Action at a Distance for Specific Repression of the \textit{glpD} and \textit{glpTQ} Genes, and Organization of the \textit{glpEGR} Genes of \textit{Escherichia coli} K-12

by

Bing Yang

Dissertation submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry

APPROVED:

\[\text{Timothy J. Larson, Chairman}\]

\[\text{W.G. Niehaus}\]

\[\text{E.M. Gregory}\]

\[\text{T.O. Sitz}\]

\[\text{S.M. Boyle}\]

January, 1997

Blacksburg, Virginia

\textbf{Key words}: Transcription, Regulation, Glycerol, Repression, \textit{E. coli}
Action at a Distance for Specific Repression of the \textit{glpD} and \textit{glpTQ} Genes, and Organization of the \textit{glpEGR} Genes of \textit{Escherichia coli} K-12

by

Bing Yang

Committee Chairman: Timothy J. Larson

Biochemistry

(ABSTRACT)

Aerobic \textit{sn}-glycerol 3-phosphate (glycerol-P) dehydrogenase is a cytoplasmic membrane-associated respiratory enzyme encoded by the \textit{glpD} gene of \textit{Escherichia coli}. The divergently transcribed \textit{glpACB} and \textit{glpTQ} operons encode the subunits of the anaerobic glycerol-P dehydrogenase, glycerol-P transporter, and glycerophosphodiesterase, respectively.

The \textit{glpD} operon is controlled by cooperative binding of \textit{glp} repressor (GlpR) to tandem operators that cover the -10 promoter element and 30 bp downstream of the transcription start site. The \textit{glpACB-glpTQ} operons operons are controlled by GlpR binding to operators that overlap the \textit{glpA} promoter elements. In this study, two additional operators were identified within the \textit{glpD} structural gene, and three additional operators were identified within the \textit{glpT} structural gene. The internal \textit{glpD} operators
contributed about five- to sevenfold to repression of $glpD$, in the presence of the tandem operators. The internal $glpT$ operators contributed about three- to fourfold to repression of $glpT$, in the presence of $glpA$ operators. Integration host factor (IHF) was found to contribute about twofold to repression of $glpT$ (in the presence of $glpA$ and internal $glpT$ operators) but had no effect on control of $glpD$. Another histone-like protein, HU protein, was found to contribute fourfold to repression of $glpD$ (in the presence of all $glpD$ operators) but had no effect on repression of $glpT$. The results suggest that both $glpD$ and $glpT$ are controlled by formation of repression loops between proximal and distal operators. HU and IHF may assist repression by bending DNA to facilitate loop formation.

The $glpEGR$ operon encodes the $glp$ repressor (GlpR) and two proteins with unknown functions (GlpE and GlpG). The $glpEGR$ genes were previously determined to be co-transcribed only from the $glpE$ promoter. In the work described here, additional promoters downstream of the translational start for $glpE$ were discovered. The strongest of these promoters was located just upstream of $glpG$. In addition, two very weak promoters were found within $glpG$, upstream of the translational start codon of of $glpR$. There are no apparent transcriptional terminators within the $glpEGR$ genes. Therefore, the $glpEGR$ genes form a single operon. The strengths of the individual promoters and combinations of the promoters were compared in detail.
Acknowledgements

I would like to express my deepest appreciation to my advisor, Dr. Timothy J. Larson, for his guidance, patience, and support during the past several years. He has provided a perfect example as a brilliant, decent, diligent, persistent, calm, modest, and honest scientist.

I would also like to thank my committee members, Dr. T. O. Sitz, Dr. W. G. Niehaus, Dr. E. M. Gregory, and Dr. S. M. Boyle for their wise advice and numerous help. Thanks, too, to the generous help from Dr. B. M. Anderson, Dr. W. E. Newton, Dr. Jiann-shin Chen, Dr. J. L. Hess, Dr. B. Storrie, Dr. R. E. Ebel, and Dr. D. R. Bevan. Especially I would like to thank Mrs. Connie Anderson for her kind help, she is a warm-hearted person that I will never forget. Thanks to Dr. P. Bender for his valuable advice and his generous personal help. I also would like to thank Peggy Arnold and Mary Jo Smart for their professional help.

Thanks to the members of the lab, Ali Bhattacharya, Sergey Podkovyrin, Ningyue Zhao, Steven Solow, Keith Ray, Tsuneo Hill, Andy Ewens, Christina Pao, Matt Bunce, for making the lab pleasant.

My personal deepest appreciation goes to my wife, Xiao Wen, for her love and support.
List of Abbreviations

glycerol-P  sn-glycerol 3-phosphate
FDP  fructose 1, 6-bisphosphate
NAD  nicotinamide adenine dinucleotide
ATP  adenosine 5'-triphosphate
dNTP  deoxynucleoside triphosphate
bp  base pairs
kb  kilobase pairs
cAMP  cyclic adenosine 3', 5'-monophosphate
amp  ampicillin
kan  kanamycin
cam  chloramphenicol
tet  tetracycline
sp  spectinomycin
DTT  dithiothreitol
EDTA  ethylenediaminetetraacetate
Tris  tris(hydroxymethyl)aminomethane
IPTG  isopropylthio-β-D-galactopyranoside
ONPG  orthonitrophenyl-β-D-galactopyranoside
X-gal  5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
SDS  sodium dodecyl sulfate
PAGE  polyacrylamide gel electrophoresis

$K_m$  Michaelis constant
# Table of Contents

**LITERATURE REVIEW** .................................................................................................................. 1

Glycerol Metabolism in *Escherichia coli* ....................................................................................... 1

Genetic Regulation of the *g/p* Regulon........................................................................................ 8

DNA Looping in Negative Regulation ............................................................................................ 16

Accessory Proteins IHF, HU and FIS ............................................................................................ 19

**PART I** ........................................................................................................................................ 25

Action at a Distance for Transcription Control by *g/p* Repressor of *Escherichia coli* .......... 25

**INTRODUCTION** ......................................................................................................................... 25

**MATERIALS AND METHODS** .................................................................................................. 28

**RESULTS** .................................................................................................................................. 49

**DISCUSSION** .............................................................................................................................. 86

**PART II** ....................................................................................................................................... 95

Transcriptional organization of the *g/pEGR* operon .................................................................. 95

**INTRODUCTION** ......................................................................................................................... 95

**MATERIALS AND METHODS** .................................................................................................. 98

**RESULTS** .................................................................................................................................. 106

**DISCUSSION** .............................................................................................................................. 120

**LITERATURE CITED** .................................................................................................................. 123
List of Illustrations

Fig. 1. Locations of the glp operons on the E. coli linkage map........................................2
Fig. 2. Glycerol-P metabolism in E. coli.............................................................................4
Fig. 3. Alignment of GlpR with some of the members of the DeoR family...........13
Fig. 4. DNA loop model.................................................................................................17
Fig. 5. Nucleotide sequence of the glpD control region..............................................30
Fig. 6. Nucleotide sequence of the divergently transcribed glpACB-glpTQ control region.................................................................34
Fig. 7. Identification of glpD operators by DNase I footprinting..............................57
Fig. 8. Lack of binding of the glp repressor to mutated O_{1,4} in the presence of O_{3,1} and O_{3,2}.................................................................60
Fig. 9. DNase I footprinting of the glpD internal operators (in the absence of O_{3,1} and O_{3,2}).................................................................62
Fig. 10. Effects of internal glpT operators, IHF, and HU on repression of glpT-lacZ transcriptional fusions.........................................................66
Fig. 11. DNase I footprinting of the wild type glpT internal operator region............71
Fig. 12. DNase I footprinting of the mutated internal glpT operators.................75
Fig. 13. DNase I footprinting of the wild-type IHF sites within glpT...................81
Fig. 14. DNase I footprinting of the mutated IHF sites.............................................85
Fig. 15. Model of the IHF-GlpR-operator complex for control of glpT
transcription

Fig. 16. Nucleotide sequence of the glpE, glpG, and glpR genes and the deduced amino acid sequence.

Fig. 17. Localizations and contributions of the glpEGR promoters.

Fig. 18. Primer extension analysis of the glpEGR promoters.

Fig. 19. Identification of the glpG promoter.
List of Tables

Table 1. Oligonucleotides used in the study of *glpD* ............................................. 29
Table 2. Oligonucleotides used in the study of *glpT* ............................................. 31
Table 3. Plasmids used or constructed in the study of *glpD* ..................................... 35
Table 4. Plasmids used or constructed in the study of *glpT* ..................................... 38
Table 5. Strains of *E. coli K-12* used or constructed in this study ............................. 41
Table 6. Titration of chromosomally-encoded GlpR by different combinations of multiple *glpD* operators .......................................................... 50
Table 7. Effect of internal *glpD* operators on regulation of *glpD-lacZ* transcriptional fusions .......................................................... 53
Table 8. Effect of IHF and HU on regulation of *glpD* transcription ............................ 63
Table 9. IHF encoded by a plasmid complements the effect of chromosomal *himA* deletion on *glpT* transcription .............................................. 77
Table 10. Oligonucleotides used in this study ............................................................. 99
Table 11. Plasmids constructed in this study ............................................................. 104
Table 12. Repression of Φ(*glpK-lacZ*) in Δ(*glpEGR*) strain NZ45 by plasmids carrying portions of *glpEGR* ......................................................... 107
Table 13. Mapping of promoters within the *glpEGR* genes ..................................... 110
Table 14. Expression of *glpEGR* is not controlled by GlpR .................................... 119
LITERATURE REVIEW

Glycerol Metabolism in Escherichia coli

sn-Glycerol 3-phosphate (glycerol-P) and its derivatives are ubiquitous in nature. Acylated derivatives of glycerol-P, glycerophospholipids such as phosphatidylglycerol, phosphatidylethanolamine and cardiolipin, are present in virtually all biological membranes (1). Thus, glycerol-P is a direct precursor for phospholipid biosynthesis (1). Glycerol-P and its precursors, when present in excess, are used as energy sources by a variety of microorganisms (2). In order to maintain the cellular concentration of glycerol-P at a level allowing an optimal rate of phospholipid synthesis, organisms must coordinate the levels of the catabolic activities with the levels of the phospholipid biosynthesis activities.

The genes encoding the enzymes responsible for catabolism of glycerol-P and its precursors comprise the glp regulon of E. coli and are arranged in five different operons (Fig. 1; (3)). The glp genes comprise a regulon because they are all subject to negative transcriptional regulation by the glp repressor. The affinity of the glp repressor for its operator sites on the DNA is decreased upon binding of glycerol-P, the inducer for the regulon (2,3).

The glp genes have been mapped to three different regions on the E. coli
Fig. 1. Locations of the glp operons on the E. coli linkage map.
chromosome (Fig. 1: (3)). The *glpF*K operon, encoding the glycerol diffusion facilitator (GlpF), glycerol kinase (GlpK), and a protein of unknown function (GlpX), respectively, maps to min 89 (4-6). The *glpTQ* operon encoding the glycerol-P permease (GlpT) and the periplasmic glycerophosphodiesterase (GlpQ), respectively, maps to min 51 (6-9). The divergently transcribed *glpACB* operon encodes the three subunits of the anaerobic glycerol-P dehydrogenase (GlpA,GlpB,GlpC) (10,11). The *glpD* gene, encoding the aerobic glycerol-P dehydrogenase (GlpD), is located near min 77 (3,6,12,13), and is adjacent to the divergently transcribed *glpEGR* operon (14,15). The *glpR* gene encodes the specific repressor (GlpR) for the *glp* regulon (14). The functions of the proteins GlpE/GlpG encoded by the *glpE* *glpG* genes are still unknown. Comparisons of the amino acid sequence revealed that the proteins most similar to GlpE/G were the corresponding proteins of the Gram-negative bacterium *Haemophilus influenzae* (16). Except *glpD*, all the other *glp* genes are present in *H. influenzae*. One major difference in that case is that *glpG* is upstream of *glpR* while *glpE* is convergently transcribed toward *tpIA* (encoding triosephosphate isomerase) at a different locus (16).

The metabolism of glycerol-P in *E. coli* is summarized in Fig. 2. The entry of glycerol into the cell is mediated by GlpF (3). GlpF is the only known cytoplasmic membrane porin in *E. coli* that catalyzes the energy-independent equilibration of intracellular and extracellular glycerol concentrations (3). In contrast to the typical twelve membrane spanning segments found in other transport proteins which transport substrates through hydrophilic cores, analysis of the 30 kDa of GlpF revealed only six
Fig. 2. Glycerol-P metabolism in *E. coli.*
potential membrane spanning helices (17). It is conceivable that a homodimer of GlpF could be arranged to form a similar transport channel containing 12 membrane-spanning helices. The estimated pore size of the channel is 0.4 nm (3). In addition to glycerol, GlpF also facilitates the transport of glyceraldehyde, glycine, urea, and several polyhydric alcohols (18).

Once glycerol enters the cell, it is phosphorylated by GlpK to form glycerol-P. The addition of the charged phosphate group prevents the passage of glycerol-P back through the cytoplasmic membrane and therefore the substrate is trapped inside the cell. Thus GlpK is the pace-maker for the dissimilation of glycerol. The active glycerol kinase (EC 2.7.1.30, ATP: glycerol 3-phosphotransferase; $K_m$ glycerol: 10 μM; $K_m$ ATP: 80-100 μM) is a homotetramer of 56 kDa subunits (19) and is subject to two forms of noncompetitive inhibition (2i). One inhibitor is fructose 1,6-bisphosphate ($K_i = 0.5$ mM) (19), an intermediate of glycolysis and gluconeogenesis. This feedback inhibition ensures that the cell reduces production of glycerol-P when too much fructose 1,6-bisphosphate accumulates. Growth on glucose results in a second form of allosteric inhibition exerted by the nonphosphorylated form of the enzyme IIα of the bacterial phosphotransferase system (PTS) (22,23). In the presence of glucose, this enzyme form is predominant. This also prevents the cell from using glycerol in the presence of the most favorable substrate, glucose. The activity of glycerol kinase is greatly reduced upon binding of enzyme IIα, in the presence of glycerol (22,24). The specific interactions between glycerol kinase and enzyme IIα were recently revealed by the analysis of the
crystal structure of the kinase-enzyme IIA
c complex (25).

The activity of glycerol kinase is activated by interaction with GlpF (26).
Comparison of the glycerol transport kinetics of the wild type strain with those of the
\( glpI^- \) mutant shows that both the \( V_{\text{max}} \) and the \( K_m \) of glycerol kinase are changed upon
interaction with GlpF. The \( K_m \) for glycerol was lowered from 50 to 5.6 \( \mu \)M, and the \( V_{\text{max}} \)
increased from 0.1 to 13.1 nmol/min/10^9 cells. This effect is similar to that discovered in
mitochondria where the glycerol kinase is activated by interaction with the mitochondrial
porin (27,28). It is surprising that the \( glpF \) mutant strain still can grow on glycerol,
despite the low activity of the glycerol kinase (26). This may be due to the action of
other undefined glycerol utilization pathways that are not initiated by phosphorylation
(2).

Extracellular glycerophosphodiesters diffuse through the outer membrane porins
and are hydrolyzed by the periplasmic enzyme GlpQ to give an alcohol and glycerol-P
(7). The \( K_m \) for phosphodiesters varies from 0.24 mM for glycerophosphoethanolamine
to 1.0 mM for glycerophosphoinositol (29). The native form of GlpQ is a homodimer of
40 kDa subunits.

The transport of glycerol-P into the cytoplasm of the cell is specifically carried
out by the GlpT permease with the exchange of two P, anions (30). GlpT displays a size
of 50 kDa on SDS-PAGE (31). The fact that missense mutations in the C-terminal of the
protein result in a negative dominant phenotype suggests that this transporter might have
an oligomeric structure (31). Analysis of the amino acid sequence suggests that GlpT is
similar to other transport proteins, with six hydrophobic domains traversing the cytoplasmic membrane in an α-helical conformation flanking a central cytoplasmic hydrophilic region (31). Import of glycerol-P is also carried out by the \textit{ugp} transport system (32) or the hexose-phosphate transporter \textit{UhpT} (30). These two are not the dominant pathways for transport of glycerol-P, and are only induced under conditions of Pi limitation or when hexose-P is present, respectively (30,32).

Under aerobic conditions, aerobic glycerol-P dehydrogenase (GlpD) converts glycerol-P to dihydroxyacetone phosphate (DHAP) with concurrent reduction of FAD to FADH\textsubscript{2} (13). DHAP enters the glycolytic pathways. FADH\textsubscript{2} passes electrons on to ubiquinone and ultimately to oxygen or nitrate (3). GlpD is a primary dehydrogenase in the respiratory chain and is associated with the cytoplasmic membrane. The purified GlpD is a homodimer of 58-kDa subunits with a \( K_m \) of 0.4 mM for glycerol-P (33).

Under anaerobic conditions, glycerol-P is converted to DHAP by anaerobic glycerol-P dehydrogenase (GlpACB) (11). GlpA and GlpC are the catalytic subunits and the GlpB is the membrane anchor. The anaerobic enzyme has an apparent \( K_m \) of 0.3 mM for glycerol-P. Glycerol-P is oxidized to DHAP with the concurrent reduction of cofactors FAD/FMN (detailed mechanism unknown) (11). The electrons are finally accepted by either fumarate or nitrate, catalyzed by fumarate or nitrate reductase (34).

Glycerol-P can also be synthesized from DHAP by \textit{gpsA}-encoded glycerol-P synthase (35). This is a reverse reaction of the one catalyzed by glycerol-P dehydrogenase. Since glycerol-P synthase is feedback regulated by the product, glycerol-
P, the actual level of glycerol-P in the cell is maintained tightly with or without the exogenous glycerol-P obtained by the pathways described above. The biosynthetic role of glycerol-P is to serve as a precursor of the phospholipids for the cell membrane. The committed step is the acylation of glycerol-P, catalyzed by the plsb-encoded acyltransferase with a $K_m$ of 150 $\mu$M for glycerol-P (36). The second step further acylates 1-acylglycerol-P and the resulting product is phosphatidic acid (37), which is an intermediate required for synthesis of the various species of phospholipids (1,38).

