and MTKSE for the smaller polypeptide. These sequences correlate exactly with the N-terminal sequences of the α and β subunits of IHF, respectively (8, 27); the initial methionine residue is absent from the α subunit. The total amount of protein in pool III was approximately 600 μg. Assuming that IHF is at least 90% pure, on the basis of polyacrylamide gel electrophoresis, this would indicate that IHF constitutes 0.02% of the soluble protein in the anaerobic *E. coli* cell.

IHF binds to the *pfl* promoter-regulatory region. To confirm that it is indeed IHF which binds to the regulatory region and not one of the minor contaminating proteins in pool III, a homogeneous preparation of IHF purified from an IHF-overproducing strain (see Materials and Methods) was tested

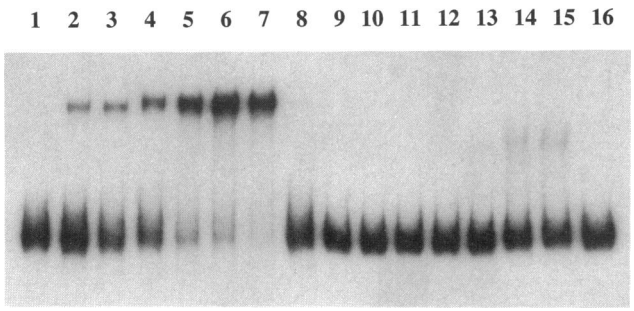

FIG. 3. Specific binding of IHF to the *pfl* promoter-regulatory region. Shown is a mobility shift assay of labeled DNA fragments isolated from pSN15 (wild-type IHF binding site), lanes 1 to 8, or from pSN15A1 (mutagenized IHF binding site), lanes 9 to 16. The concentrations of IHF used were as follows: 0 nM, lanes 1 and 9; 1 nM, lanes 2 and 10; 2.5 nM, lanes 3 and 11; 5 nM, lanes 4 and 12; 10 nM, lanes 5 and 13; 20 nM, lanes 6 and 14; 40 nM, lanes 7, 8, 15, and 16. An IHF-specific competitor DNA (see Materials and Methods) was added to the reaction mixtures in lanes 8 and 16.

in a mobility shift experiment. The results in Fig. 3 show that a single retarded DNA-protein complex was progressively formed as the concentration of IHF was increased from 0.5 to 40 nM. Approximately 5 to 10 nM IHF was required for 50% binding to the fragment. DNase I footprinting (13) was used to determine precisely where IHF binds in this region. The data in Fig. 4 show the protection by IHF of the bottom (Fig. 4A) and top (Fig. 4B) strands of the DNA fragment containing the *pfl* upstream regulatory region. A single protected area can be seen from approximately −132 to −94 bp on the top strand and from approximately −141 to −103 bp on the bottom strand. This site is located between the *Dra*I and *Sau*96I restriction enzyme recognition sequences identified as the region bound by IHF purified from the wild-type extract (Fig. 2) and contains a sequence (Fig. 1B and C) which is identical to the 13-bp IHF consensus sequence (23). The extended IHF consensus sequence of Goodrich et al. (15) was used to obtain a similarity score of 60.2 for the *pfl* IHF site. This score is slightly above the average of the 27 IHF binding sites analyzed previously (15).

IHF is required for maximal anaerobic expression of *pfl*. The in vitro binding of IHF to the *pfl* promoter region suggested that IHF might be involved in the expression of *pfl* in vivo. We examined this possibility by transducing a *himD* mutation into strain RM123 which has an operon fusion between the *pfl* promoter region and the *lacZ* structural gene at the λ attachment site on the *E. coli* chromosome (33). β-Galactosidase activity was measured in the IHF mutant and in the parental strain grown under a variety of conditions that are known to alter *pfl* expression. The IHF mutation had little effect on the large increase in *pfl* expression due to anaerobiosis (Fig. 5). However, the 2.3-fold induction of *pfl* due to addition of pyruvate to cells grown anaerobically in rich media was completely absent in the IHF mutant (Fig. 5). As noted previously (33), pyruvate induction is also completely dependent on FNR, the major activator of anaerobic *pfl* expression; however, pyruvate induction was only marginally affected (1.5-fold induction compared with 2.3-fold for the wild type) by an *arcA* mutation (Fig. 5).