**Genetic Regulation of the glp Regulon**

The glp operons are subject to three types of transcriptional regulation: 1. Catabolite repression in the presence of glucose. 2. Respiratory control in the presence of oxygen versus alternate terminal electron acceptors. 3. Specific repression by GlpR in the absence of the inducer, glycerol-P.

1. Catabolite repression

The presence of glucose in the growth medium results in a decrease in the intracellular level of cAMP. cAMP is the effector molecule for cAMP receptor protein (CRP). cAMP-CRP complex is a global regulator that activates numerous genes, generally those involved in utilization of energy sources other than glucose. All genes subject to positive control by CRP form a modulon. All of the glp genes are members of the CRP modulon. Binding sites for CRP were identified in the control regions for each of the glp operons by using DNase I footprinting (17,39,40). The order of the sensitivity of the glp operons to the presence of glucose in medium is $glpTQ > glpFKX = glpACB >$
glpD (17). The differential sensitivity might be due to the differences in the number and position of the CRP sites in the promoters and the degree of the similarity of the various sites to the CRP consensus binding sequence. In the cases of the glpTQ, glpFKX and glpACB promoters which do not contain typical consensus -35 promoter regions, CRP binding sites overlap the -35 region and thus belong to the class II CRP sites (centered at -41.5, as in the gal operon). In the case of the glpD promoter, one CRP site is centered at -63.5. This places the glpD promoter in the class I type of CRP-activated promoters (similar to the lac operon). Class II promoters are typically more dependent on CRP than the class I promoters (41). This may be one reason why the glpD promoter is the least sensitive among the glp operons to catabolite repression. Overall, the glpD promoter has the most recognizable -10 and -35 promoter consensus sequences and therefore transcription is not very dependent on CRP function. Comparison of the sequences of all the CRP sites revealed that the CRP site in the glpT promoter matches the consensus most closely. Binding of CRP to three widely separated sites in the glpACB-glTQ control region might contribute to co-activation of the divergent promoters. Binding of CRP to similarly spaced sites in the ans promoter synergistically activates transcription (42). CRP binds to tandemly repeated sites in the glpFKX promoter.

2. Respiratory control

Unlike catabolite repression, which affects all of the glp operons, respiratory signals are sensed primarily by glpACB and glpD genes (17,34,43,44). Under anaerobic growth conditions, the FNR protein activates glpACB expression (about 20-fold, (17))
and thus the anaerobic glycerol-P dehydrogenase level is maximal anaerobically with fumarate present as electron acceptor (43). FNR is a global transcriptional regulator which specifically activates genes encoding anaerobic respiration proteins and represses some genes encoding aerobic proteins in the absence of oxygen. The redox state of an Fe-S center in FNR is the signal for activation of FNR. In the absence of oxygen, the reduction of the Fe-S center is thought to increases the dimerization of the FNR protein and hence DNA-binding affinity of the protein (45-47).

Under aerobic growth conditions, expression of the glpD-encoded dehydrogenase is maximal. Twofold anaerobic repression of the glpD gene is mediated by the arcA/arcB-encoded two-component regulatory system (34,43,44,48-50). The arcB gene encodes a sensor kinase that controls the phosphorylation of the arcA-encoded response regulator in response to the redox state of the cell (51). The phosphorylated form of ArcA, which predominates anaerobically, binds the regulatory regions of various genes and thus activates or represses different genes to adjust to the aerobic-anaerobic environmental changes. Recently the consensus sequence of the ArcA-Pi binding site was determined by DNA-binding studies using different arc-regulated genes and purified ArcA-Pi protein (51). The sequence protected by DNase I footprinting was also compared to numerous ArcA-Pi binding sequences in other arc-regulated genes (51). For binding of monomeric ArcA-Pi, the 10 bp consensus sequence ([A/T]GTTAATTA[T/A]) is flanked by random 10 bp sequences. Binding of the dimeric ArcA-Pi covers about 60 bp, with the 10 bp consensus sequence in approximate
palindromic positions (51). One plausible ArcA-Pi binding sequence in glpD was found overlapping the -10 sequence of the promoter. Because the mismatches of this site compared to the consensus sequence is 4 bp, the anaerobic repression conferred by arcA/arcB on glpD is only twofold (52).

Previous studies also showed that glpT is subject to respiratory control of anaerobic repression (17,43). Repression is fivefold for glpT (17). This is not unexpected, since the divergent glpT and glpACB genes apparently share transcriptional control elements. Binding of FNR at the glpA promoter might simultaneously activate the transcription of glpT by interaction with CRP bound at the glpT promoter.

3. Specific repression by GlpR

All of the glp structural genes are negatively regulated by GlpR. It is assumed that binding of GlpR to the control regions blocks the function of RNA polymerase. The operator sites in the control regions of each of the glp genes have been identified by DNase I footprinting (17,39,40). By using site-directed mutagenesis, the 20-bp palindromic operator consensus sequence was revealed as having the left half-site WATKYTCGWW (W = A or T, K = G or T, Y = C or T) (53). All substitutions at positions 3, 4, 5, and 8 (starting from the center of symmetry of the palindrome) caused a severe decrease in repressor binding. Substitutions at other positions had modest or no effect on GlpR binding. In most glp operons (glpD, glpACB, and glpFKX), GlpR binds to tandemly repeated operators of hyphenated dyad symmetry in the promoter regions. Upon binding of glycerol-P by GlpR, the affinity of GlpR for the operators is decreased
and thus transcription is derepressed. Therefore, the system for specific negative control senses the amount of the substrate (glycerol-P) and maintains optimal expression of the catabolic genes. Previous studies by Lin and coworkers found that the order of sensitivity of the glp operons to GlpR is glpD > glpT > glpK (2,3). This relative order of sensitivity was obtained by measuring the activity of each enzyme in a strain harboring a thermolabile glp repressor following growth at various temperatures (52). More recent studies using transcriptional glp-lacZ fusions indicated that the glpFKX operon is controlled most tightly by GlpR (17). However, my discovery of the contribution of the internal glpD operators to repression as outlined below, might shed light on the actual order of the sensitivities of the glp operons to GlpR. The previously studied glpD-lacZ fusion (17) did not contain the internal glpD operators and therefore its regulation might not reflect the natural situation, where repression may be stronger. Three internal glpT operators were also discovered within the structural gene, as described below. These operators were found to contribute about three- to fourfold to glpT repression. In the cases of both glpD and glpT, GlpR binding at widely separated operators, with the intervening DNA looped out, seems to be responsible for full repression.

The glpR gene has been cloned and sequenced (54). The deduced amino acid sequence of GlpR displays similarity to the DeoR family of bacterial transcriptional regulators (Fig. 3; (54,57)). Members of this family contain approximate 250 amino acids with the helix-turn-helix (HTH) DNA-binding motif located near the amino terminus, and sugar phosphate inducer binding and oligomerization functions located in
Fig. 3. Alignment of the members of the DeoR family (there are at least 15 additional members, from a variety of bacterial species, present in the data bases (54). Completely conserved and highly conserved (4 of 5) residues are indicated with "**" and "*", respectively. DeoR is the repressor of the deo system of E. coli. HI GlpR is the glp repressor of Haemophilus influenzae (16), PA GlpR Is the glp repressor of Pseudomonas aeruginosa (55), and o269R is an hypothetical repressor of E. coli that is closely related to the glp repressor (56). Positions of preferential trypsin cleavage in glp repressor with or without bound inducer are indicated above the alignments. Positions of substitutions yielding GlpR* variants (T10II, A140T, D207V) are indicated.
the carboxyl-terminal domain. Upon binding of the sugar phosphate inducer, the binding affinity of the repressor for the operator decreases and thus transcription turns on. The purified GlpR is a tetramer of the identical 30-kDa subunits under native conditions (58). In the action of a typical prokaryotic repressor, the amino acid side chains of the second helix of the HTH (the “recognition” helix) make specific contacts with base pairs in the major groove of the operator DNA (59). A repressor dimer binds one palindromic operator with each monomer binding to half of the palindromic sequence. The cooperative binding of the two monomers increases the binding affinity of the repressor.

Comparison of the sequences of the helix-turn-helix motif of GlpR and DeoR revealed that the second helices are strikingly similar:

<table>
<thead>
<tr>
<th>Recognition Helix turn Helix</th>
<th>Operator half-site</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>deo</em> LKDAAALLGVSEMTIRRDLE</td>
<td>T A T G T A G A A *</td>
</tr>
<tr>
<td><em>glp</em> 22-TEELVEHFSVPQTRRDLENN-42</td>
<td>T A T G T C G A W *</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q34M</th>
<th>F33E, Q34M</th>
<th>E23K, F33E, Q34M</th>
<th>R35A (neg. control)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>EM</td>
<td>EM</td>
<td>A</td>
</tr>
<tr>
<td>M</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* indicates center of operator symmetry

It seems that amino acids 33 and 34 (PQ of GlpR) bind the C-G base pairs in _glp_ operator, while the corresponding amino acids (EM of DeoR) bind the A-T base pairs in _deo_ operator (position 4). However, site-specific change of PQ to EM in GlpR resulted a repressor with no binding affinity to either _glp_ or _deo_ operators (54). The same results were obtained with Q34M and the E23K, F33E, Q34M forms of _glp_ repressor (54). The altered repressor encoded by a plasmid had a strong negative dominant effect on a normal, chromosomally encoded repressor, suggesting the altered repressor protein is
stable and able to oligomerize (54). These results prove that the proposed helix-turn-helix is important for DNA-binding. The lack of the binding of GlpR^{P3E-Q3M} to the deo operator further suggests that besides the second helix, other aspects of the 3-dimensional structure of the repressor protein are important for the assembly of the DNA-binding domain. More detailed studies of the DNA-binding mechanism are underway.

The small helix-turn-helix DNA-binding domain and the large effector binding domain of many DNA-binding proteins can be separated from one another by limited proteolysis (60,61). This is also the case for GlpR using trypsin treatment (Bhattacharya and Larson, unpublished data). In the absence of glycerol-P, trypsin digestion of GlpR yielded two domains of 22 and 8 kDa, respectively, as assessed by SDS PAGE. In the presence of glycerol-P, digestion of GlpR with trypsin yielded only a 24 kDa fragment, as assessed by SDS PAGE. N-terminal sequence analysis of the 22 and 24 kDa fragments revealed that trypsin preferentially cleaves after K71 and R51 in the absence and presence of glycerol-P, respectively (Fig. 3). Binding of glycerol-P apparently changes the conformation of GlpR in the glycerol-P binding domain, and destabilizes the HTH DNA-binding domain.

During the early studies by Lin on the glycerol system, a noninducible strain ECL89 presumably carrying an altered GlpR was isolated (62). The glpR12 gene encoding the altered GlpR was subcloned and sequenced. Although several mutations were found compared to the wild-type glpR sequence, one substitution at T1011 (Fig. 3)
was found to be sufficient for conferral of noninducibility (63). The residue at position 101 is either threonine or serine for all the members of the DeoR family, and is highly conserved in these repressors (Fig. 3). Changes of T101S and T101A of GlpR also resulted in noninducibility, indicating that serine cannot replace threonine in this case. The altered GlpR carrying the T101I substitution was partially purified and the noninducibility was shown to be due to the inability of GlpR$^{T101I}$ to bind the inducer compared with the wild type purified GlpR (Hill and Larson, unpublished data). The trypsin cleavage pattern of this altered GlpR with or without the inducer was identical to the wild-type GlpR without the inducer, providing additional evidence that the protein is unable to interact productively with the inducer. Random mutagenesis of the glpR gene provided other noninducible variants of GlpR (A140T and D207V). These two positions (140 or 207) are close or identical to positions of LacR that are important for inducer binding (64).

**DNA Looping in Negative Regulation**

Repressor proteins are commonly believed to repress gene expression by sterically hindering the binding of RNA-polymerase to the promoter. Thus, operator sites are generally adjacent to or overlap with the promoter (65). However, more complete characterization of repression systems has revealed multipartite operator sites in nature. To explain how widely separated operators function in regulation, DNA looping models have been proposed (66-69). Fig. 4 describes the basic model of DNA looping. In this model, dimeric repressor proteins bound at distant operators cooperate by direct
Fig. 4. DNA loop model. Solid circles represent GlpR monomers (R = GlpR). Double solid lines represent DNA. $O_1$ and $O_2$ represent two widely separated operator regions on the DNA. This model does not intend to depict any of the possible topological forms but simply shows a loop.
protein-protein interactions to form repressor tetramers, with the intervening DNA
looped out. The so-formed higher order DNA-protein complex is believed to be more
effective in blocking transcriptional isomerization, initiation or promoter clearance by
RNA polymerase (65). Three classic examples of DNA looping involved in repression of
the lac, gal, and deo operons are summarized below.

1. Tripartite Operators in the lac Operon

   There is one CRP site (centered at -61.5) and three operators at positions +11
   (O₁), +401 (O₂) and -82 (O₃) in the lac operon that exert regulatory functions. O₁ was
   thought to be the sole operator responsible for repression mediated by the tetrameric lac
   repressor (LacI) (70). However, sequence analysis revealed two additional operator-like
   sequences (O₂ and O₃), which, when mutated, resulted in a two to three fold loss of
   repression, indicating that these two operators have repressive functions in vivo (71).
   Mutation of O₁ alone resulted in a five to fifty fold loss of repression in the presence of
   O₂/O₃, instead of a 500-fold derepression in the absence of the O₂/O₃ (71), further
   suggesting a contribution of O₂/O₃ in vivo. Under the control of a nonlooping mutant
   LacI (functional in dimer but defective in tetramer formation), repression was decreased
   similarly by about 3-fold. These data suggest that cooperative binding of a LacI tetramer
to either O₁/O₂ or O₁/O₃ with the intervening DNA forming a loop might strengthen the
steric inhibition of RNA polymerase and thus cause full repression.

2. Bipartite Operators in gal

   The gal operon contains two promoters P1 and P2 with the transcription start
point separated by 5 bp (72,73). CRP protein binding at its recognition site (centered at -41.5) activates P1 but represses P2. There are two operators O_\text{E} and O_\text{I} centered at positions -60.5 and +53.5, respectively (67,69,74). Full repression exerted by the \textit{gal} repressor needs the presence of both operators, as occupation of either of the operator alone confers poor (O_\text{E}) or no (O_\text{I}) repression (75). If operators O_\text{I} and O_\text{E} are genetically substituted by \textit{lac} operators, the nonlooping Lacl repressor, which forms dimers but not tetramers, fails to repress the \textit{gal} promoter (76). Therefore, repressor binding at both operators with the intervening DNA forming a loop is responsible for full repression.

3. Tripartite Operators in the \textit{deo} Operon

The \textit{deo} operon contains two promoter P1 and P2. While P1 is negatively controlled by the \textit{deo} repressor (DeoR) alone, P2 is repressible by both the DeoR and CytR (77). There are three operators centered at -286 (O_\text{E}), -7 (O_\text{I}) and +592 (O_\text{O}). DeoR-mediated repression is enhanced 20- to 30-fold when repression of a \textit{deo-lacZ} fusion is measured in the presence of both O_\text{E}/O_\text{O} compared with constructs having either O_\text{I} or O_\text{O} alone, suggesting a DNA looping mechanism involving the repressor binding the two operators (78). Full repression requires all three operators. The DNA looping mechanism was further supported as repression could be achieved by varying the distance between O_\text{I} and O_\text{O} within a range of 24-997 bp (78,79) or even up to 1-5 kb (80).

**Accessory Proteins IHF, HU and FIS**

It is common knowledge that eukaryotic chromatin is organized around the
nucleosome, a bead-like structure in which DNA is wrapped around the surface of an octamer of histone protein subunits (81). This organization is believed to provide the differential accessibility to specific DNA sequences during developmental processes (82). In the past decades, bacterial histone-like proteins have been discovered and studied. These include the HU protein, integration host factor (IHF), factor for inversion stimulation (Fis) and HNS protein. The absolute physiological roles played by these proteins in the cell are still unknown. In other words, they cannot be identified as contributing to an isolated, bead-like structure. In many cases, these proteins exert their functions by assisting other enzymes or regulatory proteins, hence the name accessory protein. Cells deficient in each one of these proteins are viable with modest phenotypic changes. For example, HU-deficient strains give small, slow-growing colonies. Colonies of IHF-deficient strains look normal on plates, although the synthesis of dozens of individual protein were differentially affected compared with the wild type strain as assessed on two-dimensional gels (83). Despite the lack of specific roles in the cell, these proteins are often involved in the regulation of gene expression, presumably by bending the DNA to exert their assisting functions (84).

IHF is a heterodimeric protein (about 20 kDa) with the two subunits HimA (encoded by himA, 11.9 kDa) and HimD (encoded by himD, 0.9 kDa) (85). The expression of IHF increases about 5-fold when cells enter the stationary phase (IHF is present at a level of 0.5 to 1.0 ng per μg of cell protein during exponential growth) (86). This is probably due to the increased synthesis of the IHF in stationary phase (87). The
expression of the IHF is autoregulated (86,88). The consensus sequence for IHF binding is WATCAANNNNTTR (W = A or T, R = A or G) (89). Footprinting of IHF usually covers about 28-30 bp with the 13 bp consensus sequence asymmetrically located toward one side of the protected DNA (89). IHF binds its specific sites with an affinity of 1-20 nM (90,91). Despite the much higher intracellular concentration of IHF (about 6 μM during exponential growth), the effective intracellular concentration of IHF is only about 40 nM (92), suggesting the possibility that most IHF protein is nonspecifically bound to the DNA.

IHF plays important roles in a number of cellular processes including recombination, replication, transposition and regulation of gene expression (93). In the process of lambda integration and excision, the major player is the phage-encoded integrase (Int) protein. The lambda attachment site for recombination assumes a complicated higher-order structure where IHF is thought to permit assembly of Int into such a structure by simply bending the DNA (94,95). Replication of oriC-dependent minichromosomes requires IHF probably because IHF assists the DnaA protein to melt the 13-mer regions in the minimal origin, which is essential for replication (96). Transposition of IS10 and Tn10 also requires IHF, because IHF is needed for transposase function (97). Binding of IHF results in bending of the DNA, which brings the upstream activator and the downstream RNA polymerase together for activation of σ^54-dependent promoters, like the activation of the ndfA (98) and the glnHlp2 genes (99). This activation is postulated to involve a DNA-loop mechanism. However, IHF could also directly
repress and activate transcription via contact with RNA polymerase. For example, IHF represses the \( ilvP_c \) promoter but activates the downstream \( ilvP_c,2 \) promoter (100,101). My elucidation of a role for IHF in mediating formation of GlpR-operator complex for control of \( glpT \) transcription expands the physiological roles played by IHF.

HU is a 20 kDa heterodimer of two subunits (encoded by \( hupA \) and \( hupB \)) (93). Homodimers of either of the subunits have substantial function. Therefore, deletion of both subunits is required for study of HU function. HU is a sequence-nonspecific DNA-binding protein. The cellular content of HU is estimated to be 2 to 5 ng per \( \mu \)g of total cell protein, or about 20 \( \mu \)M (102). Transcription of both \( hupA \) and \( hupB \) genes is activated by CRP (103). FIS stimulates the transcription of the \( hupA \) genes, but represses that of the \( hupB \) gene (103).

HU was found to be important in the process of the transposition of bacteriophage Mu (104). In the early stage of the transposition reaction, the formation of a stable synaptic complex as Mu transposase cooperates to assemble a tetramer that juxtaposes the ends of Mu is essential. HU is thought to bend the DNA and thus assists the formation of the higher order structure (105). HU is also important in the process of site-specific inversion of the flagellin promoter as HU, together with another accessory protein FIS, assists the recombinase Hin to form the required structure of the invertosome (106-108). Recently it was reported that HU helps two- to threefold in the repression of the \( gal \) operon as it assists in DNA-loop formation (109). This is consistent with the results of my work where HU assists about fourfold in repression of \( glpD \)
transcription. HU probably helps DNA-loop formation as GlpR binds to widely-separated glpD operators.

Amino acid sequence analysis revealed that the subunits of IHF and HU are homologous (93). The sequence of the four subunits of HU and IHF can be aligned without major gaps and have more than 45 percent identical or similar residues within the alignment (93). An X-ray crystallographic structure has been obtained using the homodimeric HU protein from Bacillus stearothermophilus (110,111). Since then, models of DNA-protein complexes have been built and studied in detail (93). Since IHF and HU share sequence similarity, functional interchangeability of these two proteins is found in some cases. For example, assembly of a functional attL site can be accomplished by Int either with IHF or HU (112). Likewise, HU successfully replaces IHF in several other in vitro systems, e.g., during initiation of replication at oriC minichromosomes and transposition of transposon Tn10 (97). However, the results of my work show that HU helps repression of glpD, but not of glpT. IHF helps repression of glpT, but not of glpD. The interchangeability of these two proteins might be condition-dependent.

FIS (encoded by fis) has a molecular mass of 11.2 kDa and is a dimer in solution. It binds to highly degenerate sequences with the consensus [T/G]NNYRNN[A/T]NNYRNNC (Y = C or T, R = A or G) (113). The expression of FIS varies during cell growth from highest after nutritional up-shift and leveling off during exponential growth (114,115). The transcription of fis is under stringent control (RelA...
dependent) and is repressed by amino acid starvation (116).

The essential function of FIS in \textit{E. coli} is as obscure as IHF or HU. However, FIS exerts a modulatory effect on a number of cellular processes. As described above, FIS assists in site-specific DNA inversion reactions catalyzed by the phage-encoded DNA invertases Gin, Cin and Hin (113,117,118). FIS was also shown to participate in other processes such as stimulation of phage lambda excision (114,119), and \textit{oriC} function (120), probably in a way similar to that described for IHF. In addition, binding of FIS to its targets upstream of ribosomal RNA promoters was found to activate the transcription at least fivefold (84). Recent discovery by comparison of the protein expression in isogenic \textit{fis}^{-} and \textit{fis::cat} strains revealed that FIS is a transcriptional modulator involved in the regulation of many aspects of metabolism in \textit{E. coli} (84). Northern analysis suggested that FIS represses some genes (\textit{glpR, deoR, cytR}), corepresses \textit{milADR} and activates some other genes (\textit{cdd, udp} and \textit{rbsR}) (84). Computer analysis of the \textit{glpEGR} sequence revealed several FIS binding sites, with one of them overlapping the CRP site for the \textit{glpE} promoter and several other sites located further downstream. Binding of FIS to this region could interfere with or exclude the binding of CRP and thus decreases or eliminate the activation by CRP. This could explain why \textit{glpR} expression is apparently higher in a \textit{fis}-deficient strain compared with the expression in an isogenic wild-type strain (84).
PART I

Action at a Distance for Transcriptional Control by glp Repressor of

Escherichia coli

INTRODUCTION

Negative regulation of glpD transcription is conferred by GlpR binding to tandem glpD operators overlapping the glpD promoter (39). Lin and coworkers found that the glpD operon was controlled more tightly by GlpR than was either the glpTQ or glpFKX operon (30-, 26.5-, 3.2 fold, respectively (52)). This conclusion was reached after measuring the activities of enzymes encoded by these genes (glycerol-P dehydrogenase, glycerol-P permease and glycerol kinase) in extracts of a strain harboring a thermolabile glp repressor after growth at various temperatures (52). In contrast, it was recently found that the glpFKX operon was more sensitive than the glpD operon to repressor by using transcriptional fusions of the promoter-operator regions to lacZ (17). The discrepancy may be due to the presence of other potential operators within the glpD coding region that contribute additionally to negative regulation or to differential posttranscriptional regulation. Computer analysis revealed four potential operator sites within the glpD structural gene.

Common cis-acting regulatory elements are important for coordinate control of transcription of the divergent glpACB (glpA) and glpTQ (glpT) operons, whose initiation
sites are separated by 132 bp (10,40). Previous footprinting studies revealed four GlpR binding sites within the *glpA*-proximal portion of the control region and a weak GlpR binding site (O7) that overlapped the CRP site for the *glpT* promoter (40). Mutation of the *glpA* operators or insertions between the two transcriptional start sites had differential effects on repression of both operons (40). Repressor control of *glpT* transcription was mostly dependent on the *glpA* operator O1, centered at -81.5 with respect to the *glpT* transcription initiation site and overlapping the CRP/FNR site for *glpA*. Individual mutation of the more distal operators O2, O3 and O4 strongly reduced repression of *glpA* but also influenced control of *glpT*. There was no apparent dependence on O7 for repressor control of *glpT*. The results of these studies suggested that repressor-mediated DNA looping was involved in control of both operons (40).

Since previous studies on repression of *glpT* were done using fusions with the *glpT* gene truncated upstream of the translational start point, the contributions of putative distal regulatory elements may have been overlooked. In order to obtain a more complete picture of *glpT* transcriptional regulation, the entire *glpT* sequence was searched for sequences resembling the *glp* operator consensus sequence. Three putative *glp* operators were found within the *glpT* coding region.

The distance between the tandem *glpD* operators and the putative *glpD* internal operators is several hundred bp. The distance between *glpA* and the putative internal *glpT* operators is at least 394 bp (the distance between O1 and O7). GlpR binding to the widely separated operators in both cases with the intervening DNA looped out might
increase the repression conferred by GlpR, similar to repression of the lac, gal, and deo operons described above. IHF or HU might play a role in GlpR-mediated repression of glpD and glpTQ at a distance. If a repressor-mediated DNA loop or nucleo-protein complex forms as GlpR binds at widely separated operators. Binding of IHF or HU might facilitate the formation of such a looped protein-DNA structure by bending the DNA. A computer search revealed five sequences resembling the IHF consensus sequence within the glpAT control region and the adjacent glpT coding region. Several potential IHF binding sites were also found within glpD. The work described in Part I of this dissertation was done in order to discover more details of repression control of these two operons.
MATERIALS AND METHODS

Materials. The reagents for standard recombinant DNA techniques and DNA sequencing were obtained from New England Biolabs or Amersham Life Science, Inc. Oligonucleotides (Tables 1 and 2) were synthesized on an Applied Biosystems 381A DNA synthesizer and purified as recommended by the manufacturer. New England Nuclear Corp. supplied $[\gamma^{32P}]$ATP and $\alpha^{35S}$-dATP. The $glp$ repressor was purified to homogeneity as previously described (58).

Construction of recombinant plasmids. Standard molecular cloning methods were used for the construction of recombinant plasmids. The plasmids used or constructed for this study are listed in Tables 3 ($glpD$ region) and 4 ($glpT$ region).

1. Plasmids containing portions of the $glpD$ region: Plasmids pDW65, pDA1104, pSY15, and pBY100 contained various regions of the $glpD$ gene cloned into pBluescript KS$^-$ (pBS; Stratagene) as indicated in Table 3. Plasmids pBY101 ($O_{D2}$ mutated), pBY102 (wild type), pBY103 (box 1 mutated), pBY104 (box 2 mutated), pBY105 ($O_{D2}$, box 1 and box 2 mutated), pBY109 ($O_{D4}$ mutated), and pBY150 ($O_{D1}$ and $O_{D2}$ mutated) contained an intact $glpD$ gene cloned into the vector pACYC177 (121), except for the indicated point mutations introduced by PCR primers (Table 1). During the construction of these plasmids, replacing the BamHI-BssHII region of pSH56 (37) with each PCR product deleted about 1,500 bp of the $glpEG$ region of pSH56 and left $glpD$ intact. The product of PCR using primers 85488 and 168198 with pSH56 as the template was used to construct pBY102 which carries the wild-type $glpD$ gene. The construction of pBY101
<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence*</th>
<th>Position^b</th>
</tr>
</thead>
<tbody>
<tr>
<td>85488</td>
<td>CGTTAATACATTCAGAAGATCCATGCGTCTC</td>
<td>-10 to 22</td>
</tr>
<tr>
<td>168198</td>
<td>GGCCTTAACCAGCCGGCCGCCTTGCCAGGC</td>
<td>820 to 791</td>
</tr>
<tr>
<td>104322</td>
<td>AACGAACATTTAAGAGTTATAAgcAAAGTGAAATG</td>
<td>167 to 196</td>
</tr>
<tr>
<td>85436</td>
<td>GGTCTGTTTATGTAcACATCTGGGTAACGC</td>
<td>514 to 546</td>
</tr>
<tr>
<td>Bing2</td>
<td>GCCTTTACCAGATGgTAACATACAACACAGACC</td>
<td>546 to 514</td>
</tr>
<tr>
<td>85423</td>
<td>CGAA7ATTCTGACTGcTTGAGGACGACGCCC</td>
<td>621 to 652</td>
</tr>
<tr>
<td>Bing3</td>
<td>GGCGTGCTGTCaACCCAGCAgCTCAGAAAATATTGC</td>
<td>652 to 621</td>
</tr>
<tr>
<td>Bing1^c</td>
<td>GAACCGCCGAGGATCCATGCGTCTCTC</td>
<td>9 to 22</td>
</tr>
<tr>
<td>207153</td>
<td>CCGGAAATTAAaCGgGATTgTGAATATTCTGAC</td>
<td>601 to 633</td>
</tr>
<tr>
<td>181347</td>
<td>GTCAAGAATATTcAAATCCcCGTTAATTTCCGG</td>
<td>633 to 601</td>
</tr>
<tr>
<td>241326</td>
<td>CTTTGCTAAATAgGTgAgATAAAgcAAAAATTTAACAG</td>
<td>144 to 178</td>
</tr>
<tr>
<td>241183</td>
<td>TAAATGTTgcTTATgACATATTGcAAAGAATCG</td>
<td>174 to 139</td>
</tr>
<tr>
<td>207055</td>
<td>CGAGTAGACACTTCCGGCCTTTACGCAC</td>
<td>711 to 682</td>
</tr>
<tr>
<td>254069</td>
<td>CATATCACTCTAAAGATCTT7TTTTCAATGT TA</td>
<td>98 to 129</td>
</tr>
</tbody>
</table>

* Lower case letters indicate bases substituted for the creation of mutations or restriction sites.

^b See Fig. 5.

^c The first 8 bases of this oligonucleotide are from the vector(pACYC177).
Fig. 5. Nucleotide sequence of the glpD control region. The transcriptional start site (+1) with the -10 and -35 promoter elements, the CRP site, and the translational initiation region with the ribosome binding site are indicated (39). The internal operators (O₉3 and O₉4) protected from DNase I digestion are indicated by dashed lines. Box 1 and box 2 are indicated by double dashed lines. The BamHI site used for cloning purposes was introduced using a T to G substitution at position 8. The restriction sites used for the construction of plasmids (Table 3) are also indicated.
<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>223822</td>
<td>CGAGTTTTCATTGTTATCCCG</td>
<td>4 to 23</td>
</tr>
<tr>
<td>311333</td>
<td>TCGCGGAggTTCAATTGTTATCCC</td>
<td>1 to 23</td>
</tr>
<tr>
<td>311334</td>
<td>ACTGCCGCGTcCAGAATCAAAACCTGCGGC</td>
<td>603 to 574</td>
</tr>
<tr>
<td>311355</td>
<td>GTGGCGTGGTTTAAATACTC</td>
<td>314 to 292</td>
</tr>
<tr>
<td>366852</td>
<td>GAACATTTGGAATTCTTATTTAATAGG</td>
<td>132 to 160</td>
</tr>
<tr>
<td>353988</td>
<td>TCTTCACCTGcaggGCGCTGTATACG</td>
<td>229 to 254</td>
</tr>
<tr>
<td>354048</td>
<td>ATTAACAGCGccggCAGTGAGAGAAACC</td>
<td>253 to 226</td>
</tr>
<tr>
<td>368021</td>
<td>GGCTATGCCGCCGTTACTATcTcGGTGACGTAAGAAGTCTTTGC</td>
<td>395 to 433</td>
</tr>
<tr>
<td>366933</td>
<td>AAAGTTCTTCACCGGAgAGATAGTAGCCGCATAGCCAAAG</td>
<td>430 to 391</td>
</tr>
<tr>
<td>368037</td>
<td>TATGGATTttgATccAAATCCATCATGGG</td>
<td>511 to 537</td>
</tr>
<tr>
<td>353953</td>
<td>TGATGAATTtggcAAATCCATAGCC</td>
<td>533 to 508</td>
</tr>
<tr>
<td>353993</td>
<td>ATCGGATCGCagcAAATCCGCGGTTTACTGCCC</td>
<td>544 to 576</td>
</tr>
<tr>
<td>366926</td>
<td>GGAACACCGGgGGAATTttggcGCGATCCCGATACCG</td>
<td>572 to 540</td>
</tr>
<tr>
<td>198093</td>
<td>AGCGGAACGGCCGGCGCTGT</td>
<td></td>
</tr>
<tr>
<td>109180</td>
<td>GATTGGGAGTTCGGAGGGATTCCTTATGTAATCCCTCGGAAATTC</td>
<td>14 to 33</td>
</tr>
</tbody>
</table>

*a* Lower case letters indicate bases substituted for creation of mutations or restriction sites.

*b* See Fig. 6.

*c* The sequence of this oligonucleotide matches the sequence downstream of the multiple
cloning site in pSP417 and was used to prepare radio-labeled DNA for footprinting.

The first 22 bases underlined in this oligonucleotide are not g/lpAT sequence and were used for other purposes. The mutation indicated with lower case was also used for other purposes. This oligonucleotide was used only for preparing radio-labeled DNA for footprinting.
Fig. 6. Nucleotide sequence of the divergently transcribed glpACB-glptQ control region.