In addition to the role of IHF in pyruvate induction, the results in Fig. 5 also show that IHF is required for complete

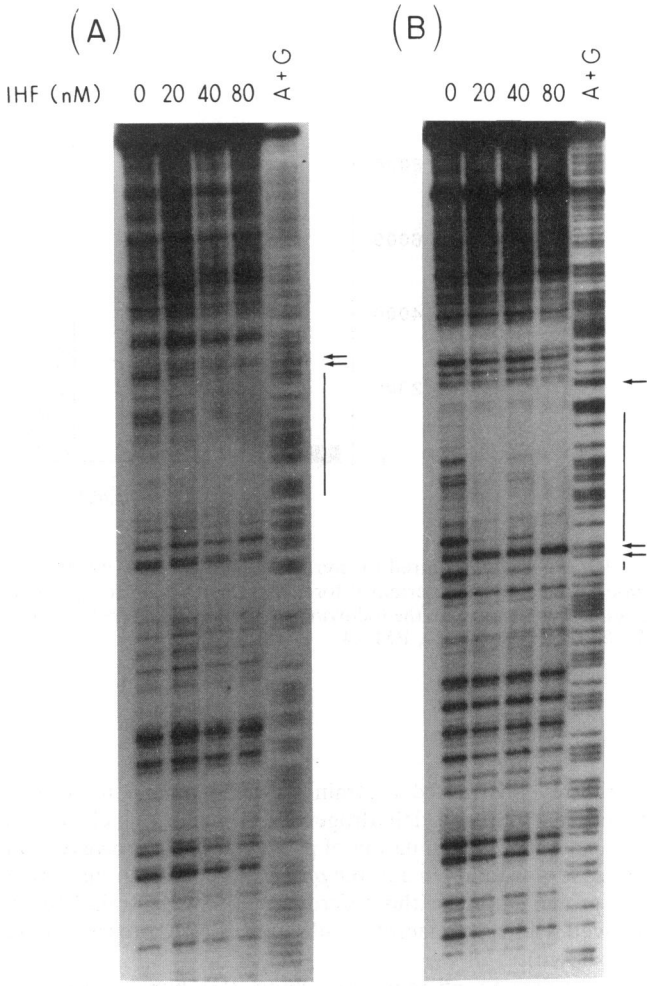

FIG. 4. DNase I protection pattern of the IHF binding site in the *pfl* regulatory region. The *Ase*I-*Bam*HI fragment from pSN15 was isolated and labeled at its 3′ end with the Klenow fragment of DNA polymerase I at the *Ase*I site (bottom strand) (A) or at the *Bam*HI site (top strand) (B). Before DNase I digestion, incubation was carried out with different concentrations of IHF as indicated at the top of each lane. The protected regions are indicated by the vertical bars, and the regions of enhanced sensitivity to DNase I after IHF binding are indicated by the small horizontal arrows. A+G indicates the sequence ladders, which were generated by chemical-cleavage DNA sequencing (26).

anaerobic regulation of *pfl*. Anaerobic induction is reduced approximately 75% in strains lacking FNR (33), and the remaining anaerobic activity has been shown to be dependent on ArcA (33). An *fnr himD* double mutant completely lacked anaerobic induction of *pfl* (Fig. 5). This phenotype is the same as that of an *arcA fnr* double mutant (35) and suggests that the anaerobic regulation of *pfl* by ArcA may be dependent on IHF.