The glpACB operators, the positions of CRP binding and the transcriptional initiation sites (+1) are indicated (40). The internal glpT operators (O1-3) protected from DNase I digestion by GlpR are indicated by the double dashed lined. The individual glpT operators are indicated by the lines above the upper strand. The two stronger hypersensitive regions (site 1 and site 4) are indicated by the darker arrows. The weaker hypersensitive region (site 3) is indicated by two light arrows. The region (site 2) which could not be protected within the internal glpT operators is also indicated by a light arrow. The regions protected from DNase I digestion by binding of IHF are indicated by the dashed lines adjacent to the protected strand of DNA with the lower case letters indicating the sequence resembling the IHF consensus. Primer 223822 is the indicated 20-nucleotide sequence, primer 311333 is the indicated 23-nucleotide sequence, primer 311334 is the indicated 30-nucleotide sequence and primer 311355 is the indicated 23-nucleotide sequence. Primers 311333 and 311334 each have two mismatches (lower cases) to create EcoRI and XbaI sites, respectively.
TABLE 3. Plasmids used or constructed in the study of *glpD*

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Insert</th>
<th>Source</th>
<th>Vector</th>
<th>Part of <em>glpD</em> (position*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pDW1</td>
<td><em>Aan</em>-<em>HinP1II</em></td>
<td>pSY12 (48)</td>
<td>M13mp18</td>
<td>139 to 178</td>
</tr>
<tr>
<td>pDW65</td>
<td><em>HindIII</em>-<em>EcoRI</em></td>
<td>pDW1</td>
<td>pBS</td>
<td>139 to 178</td>
</tr>
<tr>
<td>pDA1104</td>
<td><em>EcoRI</em>-<em>EcoRV</em></td>
<td>pDA1101 (3)</td>
<td>pBS</td>
<td>387 to 768</td>
</tr>
<tr>
<td>pSY15</td>
<td><em>Sau3A</em>-<em>XbaI</em></td>
<td>pSY12 (48)</td>
<td>pBS</td>
<td>170 to 218</td>
</tr>
<tr>
<td>pBY100</td>
<td><em>HinfI</em>-<em>EcoRI</em></td>
<td>pSH79 (48)</td>
<td>pDA1104</td>
<td>140 to 768</td>
</tr>
<tr>
<td>pBY102</td>
<td><em>BamHI</em>-<em>BssHII</em></td>
<td>PCR</td>
<td>pSH56</td>
<td>Entire <em>glpD</em></td>
</tr>
<tr>
<td>pBY101</td>
<td><em>BamHI</em>-<em>BssHII</em></td>
<td>PCR</td>
<td>pSH56</td>
<td>Entire <em>glpD</em></td>
</tr>
<tr>
<td>pBY103</td>
<td><em>BamHI</em>-<em>BssHII</em></td>
<td>PCR</td>
<td>pSH56</td>
<td>Entire <em>glpD</em></td>
</tr>
<tr>
<td>pBY104</td>
<td><em>BamHI</em>-<em>BssHII</em></td>
<td>PCR</td>
<td>pSH56</td>
<td>Entire <em>glpD</em></td>
</tr>
<tr>
<td>pBY105</td>
<td><em>BamHI</em>-<em>BssHII</em></td>
<td>PCR</td>
<td>pSH56</td>
<td>Entire <em>glpD</em></td>
</tr>
<tr>
<td>pBY109</td>
<td><em>BamHI</em>-<em>BssHII</em></td>
<td>PCR</td>
<td>pSH56</td>
<td>Entire <em>glpD</em></td>
</tr>
<tr>
<td>pBY150</td>
<td><em>BamHI</em>-<em>BssHII</em></td>
<td>PCR</td>
<td>pSH56</td>
<td>Entire <em>glpD</em></td>
</tr>
<tr>
<td>pBY112</td>
<td><em>BamHI</em>-<em>EcoRV</em></td>
<td>pBY102</td>
<td>pSP417</td>
<td>11 to 768</td>
</tr>
<tr>
<td>pBY159</td>
<td><em>BamHI</em>-<em>EcoRV</em></td>
<td>pBY109</td>
<td>pSP417</td>
<td>11 to 768</td>
</tr>
<tr>
<td>pBY160</td>
<td><em>BamHI</em>-<em>EcoRV</em></td>
<td>pBY150</td>
<td>pSP417</td>
<td>11 to 768</td>
</tr>
<tr>
<td>pBY162</td>
<td><em>BamHI</em>-<em>EcoRI</em></td>
<td>pBY150</td>
<td>pBY159</td>
<td>11 to 768</td>
</tr>
<tr>
<td>pBY128</td>
<td><em>BamHI</em>-<em>BglII</em></td>
<td>pBY102</td>
<td>pSP417</td>
<td>11 to 218</td>
</tr>
</tbody>
</table>

*See Fig. 5.
which harbors an altered \( \text{O}_{2} \), was accomplished by a previously published method (122). Primers 104322 and 168198 were used for first-round PCR with pSH56 as the template. The product of the first-round PCR was used as a megaprimer together with primer 85488 and with pSH56 as the template in a second-round PCR. Plasmids pBY103, pBY104, pBY105, pBY109, and pBY150 were constructed by a modification of a previously published method (123). Construction of pBY103, pBY104, and pBY150 used pBY101 as the template, the construction of pBY105 used pBY103 as the template, and the construction of pBY109 used pSH56 as the template for PCR. Two first-round PCRs were used to produce two overlapping fragments of DNA by using the primer pairs as follows: primers Bing1 and Bing2 and primers 85436 and 168198 for pBY103; primers Bing1 and Bing3 and primers 85423 and 168198 for pBY104; primers Bing1 and Bing3 and primers 85423 and 168198 for pBY105; and primers 85488 and 181347 and primers 207153 and 168198 for pBY109. After the first-round PCR, the two products were mixed, denatured, and reannealed to be used as the template for the second-round PCR using primers 168198 and 85488 for pBY109 and primers 168198 and Bing1 for pBY103, pBY104, pBY105, and pBY150. The oligonucleotides used for creating operator mutations contained base substitutions only in the wobble positions of affected codons. Therefore, each plasmid was predicted to encode wild-type GlpD, which was verified with strain SH305 (\( \Delta \text{gldD}102 \)) (Table 5). All of the plasmids complemented \( \Delta \text{gldD}102 \). Plasmids pBY112, pBY159, pBY160, and pBY162 contained the same \text{BamHI-EcoRV} region of \text{gldD} (except the point mutations from the parental
plasmid (Table 2) cloned into vector pSP417 (124), which created \( glpD-lacZ \) transcriptional fusions. Plasmid pBY128 contained a shorter region of \( glpD \) (\( \text{BamH}1-Bg\text{II} \)) cloned into vector pSP417. The expected sequence of the inserted DNA in each plasmid was verified by DNA sequence analysis (125).

2. Plasmids containing positions of the \( glpT \) region: Various parts of the \( glpAT \) control region were generated either by PCR or by restriction enzyme digestion of pATHN (40), and then cloned into pSP417 (124) as described in Table 4. A 302 bp \( \text{EcoR}1-\text{XbaI} \) fragment containing the 268 bp \( \text{HaeIII-Nr}u1 \) (Fig. 6) fragment with the \( glpAT \) control region was cut from pATHN (40) and cloned into the vector pSP417, resulting in the transcriptional fusion \( glpT''-lacZ \) (pBY20). This plasmid contains the \( glpAT \) control region without the putative \( glpT \) internal operators.

Primers 366852 and 311334 were used to amplify a 472 bp DNA fragment using pGS31H5 (31) as template and a 450 bp \( \text{EcoR}1-\text{XbaI} \) DNA fragment was cleaved and cloned into pSP417, resulting in the transcriptional fusion \( glpT''-lacZ \) (pBY70), which contains only the \( glpT \) promoter and internal \( glpT \) operators without the \( glpA \) operators. In order to construct a \( glpT-lacZ \) fusion containing the \( glpA \) operators and the internal \( glpT \) operators, PCR primers 311333 and 311334 (see Fig. 6) were used for amplification of a 604 bp DNA fragment using pGS31H5 as template. A 587 bp DNA fragment was cleaved from the PCR product using \( \text{EcoR}1 \) and \( \text{XbaI} \) and cloned into pSP417 as described above, resulting in the transcriptional fusion \( glpT''-lacZ \) (pBY30).

Construction of pBY75 harboring altered \( O_{73} \) was accomplished using a
TABLE 4. Plasmids used or constructed in the study of $glpT^a$

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Relevant genotype</th>
<th>Part of $glpT$ contained$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBY20</td>
<td>$\Phi(glpT^-\cdot-lacZ)$</td>
<td>2 to 272</td>
</tr>
<tr>
<td>pBY70</td>
<td>$\Phi(glpT^-\cdot-lacZ)$ PCR</td>
<td>142 to 592</td>
</tr>
<tr>
<td>pBY30</td>
<td>$\Phi(glpT^-\cdot-lacZ)$</td>
<td>10 to 592</td>
</tr>
<tr>
<td>pBY71</td>
<td>$\Phi(glpT_{mot}^-\cdot-lacZ)$</td>
<td>10 to 592</td>
</tr>
<tr>
<td>pBY72</td>
<td>$\Phi(glpT_{ihr2}^-\cdot-lacZ)$</td>
<td>10 to 592</td>
</tr>
<tr>
<td>pBY73</td>
<td>$\Phi(glpT_{aff14AINF2}^-\cdot-lacZ)$</td>
<td>10 to 592</td>
</tr>
<tr>
<td>pBY74</td>
<td>$\Phi(glpT_{mot}^-\cdot-lacZ)$</td>
<td>10 to 592</td>
</tr>
<tr>
<td>pBY75</td>
<td>$\Phi(glpT_{mot}^-\cdot-lacZ)$</td>
<td>10 to 592</td>
</tr>
</tbody>
</table>

$^a$Each plasmid was constructed by cloning of inserts into the EcoRI and XbaI sites of the vector pSP417. The source of the insert for pBY20 was pATHN (40), and the source of the insert for all other plasmids was PCR as described in the text. The sequences present in pBY30, pBY71, pBY72, pBY73, pBY74 and pBY75 are the same except that $MIHF1$ designates the five substitutions in site IHF1, $MIHF2$ designates the four substitutions in site IHF2, $MOT1$ designates the three substitutions in O₁, and $MOT3$ designates the four substitutions in O₂. The $glpT^-\cdot-lacZ$ fusion of pBY70 contains a shorter $glpT$ sequence with only the internal $glpT$ operators. The $glpT^-\cdot-lacZ$ fusion of pBY20 contains the $glpA$ operators and $glpT$ sequences downstream to the site IHF1.

$^b$The position $-1$ of the sequence indicated in this table is the first nucleotide of Fig. 6.

38
published method (122). Primers 311333 and 366926 were used for first round PCR with pGS31H5 as the template. The product of the first round PCR was used as a megaprimer together with primer 311334 and pGS31H5 as template in a second round PCR. A 587 bp DNA fragment was cut from the final PCR product using EcoRI and XbaI and cloned into pSP417 as described above. Plasmids pBY71, pBY72 and pBY74 were constructed using a modification of a published method (123) and pGS31H5 as the template. Two first round PCR reactions were used to produce two overlapping fragments of DNA, using primer pairs as follows: primers 311333/354048 and primers 311334/353988 for pBY71; primers 311333/366933 and primers 311334/368021 for pBY72; primers 311333/353953 and primers 311334/368037 for pBY74. After the first round PCR, the two products were mixed, denatured and reannealed to be used as the template for the second round PCR using primers 311333 and 311334. A 587 bp DNA fragment was cut from the final PCR product using EcoRI and XbaI and cloned into pSP417 as described above, with the resulting plasmids each containing a specific altered sequence (pBY71 with IHF1 site mutated, pBY72 with IHF2 site mutated and pBY74 with Or1 mutated).

Construction of pBY73 required a three-piece ligation method. A 327 bp EcoRI-SacII DNA fragment cut from the second round PCR used for making pBY71 and a 260 bp SacII-XbaI DNA fragment cut from the second round PCR used for making pBY72 were ligated together with the pSP417 vector cut with EcoRI-XbaI. The resulting plasmid (pBY73) contained mutations in both the IHF1 and IHF2 sites. The oligonucleotides used for creating mutations in the operators or IHF sites within the glpT
coding region contained base substitutions in the wobble positions of affected codons. Therefore, these mutations did not change the amino acid sequences. The base pairs chosen for alteration of \( O_1 \) or \( O_3 \) were those shown to be crucial for an operator to bind GlpR (53). The base pairs mutated in IHF1 and IHF2 were chosen because they belong to the 13 bp IHF consensus sequence. The sequence of the insert in each plasmid constructed was verified by DNA sequence analysis (125).

**Growth media.** LB medium (126) with appropriate antibiotics (ampicillin, 100 \( \mu g/ml \); tetracycline, 10 \( \mu g/ml \); kanamycin, 50 \( \mu g/ml \); and chloramphenicol, 50 \( \mu g/ml \)) was used for the growth of *E. coli* strains. Transformation of competent cells was carried out as previously described (127), with selection on LB plates containing ampicillin (100 \( \mu g/ml \)) and 40 \( \mu g \) of 5-bromo-4-chloro-3-indolyl-\( \beta \)-D-galactopyranoside (X-Gal) per ml where appropriate. LB medium containing 20 mM sodium citrate and appropriate antibiotics was used for the selection and scoring of transductants. For minimal medium, the A and B salts of Clark and Maaløe (128) were supplemented with 2 \( \mu g \) of thiamine per ml, 0.2% Casamino Acids, 0.2% maltose, and the appropriate antibiotics at one-half strength.

**Bacterial strains.** The bacterial strains used or constructed are listed in Table 5. Single \( \lambda \) lysogens containing *glpD-lacZ* transcriptional fusions were constructed as follows. The various *glpD* control regions were cloned upstream of *lacZ* in pSP417 as described above. Promoter-probe vector pSP417 (124), a derivative of pRS415 (129), contains a promoterless *lacZ* gene downstream of the multiple cloning site, four
TABLE 5. Strains of *E. coli* K-12 used or constructed in this study

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source or derivation</th>
</tr>
</thead>
</table>
| IH1-6  | ΔhimA81 zdi-3123::Tn10kan lysA rpsL  
  polA1 ilvP11P2::galK | (130) |
| A5427  | N99 galK hupAI6::kan hupB11::cat  
  Δhim482::Tn10 | (131) |
| DH5αF' | F φ80lacZΔM15 recA1 endA1 gyrA96 thi-1  
  hsdRI7 supE44 relA1 Δ(argF-lacZ)U169 | (53) |
| MC4100 | F araD139 Δ(argF-lacZ)U169 rpsL130 deoC1  
  relA1 rbsR ptsF25 flbB5301 | (132) |
<p>| TS100  | MC4100 glpR2 | (7) |
| SH305  | MC4100 ΔglpD102 recA1 srl::Tn10 | (14) |
| TST3   | MC4100 malT::Tn10 | (127) |
| WO331  | MC4100 glpR12 (glpR*) recA1 srl::Tn10 | (53) |
| DL291  | TS100 gyrA Δ(glpT-glPA)593 recA1 | (31) |
| BY4000* | MC4100 Φ(glpD-lacZ) λBY112 lysogen of MC4100 |
| BY1000* | TS100 Φ(glpD-lacZ) λBY112 lysogen of TS100 |
| BY4159* | MC4100 Φ(glpD-lacZ) λBY159 lysogen of MC4100 |</p>
<table>
<thead>
<tr>
<th>Strain</th>
<th>Description</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>BY1159</td>
<td>TS100 $\Phi (glpD_{mt}^{-}-lacZ)$</td>
<td>$\lambda BY159$ lysogen of TS100</td>
</tr>
<tr>
<td>BY4128</td>
<td>MC4100 $\Phi (glpD^{-}-lacZ)$</td>
<td>$\lambda BY128$ lysogen of MC4100</td>
</tr>
<tr>
<td>BY1128</td>
<td>TS100 $\Phi (glpD^{-}-lacZ)$</td>
<td>$\lambda BY128$ lysogen of TS100</td>
</tr>
<tr>
<td>BY4160</td>
<td>MC4100 $\Phi (glpD_{mt}^{-}-lacZ)$</td>
<td>$\lambda BY160$ lysogen of MC4100</td>
</tr>
<tr>
<td>BY1160</td>
<td>TS100 $\Phi (glpD_{mt}^{-}-lacZ)$</td>
<td>$\lambda BY160$ lysogen of TS100</td>
</tr>
<tr>
<td>BY4162</td>
<td>MC4100 $\Phi (glpD_{mt}^{-2}-lacZ)$</td>
<td>$\lambda BY162$ lysogen of MC4100</td>
</tr>
<tr>
<td>BY1162</td>
<td>TS100 $\Phi (glpD_{mt}^{-2}-lacZ)$</td>
<td>$\lambda BY162$ lysogen of TS100</td>
</tr>
<tr>
<td>BY4006</td>
<td>BY4000 $hupA16::kan hupB11::cat$</td>
<td>P1(A5427) – BY4000</td>
</tr>
<tr>
<td>BY1006</td>
<td>BY1000 $hupA16::kan hupB11::cat$</td>
<td>P1(A5427) – BY1000</td>
</tr>
<tr>
<td>BY4459</td>
<td>BY4159 $hupA16::kan hupB11::cat$</td>
<td>P1(A5427) – BY4159</td>
</tr>
<tr>
<td>BY1459</td>
<td>BY1159 $hupA16::kan hupB11::cat$</td>
<td>P1(A5427) – BY1159</td>
</tr>
<tr>
<td>BY4004</td>
<td>BY4000 $\Delta himA82$:Tn10</td>
<td>P1(A5427) – BY4000(pUHS2)</td>
</tr>
<tr>
<td>BY1004</td>
<td>BY1000 $\Delta himA82$:Tn10</td>
<td>P1(A5427) – BY1000(pUHS2)</td>
</tr>
<tr>
<td>BY4259</td>
<td>BY4159 $\Delta himA81$ zdi-3123::Tn10kan</td>
<td>P1(IH1-6) – BY4159(pSC101)</td>
</tr>
<tr>
<td>BY1259</td>
<td>BY1159 $\Delta himA81$ zdi-3123::Tn10kan</td>
<td>P1(IH1-6) – BY1159(pSC101)</td>
</tr>
<tr>
<td>BY21*</td>
<td>MC4100 $\Phi (glpT^{-}-lacZ)$</td>
<td>$\lambda BY20$ lysogen of MC4100</td>
</tr>
<tr>
<td>BY22*</td>
<td>TS100 $\Phi (glpT^{-}-lacZ)$</td>
<td>$\lambda BY20$ lysogen of TS100</td>
</tr>
<tr>
<td>BY31*</td>
<td>MC4100 $\Phi (glpT^{-}-lacZ)$</td>
<td>$\lambda BY30$ lysogen of MC4100</td>
</tr>
<tr>
<td>BY32*</td>
<td>TS100 $\Phi (glpT^{-}-lacZ)$</td>
<td>$\lambda BY30$ lysogen of TS100</td>
</tr>
<tr>
<td>BY701*</td>
<td>MC4100 $\Phi (glpT^{-}-lacZ)$</td>
<td>$\lambda BY70$ lysogen of MC4100</td>
</tr>
<tr>
<td>BY702*</td>
<td>TS100 $\Phi (glpT^{-}-lacZ)$</td>
<td>$\lambda BY70$ lysogen of TS100</td>
</tr>
<tr>
<td>Strain</td>
<td>Description</td>
<td>Origin</td>
</tr>
<tr>
<td>----------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>BY711</td>
<td>MC4100 $\Phi(glpT_{\text{MHF1}}^{-}-\text{lacZ})$</td>
<td></td>
</tr>
<tr>
<td>BY712</td>
<td>TS100 $\Phi(glpT_{\text{MHF1}}^{-}-\text{lacZ})$</td>
<td></td>
</tr>
<tr>
<td>BY721</td>
<td>MC4100 $\Phi(glpT_{\text{MHF2}}^{-}-\text{lacZ})$</td>
<td></td>
</tr>
<tr>
<td>BY722</td>
<td>TS100 $\Phi(glpT_{\text{MHF2}}^{-}-\text{lacZ})$</td>
<td></td>
</tr>
<tr>
<td>BY731</td>
<td>MC4100 $\Phi(glpT_{\text{MHF1,2}}^{-}-\text{lacZ})$</td>
<td></td>
</tr>
<tr>
<td>BY732</td>
<td>TS100 $\Phi(glpT_{\text{MHF1,2}}^{-}-\text{lacZ})$</td>
<td></td>
</tr>
<tr>
<td>BY741</td>
<td>MC4100 $\Phi(glpT_{\text{MORT}}^{-}-\text{lacZ})$</td>
<td></td>
</tr>
<tr>
<td>BY742</td>
<td>TS100 $\Phi(glpT_{\text{MORT}}^{-}-\text{lacZ})$</td>
<td></td>
</tr>
<tr>
<td>BY751</td>
<td>MC4100 $\Phi(glpT_{\text{MORT}}^{-}-\text{lacZ})$</td>
<td></td>
</tr>
<tr>
<td>BY752</td>
<td>TS100 $\Phi(glpT_{\text{MORT}}^{-}-\text{lacZ})$</td>
<td></td>
</tr>
<tr>
<td>BY27</td>
<td>BY21 $\Delta himA82\text{:}Tn10$</td>
<td>P1(A5427) - BY21(pUHS2)</td>
</tr>
<tr>
<td>BY28</td>
<td>BY22 $\Delta himA82\text{:}Tn10$</td>
<td>P1(A5427) - BY22(pUHS2)</td>
</tr>
<tr>
<td>BY37</td>
<td>BY31 $\Delta himA82\text{:}Tn10$</td>
<td>P1(A5427) - BY31(pUHS2)</td>
</tr>
<tr>
<td>BY38</td>
<td>BY32 $\Delta himA82\text{:}Tn10$</td>
<td>P1(A5427) - BY32(pUHS2)</td>
</tr>
<tr>
<td>BY7410</td>
<td>BY741 $\Delta himA82\text{:}Tn10$</td>
<td>P1(A5427) - BY741(pUHS2)</td>
</tr>
<tr>
<td>BY7420</td>
<td>BY742 $\Delta himA82\text{:}Tn10$</td>
<td>P1(A5427) - BY742(pUHS2)</td>
</tr>
<tr>
<td>BY7510</td>
<td>BY751 $\Delta himA82\text{:}Tn10$</td>
<td>P1(A5427) - BY751(pUHS2)</td>
</tr>
<tr>
<td>BY7520</td>
<td>BY752 $\Delta himA82\text{:}Tn10$</td>
<td>P1(A5427) - BY752(pUHS2)</td>
</tr>
<tr>
<td>BY25</td>
<td>BY21 $\text{hupA16::kan hupB11::cat}$</td>
<td>P1(A5427) - BY21</td>
</tr>
<tr>
<td>BY26</td>
<td>BY22 $\text{hupA16::kan hupB11::cat}$</td>
<td>P1(A5427) - BY22</td>
</tr>
<tr>
<td>BY35</td>
<td>BY31 $\text{hupA16::kan hupB11::cat}$</td>
<td>P1(A5427) - BY31</td>
</tr>
</tbody>
</table>
a Lysogens BY4000, BY1000, BY4159, BY1159, BY4128, BY1128, BY4160, BY1160, BY4162 and BY1162 have the same glpD-lacZ fusions as the parental plasmids pBY112, pBY112, pBY159, pBY159, pBY128, pBY128, pBY160, pBY160, pBY160, and pBY162, respectively.

b Single lysogens of each lysogenic strain were identified by assays of the β-galactosidase activities in at least eight independently isolated lysogens. A strain was considered to be singly lysogenic for the indicated phage if its lacZ expression is the lowest, and double or triple lacZ expression was found in other independently isolated lysogens.
transcriptional terminators upstream of the multiple cloning site, and a divergently transcribed \textit{bla} gene. Each plasmid-borne \textit{glpD-lacZ} fusion was transferred to phage \textit{\lambda RS45} (containing a segment of the \textit{bla} gene and divergent promoterless \textit{lacZ} gene, oriented as in pSP417) by homologous recombination (127,129). Strain SH305 (\textit{\Delta glpD102 recA}) was used as the host (avoiding recombination between the mutated plasmid and the chromosome) for selection and purification of the resulting phages (\textit{\lambda BY128, \lambda BY112, \lambda BY159, \lambda BY160, and \lambda BY162}), which contain the same fusions as pBY128, pBY112, pBY159, pBY160, and pBY162, respectively. Single lysogens of MC4100 \textit{(glpR\textsuperscript{-})} and TS100 \textit{(glpR2)} with each phage are listed in Table 5.

Single \textit{\lambda} lysogens containing \textit{glpT-lacZ} transcriptional fusions were constructed in a similar way. The various \textit{glpAT} control regions were cloned upstream of \textit{lacZ} in pSP417 as described above. Each plasmid-borne \textit{glpT-lacZ} fusion was transferred to phage \textit{\lambda RS45} by homologous recombination. Strain DL291 (\textit{\Delta glpT1 recA}) was used as the host (to avoid recombination between the mutated plasmids and the chromosome) for selection and purification of the resulting phages (\textit{\lambda BY20, \lambda BY30, \lambda BY70, \lambda BY71, \lambda BY72, \lambda BY73, \lambda BY74, and \lambda BY75}) which contain the same fusions as pBY20, pBY30, pBY71, pBY72, pBY73, pBY74, and pBY75, respectively. Single lysogens of MC4100 \textit{(glpR\textsuperscript{-})} or TS100 \textit{(glpR2)} using each phage are listed in Table 5.

The \textit{himA} deletion strains were constructed as follows. Plasmid pSC101 was introduced into strains BY4159 and BY1159 by selection for \textit{Tet\textsuperscript{r}}, and the resulting transformants were used as recipients in P1 crosses with strain IH1-6 (\textit{\Delta himA81:kan})
as the donor and with Kan' selection. Transductants containing the himA deletion were identified by their Tet' phenotype (replication of plasmids with the pSC101 origin requires IHF protein (130,133)). The resulting strains were named BY4259 and BY1259, respectively (Table 5). The ΔhimA derivatives of strains BY4000, BY1000, BY21, BY22, BY31, BY32, BY741, BY742, BY751 and BY752 were constructed by the same method except that plasmid pUHS2 (a Kan' pSC101 derivative, R. Lutz and H. Bujard, University of Heidelberg) was used for transformation, and the resulting transformants selected by Kan' were used as recipients in P1 crosses with strain A5427 (ΔhimA82::Tn10 (131)) as the donor and with Tet' selection followed by screening to identify Kan' transductants. The resulting strains were named BY4004, BY1004, BY27, BY28, BY37, BY38, BY7410, BY7420, BY7510 and BY7520, respectively (Table 5).

To construct strains lacking HU, strains BY4000, BY1000, BY4159, BY1159, BY21, BY22, BY31 and BY32 were used as recipients in two consecutive P1 crosses with strain A5427 (hupA16::kan hupB11::cat (131)) as the donor and with Kan' selection followed by Cam' selection. The resulting strains were named BY4006, BY1006, BY4459, BY1459, BY25, BY26, BY35 and BY36, respectively (Table 5).

DNase I footprinting. The conditions used for GlpR footprinting were modifications of previously described methods (17,39). The end-labeled DNA substrates were prepared by PCR using appropriate templates, one unlabeled primer, and one 32P-end-labeled primer as indicated. Purified repressor was added to the radiolabeled DNA in a reaction mixture containing 10 mM Tris-HCl (pH 7.4), 1 mM
EDTA, 50 mM KCl, 2 mM CaCl₂, 4 mM MgCl₂, 5% glycerol, 0.025% Triton X-100, 100 μg bovine serum albumin per ml, and 5 μg sonicated salmon sperm DNA per ml in a total volume of 0.1 ml. The reaction mixtures were incubated for 30 min at 37°C, with the subsequent addition of 0.5 ng of DNase I (RNase free, from Boehringer Mannheim). Incubation was continued for 3.5 min. The DNA was precipitated by the addition of 5 μg of sonicated salmon sperm DNA, 75 μl of 6.5 M ammonium acetate plus 34 mM EDTA, and 0.4 ml of 95% ethanol. The mixture was placed at minus 70°C for 20 min, centrifuged at 4°C, and rinsed with 80% ethanol. The samples were dried, dissolved in 4 μl of 10 mM EDTA in 95% formamide, and analyzed on a sequencing gel.

For IHF footprinting, purified IHF was added to the end-labeled DNA as indicated in the figure legends in reaction mixtures the same as those used for GlpR footprinting except that salmon sperm DNA was omitted. After incubation at 22°C for 30 min, 2.5 ng DNase I was added. The reaction with no IHF was continued for 1 min and reactions with IHF were continued for 5 min. The samples were processed and analyzed on a sequencing gel.

**Assay of β-galactosidase.** Cells were grown with vigorous aeration in minimal medium containing 0.2% maltose and 0.2% Casamino Acids as carbon sources. Enzyme activities were determined by using at least four different logarithmically growing cultures on at least two different occasions and were expressed in Miller units (126). Because it is possible that cells with differences in the *hup, himA*, and *glpR* alleles differ in sizes, thereby introducing error into the calculation of β-galactosidase specific activity.
when culture optical density is used as the measure of protein content, the specific activities were also determined relative to protein concentrations measured (Bradford method (134)) in extracts of sonicated cells. Repression ratios were found to be essentially the same as those calculated using Miller units.
RESULTS

1. Regulation of glpD gene expression.

Apparent binding affinity of the glp repressor to various segments of glpD.

Computer analysis revealed four potential glp operators within the glpD coding region, box 1, box 2, O₁₃, and O₁₄ (Fig. 5). To find out if any of them bind GlpR in vivo, plasmids containing different combinations of glpD operator sequences were constructed (Table 3). These plasmids were introduced into strain WO331 (Table 5) (53), a strain that contains a glpD-lacZ fusion and a noninducible repressor gene, glpR⁺. Binding of GlpR⁺ to operators carried by multicopy plasmids releases repression of the glpD-lacZ fusion (53). Therefore, the level of expression of lacZ is proportional to the binding strength of the operators carried by plasmids introduced into this strain (53). From the results shown in Table 6, pSY15, which harbors the tandem operators O₁₁ and O₁₂, bound the repressor 10 times better than did pDW65, which harbors only O₁₁ (2,210 versus 190 Miller units); this is consistent with the previous discovery that GlpR binds tandem operators cooperatively (53). Plasmid pDA1104 contains only the potential internal glpD operator sequences and had almost no apparent binding affinity for GlpR⁺, because the level of β-galactosidase expression was almost the same as that of WO331 which contains the vector pBS (80 versus 50). Plasmid pBY100 contains the potential internal glpD operator sequences along with O₁₁ and O₁₂. Apparently, pBY100 bound GlpR twice as strongly as did pSY15, which contains only O₁₁ and O₁₂ (3,800 versus 2,210). The effect seen with pBY100 was much higher than the sum of pDA1104 and
TABLE 6. Titration of chromosomally-encoded GlpR° by different combinations of multicopy glpD operators

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>glpD operators present</th>
<th>β-Galactosidase sp act†</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBS</td>
<td>None</td>
<td>50±3</td>
</tr>
<tr>
<td>pDW65b</td>
<td>O,1</td>
<td>190±9</td>
</tr>
<tr>
<td>pSY15p</td>
<td>O,1,O,2</td>
<td>2,210±20</td>
</tr>
<tr>
<td>pDA1104°</td>
<td>O,3,O,4</td>
<td>80±7</td>
</tr>
<tr>
<td>pBY100b</td>
<td>O,1,O,2,O,3,O,4</td>
<td>3,800±140</td>
</tr>
<tr>
<td>pACYC177</td>
<td>None</td>
<td>50±1</td>
</tr>
<tr>
<td>pBY102c</td>
<td>O,1,O,2,O,3,O,4</td>
<td>380±6</td>
</tr>
<tr>
<td>pBY101°</td>
<td>O,2 mutated</td>
<td>200±3</td>
</tr>
<tr>
<td>pBY103c</td>
<td>O,2 and box 1 mutated</td>
<td>200±8</td>
</tr>
<tr>
<td>pBY104°</td>
<td>O,2 and box 2 mutated</td>
<td>250±5</td>
</tr>
<tr>
<td>pBY105c</td>
<td>O,2, box 1 and box 2 mutated</td>
<td>200±12</td>
</tr>
<tr>
<td>pBY106c</td>
<td>O,2, box 1, box 2 and O,4 mutated</td>
<td>90±4</td>
</tr>
<tr>
<td>pBY109c</td>
<td>O,4 mutated</td>
<td>100±1</td>
</tr>
</tbody>
</table>

†Data are the means ± standard deviations derived from at least four independently grown cultures of WO331 containing the indicated plasmids.

bPlasmid was constructed by using vector pBS.

cPlasmid was constructed by using vector pACYC177 and contained the same glpD sequence as pBY102 except for the indicated mutation.
pSY15. This synergistic effect may be due to the formation of a DNA loop, similar to that for repression of the lac system in which the effect of $O_1$ and $O_2$ together was much higher than the additive effect of the individual operators (71).

In order to find out which potential operators function in vivo, several plasmids were constructed in which each potential operator was mutated on the basis of the previous finding that substitutions at positions 3, 4, 5, and 8 from the center of symmetry of the palindromic operator caused severe decreases in GlpR binding (53). These plasmids had the same glpD sequences except for multiple substitutions predicted to severely impair operator function (Tables 1 and 3). The assay system used was the same as that described above with strain WO331, except that the vector used in this case was a lower-copy-number plasmid (pACYC177). Therefore, the titration effect on GlpR$^+$ is much less than that obtained with the high-copy-number vector pBS. From the results shown in Table 6, mutations in $O_{2}2$ caused a twofold decrease in apparent GlpR$^+$ binding affinity (380 for pBY102 versus 200 for pBY101). Further mutation of box 1 and/or box 2 did not have significant effects on GlpR$^+$ binding (compare the expression in WO331 harboring pBY103, pBY104, or pBY105 with that in WO331 harboring pBY101), indicating that these two sites probably are not functional repressor binding sites.

However, mutation of $O_{2}4$ (pBY109) caused a fourfold decrease in the binding of GlpR$^+$ compared with that of the wild-type plasmid pBY102 (100 versus 380). The effect of mutating $O_{2}4$ was stronger than was the effect of mutating the tandem operator $O_{2}2$ (pBY101), suggesting that cooperative interaction between $O_{2}1$ and $O_{2}4$ is equivalent or
better than interaction between \(O_{D1}\) and \(O_{D2}\), as determined by using this plasmid system. The fact that synergistic interactions between \(O_{D1}\) and \(O_{D2}\) were not apparent using the pACYC177 derivatives suggests that the concentrations of these plasmids may be too low to effectively titrate GlpR\(^n\).

**Role of the internal glpD operators in repression.** As mentioned above, the potential glpD internal operators bound GlpR\(^n\) only in the presence of the tandem glpD operators \(O_{D1}\) and \(O_{D2}\). To determine the effects of the potential operators on glpD transcription in vivo, lysogens harboring glpD-lacZ fusions with various combinations of operators were constructed as described in Materials and Methods (Table 5). BY4000, BY1000, BY4159, and BY1159 have the same glpD-lacZ sequences, including the internal glpD region, except that both BY4159 and BY1159 carry the \(O_{D4}\) mutation. As the control, BY4128 and BY1128 have a shorter glpD sequence, with only the tandem operators (\(O_{D1}\) and \(O_{D2}\)). From the results shown in Table 7, mutations in \(O_{D4}\) lowered the repression from 56-fold (expression of BY1000 divided by that of BY4000) to 8-fold (expression of BY1159 divided by that of BY4159), a repression ratio similar to that of strains with only the tandem operators (BY4128 and BY1128, repression ratio of 11). This suggests that the internal operator(s) contributes five- to sevenfold to repression (56 compared with 8 or 11) in vivo. The absolute level of lacZ expression in glpR2 strains was not the same with or without the mutations for unknown reasons. A possible reason is that the mRNA transcribed from the mutated DNA is less stable compared with that from the wild type DNA and results in lower lacZ expression.
TABLE 7. Effect of internal glpD operators on regulation of glpD-lacZ transcriptional fusions

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant genotype</th>
<th>β-Galactosidase</th>
<th>Repression $^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>sp act</td>
<td></td>
</tr>
<tr>
<td>BY4000</td>
<td>glpR $^c$ Φ(glpD-lacZ)</td>
<td>10±5</td>
<td></td>
</tr>
<tr>
<td>BY1000</td>
<td>glpR2 Φ(glpD-lacZ)</td>
<td>6,200±150</td>
<td>56</td>
</tr>
<tr>
<td>BY4159</td>
<td>glpR $^c$ Φ(glpD$_{mt.2}$-lacZ)</td>
<td>130±5</td>
<td></td>
</tr>
<tr>
<td>BY1159</td>
<td>glpR2 Φ(glpD$_{mt.2}$-lacZ)</td>
<td>1,000±50</td>
<td>8</td>
</tr>
<tr>
<td>BY4128</td>
<td>glpR $^c$ Φ(glpD$^{''}$-lacZ)</td>
<td>190±14</td>
<td></td>
</tr>
<tr>
<td>BY1128</td>
<td>glpR2 Φ(glpD$^{''}$-lacZ)</td>
<td>2070±6</td>
<td>11</td>
</tr>
<tr>
<td>BY4160</td>
<td>glpR $^c$ Φ(glpD$_{mt.2}$,'-lacZ)</td>
<td>150±21</td>
<td></td>
</tr>
<tr>
<td>BY1160</td>
<td>glpR2 Φ(glpD$_{mt.2}$,'-lacZ)</td>
<td>150±15</td>
<td>1</td>
</tr>
<tr>
<td>BY4162</td>
<td>glpR $^c$ Φ(glpD$_{mt.2}$,'-lacZ)</td>
<td>132±2</td>
<td></td>
</tr>
<tr>
<td>BY1162</td>
<td>glpR2 Φ(glpD$_{mt.2}$,'-lacZ)</td>
<td>100±11</td>
<td>1</td>
</tr>
</tbody>
</table>

$^a$ The sequences present in (glpD-lacZ), (glpD$_{mt.2}$,-lacZ) and (glpD$_{mt.2}$,-lacZ) are the same except that glpD$_{mt.2}$ contains 3 substitutions in O$_{2}$A and glpD$_{mt.2}$ contains 11 substitutions in O$_{5}$A and O$_{5}$B. The glpD$^{''}$-lacZ fusion contains a shorter glpD sequence with only O$_{5}$A and O$_{5}$B.