IHF is not required for aerobic *pfl* expression. As was noted previously (33), in contrast to anaerobic growth, added pyruvate did not increase aerobic *pfl* expression, nor does a *himD* mutation affect *pfl* expression under these conditions (data not shown). It is possible that pyruvate is metabolized very rapidly under aerobic conditions and that a sufficient amount of pyruvate is unavailable for *pfl* induction. We

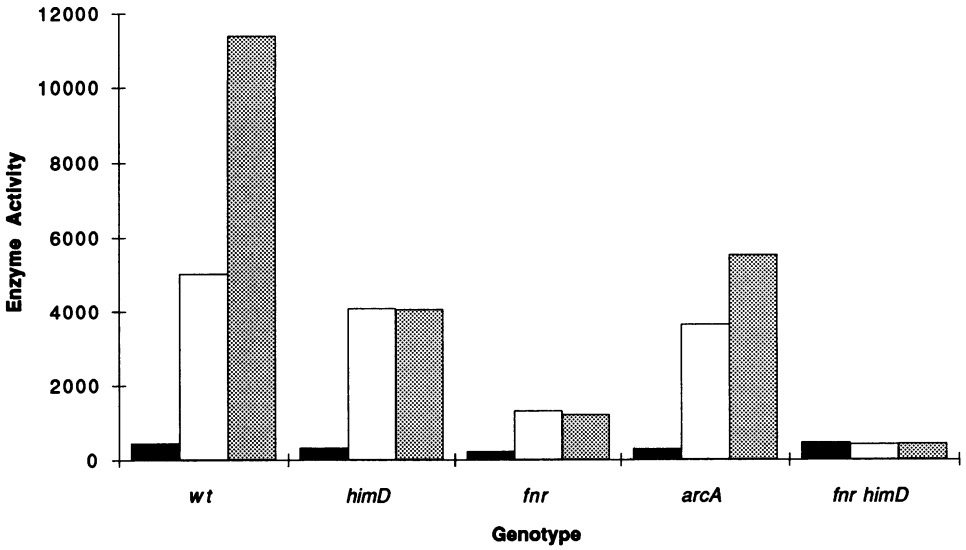

FIG. 5. IHF is required for pyruvate induction and ArcA-dependent anaerobic regulation. β-Galactosidase specific activity (28) from a *pfl-lacZ* fusion was determined for cells grown aerobically (solid bars), anaerobically (open bars), or anaerobically with pyruvate (stippled bars) in rich media with the following strains: *wt* (wild type), parent, RM123; *himD*, IHF$^-$, RM444; *fnr*, FNR$^-$, RM135; *arcA*, ArcA$^-$, RM320; *fnr himD*, FNR$^-$ IHF$^-$, RM324.

therefore constructed a strain (RM325) unable to form a functional pyruvate dehydrogenase complex, which results in the aerobic accumulation of pyruvate (5, 9). However, no increase was found in aerobic *pfl* expression when this strain was grown with or without pyruvate, although strain RM325 showed wild-type anaerobic induction of *pfl* expression (data not shown).

IHF effects on *pfl* require an intact IHF binding site. The results obtained with *pfl-lacZ* fusion strains show that IHF is necessary for maximal anaerobic expression of *pfl*. However, these experiments do not differentiate between direct and indirect effects of IHF on this operon. This was examined by altering the IHF binding site in the *pfl* promoter region by site-directed mutagenesis. A 4-bp change was made in the 13-bp IHF consensus sequence (Fig. 1). This change strongly reduced IHF binding in vitro (Fig. 3). The effect of these mutations on in vivo *pfl* expression was investigated by introducing them into the IHF binding site in the *pfl* promoter region of the *pfl-lacZ* fusion in strain RM123. Anaerobic expression of *pfl* in the IHF binding site mutant was altered in the same way and to the same extent as in the strain that lacked IHF (Fig. 6). That is, pyruvate was no longer able to induce *pfl* and the ArcA-dependent anaerobic induction of the operon was completely blocked (Fig. 6). The IHF binding site mutation also prevented the very high *pfl* expression found in strains lacking PFL (Fig. 6). This has been attributed to the intracellular accumulation of pyruvate due to its inability to be catabolically metabolized in these strains. The results shown in Fig. 6 are consistent with the idea that these high levels are due to pyruvate induction and that IHF is required for this process. These results strongly suggest that IHF is a direct effector of the *pfl* operon and that the effects of IHF on *pfl* expression found in vivo are a direct consequence of the binding of IHF to the site characterized in vitro.