$^b$ Repression was defined as the specific activity in the glpR2 strain divided by that in the glpR$^c$ strain.
Tandem operators \(O_{01}\) and \(O_{02}\) dominate transcriptional repression.

Plasmid-borne internal \(glpD\) operators \(O_{03}\) and \(O_{04}\) have no apparent affinity for GlpR\(^n\) in the absence of \(O_{01}\) and \(O_{02}\), as described above. In order to find out whether the in vivo effect of internal operators on negative regulation depends on the tandem \(glpD\) operators, lysogens BY4160 and BY1160 were constructed. These strains contain the same \(glpD\)-\(lacZ\) sequences except that the \(O_{01}\) and \(O_{02}\) operators were inactivated by specific mutations. The substitutions in \(O_{01}\) affected the minus 10 sequence of the promoter and resulted in decreased promoter activity. We found that it was necessary to decrease the level of \(glpD\) expression because constitutive expression of \(glpD\) from multicopy plasmids was toxic, resulting in accumulation of random mutations in the \(glpD\) gene (data not shown). In addition to \(O_{01}\) and \(O_{02}\) mutations, strains BY4162 and BY1162 have \(O_{04}\) mutations. From the results shown in Table 7, when \(O_{01}\) and \(O_{02}\) were inactivated, there was no repression with or without the internal operators. Therefore, it is apparent that \(O_{01}\) and \(O_{02}\) are needed for repression and that the \(glpD\) internal operators are ineffective without the tandem \(glpD\) operators.

**DNase I footprinting of the \(glpD\) operators.** The exact positions of the \(glpD\) internal operators were identified by performing a series of DNase I footprinting experiments. End-labeled DNA containing tandem operators \(O_{01}\) and \(O_{02}\) was used as a positive control for GlpR function (48). The results are shown in Fig. 7A. Repressor completely protected the tandem operators \(O_{01}\) and \(O_{02}\) at 1 nM GlpR (Fig. 7A, lanes 1 through 5), showing that GlpR is active and binds with high affinity to these tandem
Fig. 7. Identification of *glpD* operators by DNase I footprinting.

A. O$_D$1 and O$_D$2. DNA for footprinting was labeled by PCR amplification of pBY100
with primer 207055 and $^{32}$P- end-labeled SK primer. Reactions were carried out with the
following concentrations of GlpR tetramers and analyzed in the indicated lanes: 1, no
repressor; 2, 1 nM; 3, 2 nM; 4, 3 nM; 5, 4 nM; 6, 4 nM GlpR and 2 mM glycerol-P.

B. O$_D$3 and O$_D$4 (in the presence of O$_D$1 and O$_D$2). DNA for footprinting was generated
by PCR amplification of pBY100 with primer SK and $^{32}$P- end-labeled primer 207055.
Reactions were carried out with the following concentrations of *glp* repressor tetramers
and analyzed in the indicated lanes: 1, no repressor; 2, 2 nM; 3, 4 nM; 4, 8 nM; 5, 20 nM;
6, 40 nM; 7, 8 nM GlpR and 3 mM glycerol-P.

The size markers (A, C, G, and T) were generated by using pBY100 as the template in a
standard sequencing reaction with primer SK (A) or 207055 (B).
operators. In the presence of 2 mM inducer (glycerol-P), the binding affinity of GlpR was reduced (Fig. 7A, lane 6). Footprints of the internal operators using the same fragment of DNA labeled at the other end, are shown in Fig. 7B. Starting at 4 nM GlpR, the internal operators O₉3 and O₉4 were protected. End-labeled DNA with O₉4 mutations was used for footprinting, and the results are shown in Fig. 8. The protection previously shown for O₉3 and O₉4 disappeared, even when GlpR was present at 40 nM. The disappearance of O₉3 protection when O₉4 was mutated provides strong evidence that O₉3 and O₉4 bind GlpR cooperatively. When a truncated labeled DNA fragment without the tandem operators (O₉1 and O₉2) was used for footprinting (Fig. 9), the internal glpD operators still bound GlpR starting at 6 nM, a concentration slightly higher than that required for protection in the presence of O₉1 and O₉2 (lanes 1 through 6). This means that although the internal operators required O₉1 and O₉2 for repression in vivo, O₉1 and O₉2 were not required in order to demonstrate the binding of GlpR to O₉3 and O₉4 in vitro. Analogous footprinting experiments using labeled DNA containing box 1 and box 2 were carried out. Consistent with the in vivo evidence, box 1 and box 2 were not protected by GlpR, even at 50 nM GlpR (data not shown).

**HU protein facilitates negative regulation of glpD.** Since the distance between O₉2 and O₉3 is 395 bp, accessory proteins such as IHF and/or HU may assist in wrapping the intervening DNA to facilitate cooperative function of the distal glpD operators. To assess the potential regulatory roles of these proteins, the repression of glpD in strains deficient in IHF or HU was measured. Strains BY4006, BY1006, BY4459, and BY1459
Fig. 8. Lack of binding of the *glp* repressor to mutated O₁₄ in the presence of O₁₁ and O₁₂. DNA for footprinting was generated by PCR amplification of pBY109 with primer 254069 and ³²P-end-labeled primer 207055. Reactions were carried out with the following concentrations of the *glp* repressor and analyzed in the indicated lanes: 1, no repressor; 2, 2 nM; 3, 4 nM; 4, 8 nM; 5, 20 nM; 6, 40 nM; 7, 40 nM GlpR and 3 mM glycerol-P. The size markers (lanes A, C, G, and T) were generated by using pBY109 as the template with primer 207055 in a standard sequencing reaction.
Fig. 9. DNase I footprinting of the glpD internal operators (in the absence of O_b1 and O_b2). DNA for footprinting was generated by PCR amplification using SK primer and ^32^P end-labeled primer 207055. Reactions were carried out using the following concentrations of glp repressor tetramers and analyzed in the indicated lanes: 1, no repressor; 2, 1 nM; 3, 3 nM; 4, 6 nM; 5, 9 nM; 6, 20 nM; 7, 40 nM; 8, 9 nM GlpR and 3 mM glycerol-P. The size markers (lanes A, C, G, and T) were generated by using pDA1104 as the template with primer 207055 in a standard sequencing reaction.
TABLE 8. Effect of IHF and HU on regulation of glpD transcription

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant genotype</th>
<th>β-Galactosidase sp act</th>
<th>Repression*</th>
</tr>
</thead>
<tbody>
<tr>
<td>BY4000</td>
<td>$glpR^+ \Phi(glpD^{-}lacz)$</td>
<td>110±5</td>
<td></td>
</tr>
<tr>
<td>BY1000</td>
<td>$glpR2 \Phi(glpD^{-}lacz)$</td>
<td>6,200±150</td>
<td>56</td>
</tr>
<tr>
<td>BY4006</td>
<td>$glpR^+ \Phi(glpD^{-}lacz) hupA16 hupB11$</td>
<td>370±10</td>
<td></td>
</tr>
<tr>
<td>BY1006</td>
<td>$glpR2 \Phi(glpD^{-}lacz) hupA16 hupB11$</td>
<td>5720±30</td>
<td>15</td>
</tr>
<tr>
<td>BY4159</td>
<td>$glpR^+ \Phi(glpD_{mu}^{-}lacz)$</td>
<td>130±5</td>
<td></td>
</tr>
<tr>
<td>BY1159</td>
<td>$glpR2 \Phi(glpD_{mu}^{-}lacz)$</td>
<td>1,000±50</td>
<td>8</td>
</tr>
<tr>
<td>BY4459</td>
<td>$glpR^+ \Phi(glpD_{mu}^{-}lacz) hupA16 hupB11$</td>
<td>320±7</td>
<td></td>
</tr>
<tr>
<td>BY1459</td>
<td>$glpR2 \Phi(glpD_{mu}^{-}lacz) hupA16 hupB11$</td>
<td>1690±34</td>
<td>5</td>
</tr>
<tr>
<td>BY4004</td>
<td>$glpR^+ \Phi(glpD^{-}lacz) \Delta hisA82$</td>
<td>124±3</td>
<td></td>
</tr>
<tr>
<td>BY1004</td>
<td>$glpR2 \Phi(glpD^{-}lacz) \Delta hisA82$</td>
<td>6300±104</td>
<td>50</td>
</tr>
<tr>
<td>BY4259</td>
<td>$glpR^+ \Phi(glpD_{mu}^{-}lacz) \Delta hisA81$</td>
<td>132±2</td>
<td></td>
</tr>
<tr>
<td>BY1259</td>
<td>$glpR2 \Phi(glpD_{mu}^{-}lacz) \Delta hisA81$</td>
<td>960±50</td>
<td>7</td>
</tr>
</tbody>
</table>

* Repression was defined as in Table 7.
(Table 5) were constructed by deletion of the *hupA* and *hupB* genes of strains BY4000, BY1000, BY4159, and BY1159, respectively (Table 5). As the results in Table 8 indicate, the loss of HU lowered repression fourfold (15 versus 56) when all four operators were present but had almost no effect on repression when O$_6$4 was nonfunctional (5 versus 8).

To study the influence of IHF on repression, strains BY4004, BY1004, BY4259 and BY1259 (Table 5) were constructed by deletion of *him$A$* in strains BY4000, BY1000, BY4159 and BY1159, respectively. Repression was measured as described above. From the results shown in Table 8, it is apparent that the loss of IHF had almost no effect on repression of either wild-type (50 versus 56) or O$_6$4-mutated (7 versus 8) *glpD-lacZ* fusions. These results are consistent with the lack of protection by IHF of the potential IHF binding sites between the *glpD* operators, as assessed by DNase I footprinting (data not shown).

II. Regulation of *glpT* gene expression.

**Role of the internal *glpT* operators in repression of *glpT***. In order to find out if the internal operator-like sequences (O$_7$1-3) contribute to GlpR-mediated negative regulation, isogenic *glpR$^-$* and *glpR2* (constitutive) strains with single copy *glpT-lacZ* transcriptional fusions with or without the potential internal *glpT* operator sequences were constructed as described in “Materials and Methods” (Table 5). Strains BY31, BY32, BY741, BY742, BY751 and BY752 (Fig. 10) all have the same *glpT-lacZ* sequences [$\Phi(glpT'-lacZ)$] containing the putative internal *glpT* and the upstream *glpA*
<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant genotype</th>
<th>Fusion Construct</th>
<th>β-Galactosidase Sp act</th>
<th>Repression ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>701, 702</td>
<td>Wild type</td>
<td>&lt;-glpA  glpT --&gt;</td>
<td>210±7  210±29</td>
<td>1</td>
</tr>
<tr>
<td>21, 22</td>
<td>Wild type</td>
<td></td>
<td>110±9  710±67</td>
<td>6.5</td>
</tr>
<tr>
<td>27, 28</td>
<td>himA</td>
<td></td>
<td>180±13 970±70</td>
<td>5.4</td>
</tr>
<tr>
<td>25, 26</td>
<td>hupA11 hupB16</td>
<td></td>
<td>60±9  460±40</td>
<td>7.7</td>
</tr>
<tr>
<td>31, 32</td>
<td>Wild type</td>
<td></td>
<td>80±7  1700±150</td>
<td>21</td>
</tr>
<tr>
<td>37, 38</td>
<td>himA</td>
<td></td>
<td>200±18 2000±200</td>
<td>10</td>
</tr>
<tr>
<td>35, 36</td>
<td>hupA11 hupB16</td>
<td></td>
<td>60±5  980±28</td>
<td>16</td>
</tr>
<tr>
<td>741, 742</td>
<td>Wild type</td>
<td></td>
<td>100±11 1190±14</td>
<td>12</td>
</tr>
<tr>
<td>7410, 7420</td>
<td>himA</td>
<td></td>
<td>250±12 1360±16</td>
<td>5.4</td>
</tr>
<tr>
<td>751, 752</td>
<td>Wild type</td>
<td></td>
<td>210±2  1200±96</td>
<td>5.7</td>
</tr>
<tr>
<td>7510, 7520</td>
<td>himA</td>
<td></td>
<td>348±4  1400±39</td>
<td>4.0</td>
</tr>
<tr>
<td>711, 712</td>
<td>Wild type</td>
<td></td>
<td>120±3  1200±100</td>
<td>10</td>
</tr>
<tr>
<td>721, 722</td>
<td>Wild type</td>
<td></td>
<td>110±9  1100±60</td>
<td>10</td>
</tr>
<tr>
<td>731, 732</td>
<td>Wild type</td>
<td></td>
<td>154±3  1200±90</td>
<td>7.8</td>
</tr>
</tbody>
</table>
Fig. 10. Effects of internal $glpT$ operators, IHF, and HU on repression of $glpT$-$lacZ$ transcriptional fusions. "Relevant genotype" refers to the chromosomal genotype of the $himA$, $hupA$, and $hupB$ alleles. The pairs of $glpR^-$ and $glpR2$ strains with the indicated fusion constructs were assayed for $\beta$-galactosidase after logarithmic growth. Each construct contains the indicated region of DNA fused to $lacZ$ (black box). The repression ratio is defined as the specific activity in the $glpR2$ strain divided by the specific activity in the $glpR^-$ strain. Mutation of IHF1, IHF2, $O_{\gamma}1$ or $O_{\gamma}3$ is indicated by XXX.
control regions, except that BY741 and BY742 carry the O$_{3}$-1 mutation and BY751 and
BY752 carry the O$_{3}$-3 mutation. Strains BY21 and BY22 have a shorter glpT-lacZ
sequence [$\Phi$(glpT$^{+}$-lacZ)] containing the glpA control region and the glpT sequence
downstream to the IHF1 site in glpT. Strains BY701 and BY702 contain only the internal
glpT sequence [$\Phi$(glpT$^{+}$-lacZ)] fused to lacZ without any glpA operators. From the
results shown in Fig. 10, it is clear that repression of glpT is conferred mainly by the glpA
operators since there is no repression control without the glpA operators (BY701 and
BY702, repression ratio of 1). In the presence of the glpA operators, the internal glpT
operators increased repression of glpT three to four fold, observed by comparing the
repression of constructs without O$_{3}$-1-3 (expression of BY22 divided by that of BY21 is
6.5) with repression of constructs containing O$_{3}$-1-3 (expression of BY32 divided by
BY31 is 21). Destruction of O$_{3}$-1 or O$_{3}$-3 by mutation had differential effects on
repression of glpT. Mutation of O$_{3}$-1 (glpT$_{MOTI}$) decreased repression about two fold
(BY741 and BY742, repression ratio of 12), while mutation of O$_{3}$-3 (glpT$_{MOTI}$) decreased
repression about four fold (BY751 and BY752, repression ratio of 5.7). Repression of
constructs with O$_{3}$-3 mutated was similar to that of constructs lacking the internal glpT
operators, confirming that the internal glpT operators contributed about three to four fold
to repression in vivo.

**DNase I footprinting of the internal glpT operators.** By using DNase I
footprinting, it was shown that the internal glpT operators bound GlpR with or without
the glpA operators. In the presence of the glpA operators (all of the glpA operators were

67
protected by GlpR using DNase I footprinting in this study, data not shown), footprints of the \(glpT\) operators are shown in Fig. 11A. Starting at 3 nM, the \(O_{1}1-O_{1}3\) region was protected (+305 to +361 in \(glpT\); Fig. 6). This is a relatively weak \(glp\) operator region considering a higher concentration of GlpR was needed for protection compared with other defined \(glp\) operators which are protected at less than 1 nM GlpR (17,39,40).

Three hypersensitive sites were induced by binding of GlpR (sites 1, 3 and 4 within the protected area, indicated by arrows, see Figs. 6 and 11A). Another atypical region (marked site 2, indicated by an arrow, see Figs. 6 and 11A) located within the internal operators was not protected even at 12 nM GlpR. Site 2 was between \(O_{1}1\) and \(O_{1}2\) and site 4 was between \(O_{1}2\) and \(O_{1}3\), which suggests that GlpR binds simultaneously to each pair of tandem operators. It is somewhat unusual that site 1 is at the beginning of \(O_{1}1\) and site 3 is in the middle of \(O_{1}2\). This unusual pattern may be due to the perturbation of DNA structure caused by simultaneous binding of GlpR at internal \(glpT\) operators and the distant \(glpA\) operators. When a truncated labeled DNA fragment without the \(glpA\) operators was used for footprinting, as indicated in Fig. 11B, the internal \(glpT\) operators still bound GlpR with a similar protection pattern starting at 3 nM GlpR. This means that although the internal \(glpT\) operators required \(glpA\) operators to function in vivo, the binding of GlpR is strong enough to be demonstrated in vitro in the absence of the \(glpA\) operators. The DNase I cleavage pattern in Figs. 11A and 11B also suggested that \(O_{1}3\) bound GlpR better than \(O_{1}1\), and \(O_{1}2\) bound GlpR the weakest.