DISCUSSION

PFL is a key enzyme for the growth of *E. coli* in the absence of oxygen (19). Expression of *pfl* under anaerobic conditions is dependent on seven σ^{70}-transcribed promoters that are located approximately 60 to 1,300 bp upstream of the start of *pfl* translation (34). Anaerobic induction of each of these promoters requires FNR and ArcA (35), and anaerobic expression is further increased by growth in media containing pyruvate (33). All of the seven promoters are coordinately regulated by these factors through sequences at the 5′ end of the *pfl* promoter-regulatory region (35). In the present study, we developed an in vitro mobility shift assay to identify potential regulatory proteins which bind to the *pfl* promoter-regulatory region. This assay, in combination with heparin agarose chromatography, enabled us to purify and characterize one of these binding activities as being IHF, and we show that IHF is an additional component required for maximal in vivo *pfl* expression under anaerobic conditions.

In vitro mobility shift and DNase I footprinting experiments identified a single IHF binding site in the *pfl* upstream regulatory region. This site is centered at approximately −115 in relation to the start of transcription of promoter 6 and is located between two putative FNR binding sites (34). Mutagenesis of the 13-bp consensus sequence (23) in the IHF binding site very strongly reduced IHF binding in vitro. This mutation was introduced into the chromosome to analyze IHF effects on anaerobic *pfl* expression. IHF appears to be absolutely necessary for pyruvate induction since no increase in *pfl* expression was found when the IHF binding site mutant or a *himD* mutant was grown with pyruvate or in conditions that increase pyruvate endogenously. The same phenotype was also observed in an IHF$^+$ strain that lacked FNR, suggesting that IHF and FNR may work in consort to mediate pyruvate induction. IHF also has a positive role in the anaerobic induction of *pfl*, although this

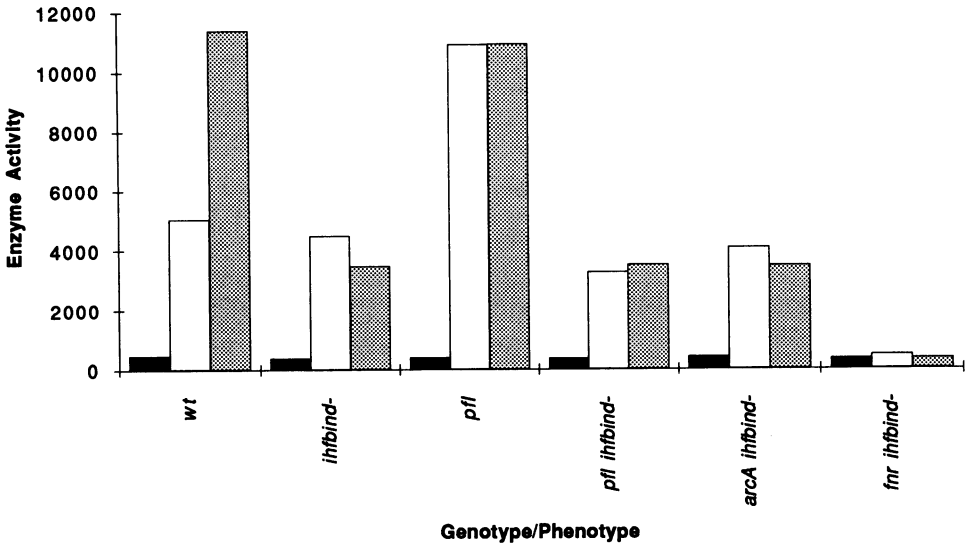

FIG. 6. An intact IHF binding site is required for maximal anaerobic *pfl* expression. β-Galactosidase specific activity (28) from a *pfl-lacZ* fusion was determined for cells grown aerobically (solid bars), anaerobically (open bars), or anaerobically with pyruvate (stippled bars) in rich medium with the following strains: *wt* (wild type), parent, RM123; *ihfbind*$^-$, IHF binding site mutant, RM381; *pfl*, PFL$^-$, RM204; *pfl ihfbind*$^-$, PFL$^-$, IHF binding site mutant, RM382; *arcA ihfbind*$^-$, ArcA$^-$, IHF binding site mutant, RM385; *fnr ihfbind*$^-$, FNR$^-$, IHF binding site mutant, RM384.

was not as obvious as was its function in pyruvate induction. Anaerobic induction was reduced only by approximately 20% in either a *himD* or an IHF binding site mutant. However, either mutation completely eliminated the residual (sixfold) anaerobic induction found in strains unable to produce FNR. This residual activity has been reported to be mediated by ArcA (35), and the *arcA*, *himD*, and IHF binding site mutants have very similar phenotypes in relation to anaerobic induction. Taken together, these results suggest that IHF may alter anaerobic induction of *pfl* primarily by influencing how ArcA functions in this process.

In a number of cases in which IHF positively regulates gene expression, IHF does not activate transcription directly but enhances the effectiveness of an activator protein. The best-characterized examples of this are at the *glnHp2* promoter where activation by NtrC requires IHF (4) and IHF-dependent NifA activation of the *nifH* promoter (16). In these cases, IHF has been postulated to facilitate looping of the DNA, optimizing interaction of the upstream DNA-bound activator with the RNA polymerase. Both of these promoters are σ^{54} dependent, where a second, promoter-proximal activator is not required. IHF also participates in the expression of some σ^{70}-transcribed promoters which are regulated by two activator proteins, where one activator binds at a site proximal to the promoter and the other has a recognition sequence which is more remote (32, 43). A particularly relevant example is the control of the *narGHJI* operon, whose expression is regulated by anaerobiosis and by nitrate (32). The FNR protein induces expression anaerobically, and evidence indicates that it probably binds to a specific recognition sequence located around position −40 bp relative to the transcription initiation site (41). Nitrate regulation is mediated by a second, distinct protein called NarL, which has been proposed to be a sequence-specific DNA-binding protein whose recognition sequence is located some 200 bp upstream of the transcription start site of the *narGHJI* operon (6). The NarL-mediated nitrate regulation of the promoter requires IHF, which recognizes and binds a DNA sequence located between the two putative activator binding sites (32, 36). In an IHF$^-$ strain, anaerobic, FNR-dependent activation of expression still occurred but NarL-dependent nitrate regulation was abolished (32, 36). The regulation of the *nar* operon has some aspects analogous to that reported here for the *pfl* operon, and it is tempting to speculate that pyruvate induction, like nitrate induction of the *nar* operon, is also mediated by a specific DNA-binding protein and that IHF functions as a coactivator facilitating interaction between the activator and either the FNR protein or RNA polymerase or both. Precisely how IHF works in *pfl* and other anaerobic operons awaits the development of in vitro systems that will allow binding and transcription analysis with all of the factors that participate in the expression of each operon.

It is interesting that of the 10 operons in *E. coli* whose expression is directly altered by IHF, 4 are involved in anaerobic metabolism (10, 32, 36, 39, 43). All four, as has been found for many anaerobic systems (38), are regulated by global and specific control elements, and IHF is a strong positive effector for the specific factor in each operon. Perhaps IHF has a special function in anaerobic gene expression. As far as we are aware, this is the first report of IHF having been purified from anaerobically grown *E. coli* cells. Because we did not use a specific assay of IHF in the crude extract, it is difficult to estimate the yield attained during the purification; however, on the basis of the procedure used here, IHF constitutes minimally 0.02% of the soluble protein. This value is 20-fold higher than that found when IHF was purified from aerobic cells (29). These findings alone warrant a more thorough investigation of IHF levels in aerobic and anaerobic cells of *E. coli*.

ACKNOWLEDGMENTS

We are greatly indebted to F. Lottspeich for determining the N-terminal amino acid sequences of the α and β subunits of IHF. We thank M. Chandler, D. Clark, and H. Nash for strains and

plasmids; M. Filutowicz for his protocol for IHF purification; and M. Sauter for performing the large-scale fermentation.

This work was supported by grant GM17152 from the National Institutes of Health to M.F. and a grant from the Deutsche Forschungsgemeinschaft (SFB-145) to G.S.