DNase I footprinting was also used to reassess binding of GlpR to \(O_{1}\), a region
Fig. 11. DNase I footprinting of the wild type *glpT* internal operator region.

A. Identification of the internal *glpT* operators in the presence of the *glpA* operators. End-labeled DNA was generated by PCR using primer 223822 and $^{32}$P end-labeled primer 311334 with pGS31H5 as template. Reactions were carried out using the following concentrations of GlpR tetramers and analyzed in the indicated lanes: 1, no repressor; 2, 0.2 nM; 3, 0.5 nM; 4, 1 nM; 5, 3 nM; 6, 6 nM; 7, 9 nM; 8, 12 nM; 9, 9 nM; 10, 12 nM GlpR. The reactions loaded in lanes 9 and 10 each contained 2 mM glycerol-P. The size standards (ACGT) were generated using pGS31H5 as template with primer 311334 in a standard sequencing reaction. Sites 1, 2, 3, and 4 shown in Fig. 1 are indicated by the arrows.

B. Identification of the internal *glpT* operators in the absence of *glpA* operators. End-labeled DNA was generated by PCR using primer 353988 and $^{32}$P end-labeled primer 198093 with pBY70 as template. Reactions were carried out using the following concentrations of GlpR tetramers and analyzed in the indicated lanes: 1, no repressor; 2, 1 nM; 3, 3 nM; 4, 6 nM; 5, 9 nM; 6, 12 nM GlpR. The size standards (ACGT) were generated using pBY75 as template with primer 198093 in a standard sequencing reaction.
overlapping the CRP site for \( glpT \) (Fig. 6). No evidence was found for binding of GlpR to this region (data not shown), which was weakly footprinted in a previous study (40). Therefore, this region is not an operator, as further supported by the observation that mutation of this site has no effect on repression of \( glpT \) (40).

DNA fragments containing the \( glpA \) operators with either \( O_{r1} \) or \( O_{r3} \) mutated were used for footprinting, with the results shown in Figs. 12A and 12B. Mutation of \( O_{r1} \) abolished the protection of \( O_{r1} \) (Fig. 12A), but GlpR still bound \( O_{r3} \) starting at 6 nM GlpR. Mutation of \( O_{r3} \) (Fig. 12B) abolished the binding to GlpR of all internal \( glpT \) operators even with 20 nM GlpR, a concentration much higher than required for protection of \( glp \) operators (17,39,40). These findings are consistent with footprinting of the wild-type sequence, and are consistent with the in vivo data indicating that destruction of \( O_{r3} \) decreased repression twice more than destruction of \( O_{r1} \) (5.7 versus 12, Fig. 10).

**IHF facilitates negative regulation of \( glpT \) transcription.** From the in vivo results shown above, the internal \( glpT \) operators contributed about three to four fold to repression of \( glpT \), but only in the presence of the \( glpA \) operators. It is possible that the internal \( glpT \) operators function by forming a DNA loop with the \( glpA \) operators when they all bind GlpR. Accessory proteins such as IHF or HU may assist bending or wrapping the DNA to facilitate such a loop formation. To assess a potential role for these proteins, repression of \( glpT \) was measured in strains deficient in IHF or HU. Strains BY27, BY28, BY37 and BY38 (Table 5) were constructed by deletion of \( himA \) in
Fig. 12. DNase I footprinting of the mutated internal \textit{glpT} operators.

A. Lack of binding of GlpR to mutated O$_{11}$ in the presence of \textit{glpA} operators. End-labeled DNA was generated by PCR using primer 109180 and $^{32}$P end-labeled primer 198093 with pBY74 as template. Reactions were carried out using the following concentrations of GlpR tetramers and analyzed in the indicated lanes: 1, no repressor; 2, 3 nM; 3, 6 nM; 4, 9 nM; 5, 20 nM. The size standards (ACGT) were generated using pBY74 as template with primer 198093 in a standard sequencing reaction.

B. Lack of binding of GlpR to mutated O$_{13}$ in the presence of the \textit{glpA} operators. End-labeled DNA was generated by PCR using primer 109180 and $^{32}$P end-labeled primer 198093 with pBY75 as template. Reactions were carried out using the following concentrations of GlpR tetramers and analyzed in the indicated lanes: 1, no repressor; 2, 3 nM; 3, 6 nM; 4, 20 nM. The size standards (ACGT) were generated using pBY75 as template with primer 198093 in a standard sequencing reaction.
strains BY21, BY22, BY31 and BY32 (Table 5), respectively. As documented in Fig. 10, the absence of the HimA subunit of IHF lowered repression two fold in the presence of glpA operators (10 versus 21), but had almost no effect on repression when glpA operators were absent (5.4 versus 6.5). To further prove that the decreased repression ratio found for strains BY37 and BY38 is due to the absence of IHF, these strains were transformed with a plasmid expressing the subunits of IHF under the control of the tac promoter. Repression was measured with different induction of the plasmid using the vector as a control under the same condition. The results are shown in Table 9. It was found that induction of IHF synthesis resulted in a two- to three-fold increase in repression when compared with the repression of strains transformed with the vector control plasmid. With no induction, similar repression was found (7.5 to 10.6).

Experiments on the effect of the himA deletion on repression were also carried out using the constructs with the O71 or O73 mutations. In this case, strains BY7410, BY7420, BY7510 and BY7520 (Table 5) were constructed by deletion of himA in strains BY741, BY742, BY751 and BY752, respectively (Table 5). Repression was measured as described above. Loss of IHF had almost no effect on repression for constructs with the O73 mutation (4.0 versus 5.7, Fig. 10), but still decreased about two fold repression in constructs with the O71 mutation (5.4 versus 12, Fig. 10). The results are consistent with in vitro data showing that mutation of O71 allowed repressor binding to O73, while mutation of O73 abolished repressor binding to all internal glpT operators. Therefore, IHF may facilitate repression in constructs with the O71 mutation since O73 was still able
TABLE 9. IHF encoded by a plasmid complements the effect of chromosomal \textit{himA} deletion on \textit{glpR} transcription

<table>
<thead>
<tr>
<th>Plasmid\textsuperscript{a}</th>
<th>IPTG\textsuperscript{b}</th>
<th>\textit{glpR}\textsuperscript{c}</th>
<th>\textit{glpR}2</th>
</tr>
</thead>
<tbody>
<tr>
<td>pHNβ α</td>
<td>0</td>
<td>168</td>
<td>1780</td>
</tr>
<tr>
<td>pHNβ α</td>
<td>50</td>
<td>103</td>
<td>1595</td>
</tr>
<tr>
<td>pHNβ α</td>
<td>100</td>
<td>41</td>
<td>1471</td>
</tr>
<tr>
<td>pAB223</td>
<td>0</td>
<td>260</td>
<td>1942</td>
</tr>
<tr>
<td>PAB223</td>
<td>50</td>
<td>299</td>
<td>2398</td>
</tr>
<tr>
<td>PAB223</td>
<td>100</td>
<td>211</td>
<td>2244</td>
</tr>
</tbody>
</table>

\textsuperscript{a} pHNβ α was a gift from H. Nash. This plasmid carries \textit{lac} and expresses the two subunits of IHF from \textit{P}_{\text{lac}}. PAB223 is vector control carrying \textit{lac} and \textit{P}_{\text{lac}} with no gene inserted.

\textsuperscript{b}Overnight cultures, grown in the absence of IPTG, were diluted 50 times into fresh medium containing the indicated concentration of IPTG. \(\beta\)-galactosidase activities are the mean values determined from two independent cultures after growth for a minimum of 5 doublings. Values for \textit{glpR} and \textit{glpR}2 were determined in strains BY37 and BY38, respectively.

\textsuperscript{c}Repression is defined as in Table 7.
to form a repression loop with the $glpA$ operators. IHF had no effect on constructs with the O$_1$3 mutation, possibly because GlpR only bound the $glpA$ operators in this case.

In order to determine if IHF exerts its effect by binding the potential IHF sites between $glpA$ and internal $glpT$ operators, specific mutations of IHF1 and/or IHF2 were introduced into the constructs containing the $glpA$ and the internal $glpT$ operators as described in the Materials and Methods. Mutation of IHF1 or IHF2 lowered the repression from 21 to 10 (Fig. 10). Mutations of both IHF1 and IHF2 lowered the repression from 21 to 7.8 (Fig. 10), a slightly larger effect than the mutation of the individual IHF sites. Therefore, it appears that IHF functions directly by binding these two sites, most likely by bending the DNA in vivo.

To study the influence of HU on repression, strains BY25, BY26, BY35 and BY36 (Table 5) were constructed by deletion of the $hupA$ and $hupB$ genes of strains BY21, BY22, BY31 and BY32, respectively (Table 5). Repression was measured as described above. From the results shown in Fig. 10, it is apparent that loss of HU had almost no effect on repression of the constructs with (16 versus 21) or without the internal $glpT$ operators (7.6 versus 6.5).

**DNase I footprinting of the IHF sites.** End-labeled DNA fragments containing the sites IHF1 and IHF2 were used in DNase I footprinting. Protection of IHF1 (+15 to +51, see Fig. 6) of the wild type DNA occurred, starting at 5 nM IHF (Fig. 13A). Footprinting of IHF2 using wild type DNA is shown in Fig. 13B. Starting at 10 nM IHF, the region IHF2 (+193 to +227 in $glpT$, Fig. 6) was protected. End-labeled DNA with
Fig. 13. DNase I footprinting of the wild-type IHF sites within glp7.

A. Radiolabeled DNA was prepared by PCR using primer 223822 and $^{32}$P end-labeled primer 311355 with pGS31H5 as template. Reactions were carried out using the following concentrations of IHF dimers and analyzed in the indicated lanes: 1, no IHF; 2, 5 nM; 3, 10 nM; 4, 20 nM; 5, 30 nM; 6, 40 nM; 7, 50 nM; 8, 100 nM IHF.

B. Radiolabeled DNA was prepared by PCR using primer 223822 and $^{32}$P end-labeled primer 311334 with pGS31H5 as template. Reactions were carried out using the following concentrations of IHF dimers and analyzed in the indicated lanes: 1, no IHF; 2, 5 nM; 3, 10 nM; 4, 20 nM; 5, 30 nM; 6, 40 nM IHF. The size markers (ACGT) for (A) and (B) were generated using pGS31H5 as template with primers 311355 and 311334 in standard sequencing reactions, respectively.
mutations in either IHF1 or IHF2 was used for footprinting and the results are shown in Fig. 14. The previously shown protection of IHF1 or IHF2 was abolished even when 200 nM IHF protein was used. Since IHF1 contains a sequence with one mismatch (AATCAGGCTGTAA, +30 to +42 in glpT, Fig. 6) and IHF2 contains a sequence with two mismatches (AACCAAATAGTAA, +209 to +197 in glpT bottom strand, Fig. 6) compared with the IHF consensus sequence (89), it is reasonable that IHF1 requires less protein for protection than IHF2. In both cases, the sequence resembling the consensus was asymmetrically located toward one side of the protected area, as noted earlier (89). Footprinting using 5 nM IHF also protected one other region that overlapped Oγ3 in glpA (the sequence resembling the IHF consensus is indicated in Fig. 6; data not shown). The two other potential IHF sites within the glpA4T promoter region were not protected even at 100 nM IHF, a non-physiological concentration for IHF footprinting (92).
Fig. 14. DNase I footprinting of the mutated IHF sites. Reactions were carried out using the following concentrations of IHF dimers for both A and B, and analyzed in the indicated lanes: 1, no IHF; 2, 25 nM; 3, 50 nM; 4, 100 nM; 5, 200 nM IHF.

A. Radiolabeled DNA was prepared by PCR using primer 366852 and $^{32}$P end-labeled primer 311355 with pBY71 as template.

B. Radiolabeled DNA was prepared by PCR using primer '366852 and $^{32}$P end-labeled primer 353953 with pBY72 as template. The size markers (ACGT) for (A) was generated using pBY71 as template with primer 311355 in a standard sequencing reaction. The size marker (ACGT) for (B) was generated using pBY72 as template with primer 353953 in a standard sequencing reaction.
DISCUSSION

Cooperative binding by a regulatory protein to distant DNA sites is a common mechanism for both negative and positive regulation of transcription in prokaryotic and eukaryotic organisms (65,135,136). In *E. coli*, it has been demonstrated that at least two spatially separated operators are necessary to give full repression of transcription of the *gal*, *deo*, and *lac* operons (65). Similar evidence of repressor binding to widely separated operators has been found for control of the *put* operon of *Salmonella typhimurium* (137). In all cases, DNA loop formation mediated by repressor binding to distal operators seems to be the mechanism for repression (65,136). As is the case for the eukaryotic enhancers, widely separated sites on bacterial DNA seem to interact by protein-mediated DNA loop formation (135,136), as further supported by the recent report that a twofold increase in repression was exerted by a second distal *lac* operator when two *lac* operators were separated by distances ranging from 600 to 1,500 bp (138). The finding of a similar situation for repression of the *glpD* and *glpT* operons agrees with the suggestion that remote operator duplications are usually associated with multiple promoters subject to more than one system of regulation (136).

In the case of *glpD*, besides regulation by GlpR, the promoter is subject to regulation by cAMP-CRP and by the ArcA/ArcB systems (43). The adjacent, divergent *glpEGR* promoter may share some of these *cis*-acting regulatory elements; therefore, it may be argued that the remote placement of operators O₆₃ and O₆₄ avoids crowding of regulatory elements near the promoters (136). An emerging general feature of negative
control by GlpR is that GlpR binds cooperatively to tandem operators (O_{D1} and O_{D2} or O_{A2} and O_{A3}) located at or near the promoter, with additional contributions to repression provided by remote operators (O_{D3} and O_{D4} or O_{T1} to 3) to form a nucleoprotein complex with the intervening DNA forming a loop.

The unique feature of the *glpT* operon, compared to the above systems, is that it does not have an operator overlapping any of its promoter elements. Instead, repression is partially controlled by the *glpA* operators located far upstream of the *glpT* transcriptional start, with the closest *glpA* operator centered at -81.5. Full repression is achieved by GlpR binding at the *glpA* and the internal *glpT* operators, possibly by forming a nucleo-protein complex with the intervening DNA forming a loop. In the *glpAT* control region, regulatory elements for control of transcription of the divergent operons are crowded upstream of the *glpT* transcription start site (Fig. 6). Besides regulation exerted by GlpR binding at four *glpA* operators centered at -81.5 (O_{A1}), -113.5 (O_{A2}), -133.5 (O_{A3}), -173.5 (O_{A4}), there is one CRP site at -41.5 in *glpT*, one CRP/FNR site overlapping O_{A1} and another CRP/FNR site overlapping O_{A3}. Because *glpA* and *glpT* both share the positive and negative regulatory elements, the remote placement of the internal *glpT* operators helps to alleviate the crowding of regulatory elements near the promoters (136).

A typical operator that binds a dimeric repressor contains a palindromic sequence, with each half of the symmetric site binding one monomer of the repressor. In the case of a repressor that binds tandem operators, the repressor dimers or tetramers lie
adjacent to each other and cooperative binding increases the local concentration of the second repressor dimer. It is apparent that a plasmid containing tandem operators $O_{D}1$ and $O_{D}2$ binds the repressor 10 times better than does a plasmid harboring only $O_{D}1$. In addition, although 21 bp intervene between $O_{D}3$ and $O_{D}4$, the observation that substitutions in $O_{D}4$ eliminated protection of $O_{D}3$ in vitro is strongly suggestive of cooperative binding of the repressor at $O_{D}3$ and $O_{D}4$. $O_{D}1$ matches the $glp$ consensus operator perfectly, but $O_{D}2$, $O_{D}3$ and $O_{D}4$ have 5, 6, and 4 mismatches, respectively, compared with the 20-bp consensus $glp$ operator sequence (53). It is reasonable that GlpR binding at a better operator like $O_{D}1$ or $O_{D}4$ facilitates cooperative binding of the weak operator, $O_{D}2$ or $O_{D}3$, respectively. It is known that GlpR binds cooperatively to tandem consensus operators (53). Cooperative binding of the repressor to the two internal operators may be possible because the distance between the centers of symmetry of $O_{D}3$ and $O_{D}4$ is 41 bp (about four helical turns), a distance that would allow GlpR binding to the same face of DNA.

Cooperative binding was also apparent as GlpR bound to the internal $glpT$ operators. Mutation of $O_{T}3$ abolished binding of GlpR to all internal $glpT$ operators in vitro. This strongly suggests that GlpR binds to the three internal operators in a cooperative fashion. Although $O_{T}1$ and $O_{T}3$ each have five mismatches compared to the consensus $glp$ operator sequence (53), they definitely play differential roles in the tandem or triple binding of GlpR, as indicated by the fact that GlpR still bound $O_{T}3$ when $O_{T}1$ was mutated, but bound no operators when $O_{T}3$ was mutated. The effect of mutating $O_{T}2$
was not studied, since O\textsubscript{12} has 8 mismatches compared to the consensus glp operator sequence (53) and probably could not bind GlpR if placed separately. Since the distance between O\textsubscript{1} and O\textsubscript{12} is 2 bp, GlpR binding at a better operator like O\textsubscript{1} could facilitate cooperative binding to a weak operator like O\textsubscript{12}. And because the distance between O\textsubscript{1} and O\textsubscript{3} is 13 bp, about one helical turn of DNA, O\textsubscript{1} and O\textsubscript{3} might function as tandem operators.