REFERENCES

1. **Begg, Y. A., J. N. Whyte, and B. A. Haddock.** 1977. The identification of mutants of *Escherichia coli* deficient in formate dehydrogenase and nitrate reductase activities using dye indicator plates. FEMS Microbiol. Lett. **2**:47–50.
2. **Birkmann, A., G. Sawers, and A. Böck.** 1987. Involvement of the *ntrA* gene product in the anaerobic metabolism of *Escherichia coli*. Mol. Gen. Genet. **210**:535–542.
3. **Casadaban, M. J., and S. N. Cohen.** 1979. Lactose genes fused to exogenous promoters in one step using Mu-*lac* bacteriophage: *in vivo* probe for transcriptional control sequences. Proc. Natl. Acad. Sci. USA **76**:4530–4533.
4. **Claviere-Martin, F., and B. Magasanik.** 1991. Role of integration host factor in the regulation of the *glnHp2* promoter of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **88**:1631–1635.
5. **Dietrich, J., and U. Henning.** 1970. Regulation of pyruvate dehydrogenase complex synthesis in *Escherichia coli* K 12: identification of the inducing metabolite. Eur. J. Biochem. **14**:258–269.
6. **Dong, X.-R., S. F. Li, and J. A. DeMoss.** 1992. Upstream sequence elements required for NarL-mediated activation of transcription from the *narGHJI* promoter of *Escherichia coli*. J. Biol. Chem. **267**:14122–14128.
7. **Filutowicz, M., and K. Appelt.** 1988. The integration host factor of *Escherichia coli* binds to multiple sites at plasmid R6K origin and is essential for replication. Nucleic Acids Res. **16**:3829–3834.
8. **Flamm, E. L., and R. A. Weisberg.** 1985. Primary structure of the *hip* gene of *Escherichia coli* and its product, the β-subunit of integration host factor. J. Mol. Biol. **183**:117–128.
9. **Flatgaard, J. E., B. Hoehn, and U. Henning.** 1971. Mutants of *Escherichia coli* K 12 which synthesise the pyruvate dehydrogenase complex constitutively. Arch. Biochem. Biophys. **143**: 461–470.
10. **Freundlich, M., N. Ramani, E. Mathew, A. Sirko, and P. Tsui.** 1992. The role of integration host factor in gene expression in *Escherichia coli*. Mol. Microbiol. **6**:2557–2563.
11. **Fried, M., and D. Crothers.** 1981. Equilibria and kinetics of *lac* repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. **9**:6505–6526.
12. **Friedman, D. I.** 1988. Integration host factor: a protein for all reasons. Cell **55**:545–554.
13. **Galas, M., and A. Schmitz.** 1978. DNase footprinting: a simple method for the detection of protein-DNA binding specificity. Nucleic Acids Res. **5**:3157–3170.
14. **Gamas, P., A. C. Burger, G. Churchward, L. Caro, D. Galas, and M. Chandler.** 1986. Replication of pSC101: effects of mutations in the *Escherichia coli* DNA binding protein IHF. Mol. Gen. Genet. **204**:85–89.
15. **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. Nucleic Acids Res. **18**:4993–5000.
16. **Hoover, T. R., E. Santero, S. Porter, and S. Kustu.** 1990. The integration host factor stimulates interaction of RNA polymerase with NIFA, the transcriptional activator of nitrogen fixation genes. Cell **63**:11–22.
17. **Iuchi, S., and E. C. C. Lin.** 1988. *arcA* (*dye*), a global regulatory gene in *Escherichia coli* mediating repression of enzymes in aerobic pathways. Proc. Natl. Acad. Sci. USA **85**:1888–1892.
18. **Knappe, J., F. A. Neugebauer, H. P. Blaschkowski, and M. Gänzler.** 1984. Post-translational activation introduces a free radical into pyruvate formate-lyase. Proc. Natl. Acad. Sci. USA **81**:1332–1335.
19. **Knappe, J., and G. Sawers.** 1990. A radical-chemical route to acetyl-CoA: the anaerobically induced pyruvate formate-lyase system of *Escherichia coli*. FEMS Microbiol. Rev. **75**:383–398.
20. **Kunkel, T. A., J. D. Roberts, and R. A. Zabour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154**:367–382.
21. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
22. **Langley, D., and J. R. Guest.** 1977. Biochemical genetics of the keto acid dehydrogenase complexes of *Escherichia coli* K-12: isolation and biochemical properties of deletion mutants. J. Gen. Microbiol. **99**:263–276.
23. **Leong, J. M., S. Nunes-Duby, C. F. Lesser, P. Youderian, M. M. Susskind, and A. Landy.** 1985. The 80 and P22 attachment sites: primary structure and interaction with *Escherichia coli* integration host factor. J. Biol. Chem. **260**:4468–4477.
24. **Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall.** 1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem. **193**:265–275.
25. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
26. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
27. **Miller, H. I.** 1984. Primary structure of the *himA* gene of *Escherichia coli*: homology with DNA-binding protein HU and association with the phenylalanyl-tRNA synthetase operon. Cold Spring Harbor Symp. Quant. Biol. **49**:691–698.
28. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. **Nash, H. A., and C. A. Robertson.** 1981. Purification and properties of the *Escherichia coli* protein factor required for lambda integrative recombination. J. Biol. Chem. **256**:9246–9253.
30. **Nash, H. A., C. A. Robertson, E. Flamm, R. A. Weisberg, and H. I. Miller.** 1987. Overproduction of *Escherichia coli* integration host factor, a protein with nonidentical subunits. J. Bacteriol. **169**:4124–4127.
31. **Ochman, H., A. S. Gerber, and D. Hartl.** 1988. Genetic applications of an inverse polymerase chain reaction. Genetics **120**:621–623.
32. **Rabin, R. S., L. A. Collins, and V. Stewart.** 1992. *In vivo* requirement of integration host factor for *nar* (nitrate reductase) operon expression in *Escherichia coli* K-12. Proc. Natl. Acad. Sci. USA **89**:8701–8705.
33. **Sawers, G., and A. Böck.** 1988. Anaerobic induction of pyruvate formate-lyase gene expression from *Escherichia coli* K-12. J. Bacteriol. **170**:5330–5336.
34. **Sawers, G., and A. Böck.** 1989. Novel transcriptional control of the pyruvate formate-lyase gene: upstream regulatory sequences and multiple promoters regulate anaerobic expression. J. Bacteriol. **171**:2485–2498.
35. **Sawers, G., and B. Suppmann.** 1992. Anaerobic induction of pyruvate formate-lyase gene expression is mediated by the ArcA and FNR proteins. J. Bacteriol. **174**:3474–3478.
36. **Schröder, I., S. Darie, and R. P. Gunsalus.** 1993. Activation of the *Escherichia coli* nitrate reductase (*nar GHJI*) operon by NarL and FNR requires integration host factor. J. Biol. Chem. **268**:771–774.
37. **Simons, R. W., F. Houman, and N. Kleckner.** 1987. Improved single and multicopy *lac*-based protein and operon fusion cloning tools. Gene **53**:85–96.
38. **Spiro, S., and J. R. Guest.** 1990. FNR and its role in oxygen-regulated gene expression in *Escherichia coli*. FEMS Microbiol. Rev. **75**:399–428.
39. **Stim, K., and G. M. Bennett.** Personal communication.
40. **Wagner, A. F. V., M. Frey, F. A. Neugebauer, W. Schäfer, and J. Knappe.** 1992. The free radical in pyruvate formate-lyase is located on glycine-734. Proc. Natl. Acad. Sci. USA **89**:996–1000.
41. **Walker, M. S., and J. A. DeMoss.** 1991. Promoter requirements for FNR-dependent activation of transcription of the *narGHJI* operon. Mol. Microbiol. **5**:353–360.

42. **Wanner, B. L.** 1986. Novel regulatory mutants of the phosphate regulon in *Escherichia coli*. J. Mol. Biol. **191**:39–58.
43. **Wu, Y., R. V. Patil, and P. Datta.** 1992. Catabolite gene activator protein and integration host factor act in concert to regulate *tdc* operon expression in *Escherichia coli*. J. Bacteriol. **174**:6918–6927.
44. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.
45. **Zinoni, F., A. Birkmann, T. C. Stadtman, and A. Böck.** 1986. Nucleotide sequence and expression of the selenocysteine-containing polypeptide of formate dehydrogenase (formate-hydrogen-lyase-linked) from *Escherichia coli*. Proc. Natl. Acad. Sci. USA **83**:4650–4654.