In both glpD and glpT transcription, the fact that the internal operators were unable to function alone in vivo even though these operators did bind GlpR in vitro suggested a further cooperative mechanism, with GlpR simultaneously binding all the glpD or glpAT operators, respectively. These internal operators are several hundred base pairs distal to the transcription start sites and thus should not directly affect the binding of RNA polymerase (65,136). This is probably the reason that most functional operators overlap or are close to the transcription initiation sites (136). The results of the present study of glpD control indicate that the internal operators O\textsubscript{p3} and O\textsubscript{p4} exert their function in vivo by cooperative binding of GlpR tetramer(s) with O\textsubscript{p1} and O\textsubscript{p2}, since the internal operators had no effect in vivo in the absence of O\textsubscript{p1} and O\textsubscript{p2} but contributed five- to sevenfold to repression in the presence of O\textsubscript{p1} and O\textsubscript{p2}. By using plasmid-encoded transcriptional fusions of the promoter-operator regions to lacZ and repressor provided by a multicopy plasmid, a repression ratio of 400 was found for transcriptional control of the glpF promoter, a ratio four times higher than that of an analogous glpD-lacZ construct containing only the promoter and tandem operators O\textsubscript{p1} and O\textsubscript{p2} (17). By
using a *glpD* fusion containing the internal operators, it is now clear that O₃ and O₄ contribute additionally five- to sevenfold to repression. Therefore, *glpD* is likely to be repressed more tightly than is *glpFKX*, consistent with Lin's observation that the *glpD* operon is controlled more tightly than is either the *glpTQ* or *glpFKX* operon, as determined by measuring enzyme activities encoded by each operon (2,3,52).

The *glpT* internal operators facilitated repression by about three to four fold, but only when the *glpA* operators were present, suggesting the internal operators also function cooperatively by forming a nucleo-protein complex with GlpR binding to all of the *glpA* and *glpT* operators. Since O₃ is more important than O₁ in vitro for cooperative binding of repressor at O₁-3, O₃ is probably more important in forming such a nucleo-protein complex as supported by the finding that mutation O₃ had about twice the effect on repression relative to mutation O₁ (5.7 versus 12).

It was shown that even without the supporting proteins, intrinsic curving and bending play an essential role in protein-DNA recognition (139). The distance between O₂ and O₃ is 395 bp, a distance suitable for DNA loop formation as GlpR binds to the four *glpD* operators. There is a predicted intrinsic bend in the DNA within *glpD* (from positions 521 to 661, covering the internal *glpD* operators [Fig. 5]) with a predicted bend of 77° (140). The intrinsic bend might contribute additionally to wrapping of the DNA to form a loop. However, the accessory protein HU contributes additionally to repression, presumably by assisting in DNA bending. The influence of HU for *glpD* was similar to that found for repression of the *gal* operon promoter P2, in which a two to
threefold effect was seen (109).

In the case of the g\(lpT\) control region, the distance between O\(\lambda\)1 and O\(\tau\)1 is 394 bp, almost the same as that between O\(\delta\)2 and O\(\delta\)3 (395 bp). IHF binding at IHF1 and IHF2 might cooperate in the formation of a looped nucleoprotein complex by bending the DNA to bring the repressor-DNA complex at g\(lpA\) operators and internal g\(lpT\) operators together. This DNA looping model is shown in Fig. 15. It is interesting that the distance between O\(\lambda\)1 and IHF1 is similar to that between O\(\tau\)1 and IHF2 (both about 110 bp). This architectural arrangement might favor the cooperative interactions of the GlpR-operator complex and thus strengthen the so formed complex. The complex containing IHF apparently provides a two-fold increase in repression by GlpR at the g\(lpT\) promoter, in the presence of the g\(lpT\) internal operators (21 versus 10 or 7.8).

Repression of the fusion containing the internal operators in the absence of IHF is still stronger than the short fusion without the internal operators in the presence of IHF (10 versus 6.5), suggesting that the major mechanism of repression is the loop formation between the distant GlpR bound operators. This is further supported by the observation that deleting IHF caused a two-fold decrease in repression in the construct with O\(\tau\)1 mutated (5.4 versus 12), suggesting that O\(\tau\)3 still bound GlpR. Repression in the construct with O\(\tau\)3 mutated was not further affected by deletion of IHF (4.0 versus 5.7) as no internal operators bound GlpR. These repression values are similar to those of the fusion contracts missing the internal operators (5.4 and 6.5 for BY27/28 and BY21/22, respectively). For the constructs with the internal g\(lpT\) operators, mutation of IHF1
Fig. 15. Model of the IHF-GlpR-operator complex for control of *glpT* transcription. The solid circle represents the GlpR protein oligomers. The solid line represents DNA of the *glpAT* control region. The degree of IHF bending is shown (40°). "A" represents the distance between IHF1 and O_{A1} (about 110 bp). "B" represents the distance between IHF2 and O_{T1} (about 120 bp). O_{A1} is used as an example of the *glpA* operators in formation of the repressor-DNA complex because of its symmetric location with respect to the *glpT* operators in the model. The directions of the transcription of *glpT* and *glpA* are indicated by the arrows positioned approximately at the positions of transcription initiation.
and/or IHF2 had a similar effect as deletion of the himA gene, proving that the sites IHF1 and IHF2 play physiological roles.

The importance of the sites IHF1 and IHF2 was further supported by showing that the decrease in repression of glpT in IHF-deficient strains is not likely due to a decrease in glp repressor concentration in these strains. Introduction of a himA deletion into strains with glpD-lacZ or glpK-lacZ fusions did not influence glp repressor-mediated repression of these fusions, suggesting that the differences of the repression in wild type or himA strains were not due to the change of the level of GlpR.

Footprinting was also carried out using GlpR and IHF together for the internal glpT operator regions. The presence of IHF, however, did not significantly change the footprinting pattern of GlpR at the internal glpT operators (data not shown). Since binding of GlpR in vitro to the internal glpT operators occurred in the absence of the glpA operators, it is conceivable that IHF has no effect on binding of the internal glpT operators when using the linearized DNA.

An effect of IHF on GlpR-mediated control of glpACB transcription was ruled out by comparing the effect of deleting himA on the glpA-lacZ transcriptional fusion in strain TL684 (10) using anaerobic growth conditions. Repression ratios of about 10 were found for both himA+ and ΔhimA strains (data not shown). A possible reason for the lack of an effect of IHF deficiency on glpACB control is that four glpACB operators overlapping the promoter elements already provide strong repressor control. An additional small contribution to repression by IHF may not be detectable in this case. Similarly, IHF has
almost no effect on repression of $glpT$ transcription in constructs without the $glpT$
internal operators or in constructs with O$_{73}$ mutated, probably because repression of $glpT$
in the absence of the internal $glpT$ operators depends only on $glpA$ operators, to which
repressor binds tightly. Thus, there is no influence of IHF in this case. Despite the
fact that HU and IHF have many structural and functional similarities, it was shown that
IHF had little or no effect on $glpD$ transcription while HU increased repression of $glpD$
about fourfold but had no effect on $glpT$ transcription. The mechanisms for the binding
of IHF and HU are predicted to be very similar, and these proteins act interchangeably in
some cases (93). Nevertheless, each protein functions specifically in one operon but has
no effect on the other operon, arguing that selective action of HU and IHF in different
operons might enhance the ability for specific regulation.

In the study of $glpT$ transcription, deletion of IHF always allowed slightly higher
constitutive expression of each $glpT$-lacZ construct, relative to that found in the himA$^-$
strain (Fig. 10). It is possible that binding of IHF at IHF1 may have a small negative
effect, in the absence of GlpR, because this site is within the range covered by RNA
polymerase (40-60 bp (141)). Because IHF1 is some distance away from the
transcription start site, the effect is not very striking in the absence of GlpR.
Part II

Transcriptional organization of the glpEGR operon

INTRODUCTION

The glpEGR genes are thought to be cotranscribed and thus form a single operon (54, 143). The nucleotide sequence of the glpEGR genes has been determined in our laboratory (54). The initiation codons for glpE, glpG and glpR as well as the translational reading frames used by these three genes have been determined (54). The subcellular locations of GlpE and GlpG were determined, with GlpE found in the cytoplasmic fraction and GlpG in the cytoplasmic membrane fraction (54). From the deduced amino acid sequence, GlpE is predicted to be an acidic protein with a molecular weight of 12,082 and GlpG is predicted to be a protein with a molecular weight of 31,278 (54). Although the calculated pl of of GlpG is 9.3, GlpG protein is relatively hydrophobic (54). The GlpE and GlpG proteins were first detected by using minicells or maxicells labeled with [35S]methionine (15).

The 5' end of the mRNA presumed to originate from the glpE promoter was previously mapped (142). The glpE promoter and the divergently transcribed glpD promoter share a CRP binding site symmetrically located between them. This is a weak CRP activation site since the activation of both promoters caused by binding of CRP-cAMP complex is only twofold (17, 142).
Previous work from our laboratory indicated that the glpEGR genes are cotranscribed only from the glpE promoter (143). However, I found that the assays and constructs used to draw that conclusion had mistakes. Therefore, the actual transcriptional organization of the glpEGR genes remains undefined. More precise work, as described below, revealed that there are relatively strong promoters upstream of both glpE and glpG, and two very weak promoters upstream of glpR. Computer analysis of the glpEGR sequence revealed no sequences that perfectly match the consensus $\sigma^{70}$ promoter elements, making it impossible to predict the location of the promoters. Therefore, different regions of the glpEGR genes were cloned into a promoter-probe vector in order to localize the regions containing the promoters. Primer extension analysis was carried out to localize the 5' ends of the newly found promoters.

The possible regulation of the glpEGR genes by CRP and by GlpR was studied by several researchers. For catabolite regulation, it was reported that the CRP site located between the divergent glpD and glpE promoters conferred twofold activation of both promoters (17, 142). However, no catabolite regulation was apparent when the expression of a glpEGR-lacZ fusion was assessed after growth in the presence of glucose, compared with cells grown in the absence of glucose (H. Schweizer, unpublished data from this lab). Discrepancies also exist regarding autoregulation by GlpR. Using a glpEGR-lacZ translational fusion carried by a high copy number plasmid, H. Schweizer found no autoregulation by glpR provided in trans on a compatible plasmid (unpublished data). However, it was reported by D. Austin in this lab that glpR is autoregulated by
using *glpE-lacZ* and *glpEGR-lacZ* transcriptional fusions carried by a high copy number plasmid with *glpR* provided in *trans* on a compatible plasmid (143). The differences described above may be due to differences in the fusion constructs used or due to the presence of multiple and/or differentially regulated promoters. The work described below was performed to find the reasons for these discrepancies and to define the actual transcriptional organization and regulation of the *glpEGR* operon.
MATERIALS AND METHODS:

Strains and growth media. The three bacterial strains used in this study were all
E. coli K-12 derivatives. DH5αF'[φ80d lacΔM15 Δ(lacZYA-argF)U169 deoR recA1
endA1 hsdR17 supE44 thi-1 gyrA96 relA1] (53) was used as the host for DNA
manipulations and plasmid expression for lacZ. Strains NZ45 [MC4100 Φ(glpK-lacZ)
ΔglpEGR::KanR recA1 srl::Tn10 λlacR tetR SpR] (constructed by N. Zhao, described in
(145)) and TL73 [MC4100 glpR2 recA1] (31) are derivatives of MC4100 (Table 5,
(132)). Appropriate growth media were used as described in Part I.

Construction of plasmids and assay of β-galactosidase. Oligonucleotides used
in this study are listed in Table 10. Plasmids pBY15, pBY200, pBY201, pBY400,
pBY300, pBY301, pBY302, pBY415, pBY311, and pBY312 were constructed by
subcloning DNA fragments into pSP417 as described in Table 11. Plasmids pBY321,
pBY322, pBY323, pBY324, BY325, pBY326, pBY331, BY332. pBY327, pBY328,
pBY341, pBY342, pBY343, pBY344, pBY345, pBY346, pBY347 and pBY348 were
constructed by using PCR amplified DNA cloned into pSP417 as described in Table 11.
Vector pSP417 is the promoter-probe vector described in Part I. Hence, the level of
expression of β-galactosidase from these plasmids indicates the existence and relative
strengths of possible promoter(s) within the region cloned. β-galatosidase specific activity
was measured as described in Part I.

Purification of RNA and primer extension analysis. Total RNA from various
strains harboring plasmids was isolated as previously described (144). For primer
TABLE 10. Oligonucleotides used in this study

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence*</th>
<th>Positions in $glpEGR^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>253617</td>
<td>TTGTAAGAAttGCAGACGCATGGATC</td>
<td>353 to 379</td>
</tr>
<tr>
<td>253592</td>
<td>TCGTAGCCCTcTaaGAGCAGGATACTGCGCC</td>
<td>623 to 594</td>
</tr>
<tr>
<td>253771</td>
<td>CAATAGCAGCGAAAttCGCGGCAGTATC</td>
<td>576 to 604</td>
</tr>
<tr>
<td>238101</td>
<td>CGCCATGTAATCtAgAAACGCGCTGCCGCACG</td>
<td>806 to 776</td>
</tr>
<tr>
<td>253613</td>
<td>GCAATTCGCGGaaTTcGCGGCGCTTTTCTG</td>
<td>977 to 1007</td>
</tr>
<tr>
<td>253715</td>
<td>CGCATACACACACTcTaAGAAAGCGCAAGCCCCCAAAAC</td>
<td>1361 to 1331</td>
</tr>
<tr>
<td>311253</td>
<td>GCGGCGCGTGAaaTTcGCGGCGCTTTTCTG</td>
<td>1321 to 1349</td>
</tr>
<tr>
<td>238253</td>
<td>CCAAGCTCTTTCGGTctaGACATAACCCCTGCTG</td>
<td>1666 to 1636</td>
</tr>
<tr>
<td>279870</td>
<td>CATATTAGCAAAAAATtcCGCTTTTAGGTAACATTTG</td>
<td>232 to 266</td>
</tr>
<tr>
<td>279484</td>
<td>CCTGCAACTTCTCaGAGCGGCTCGGCAACG</td>
<td>424 to 396</td>
</tr>
<tr>
<td>311308</td>
<td>CCACCACCGACCCGAGATAACC</td>
<td>1239 to 1219</td>
</tr>
<tr>
<td>301396</td>
<td>TATCAGGGgAAAttCAGCAAAGGCACG</td>
<td>568 to 594</td>
</tr>
<tr>
<td>333600</td>
<td>CAAAGGCCGtCtagAATGTATAGAC</td>
<td>653 to 629</td>
</tr>
<tr>
<td>301411</td>
<td>AACGTCAAGaaTtCCCGAGAGGTTG</td>
<td>662 to 685</td>
</tr>
<tr>
<td>333486</td>
<td>AAAAGAGGTTctagATCAACATCGTTGC</td>
<td>764 to 738</td>
</tr>
<tr>
<td>329718</td>
<td>CAGGTCCGaaCCTGGGTTGATGATG</td>
<td>1021 to 1046</td>
</tr>
</tbody>
</table>

99
<table>
<thead>
<tr>
<th>Accession</th>
<th>Sequence 1</th>
<th>Accession</th>
<th>Sequence 2</th>
<th>Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>329865</td>
<td>ATCACTTCTcAGATCGCCGAGAATTTG</td>
<td>333605</td>
<td>CGGGCCG7GaaTTcGCGGGCTTTCTGGC</td>
<td>1102 to 1077</td>
</tr>
<tr>
<td>336837</td>
<td>TGCCACTTTTcTaGATCGCGTTTCGCC</td>
<td>1322 to 1349</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bing4</td>
<td>GTGTTGGATCGAATGACCCAG</td>
<td>1134 to 1114</td>
<td></td>
<td></td>
</tr>
<tr>
<td>91687</td>
<td>AAATCAAAACCATCGCGCGAC</td>
<td>1474 to 1455</td>
<td></td>
<td></td>
</tr>
<tr>
<td>69882</td>
<td>CCCCCCCGcGcTTCTCGTTTTCGCGCATT</td>
<td>1582 to 1557</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11360</td>
<td>TATAAAGCGTTACGCCGTACGCCAC</td>
<td>708 to 682</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*L*owercase letters indicate bases substituted in order to create restriction sites.

*See Fig. 16.*
Fig. 16. Nucleotide sequence of the glpE, glpG, and glpR genes and the deduced amino acid sequence. The sequence beginning at the EcoRI site within the 5' end of the glpD gene is shown. Sequences resembling the consensus sequences for the -10 and -35 regions of the glpD and glpE promoters and for ribosome binding are indicated by solid underlining. The transcription initiation sites for the glpD (39) and glpE (144) genes are indicated by +1. O₁ and O₂ are the operator sites for the glpD gene, and CRP is the interaction site for CRP (dashed underlining) (39). The bases corresponding to the 5' ends determined by primer extension for the glpG promoter are double-underlined. The potential CRP binding sequences are in boldface type. The potential GlpR binding sequences are in boldface type and underlined. The potential FIS binding sites are italicized letters. (Adapted from (143))
<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Source of insert or PCR primers (sites used)</th>
<th>Vector (sites used)</th>
<th>Position$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBY15$^t$</td>
<td>pSH79 ($NruI$ - $NruI$)</td>
<td>pBS ($SmaI$)</td>
<td>453 to 2692</td>
</tr>
<tr>
<td>pBY200$^d$</td>
<td>pSH79 ($BsaAI$ - $BsaAI$)</td>
<td>pBS (self-ligated)</td>
<td>1384 to 3000</td>
</tr>
<tr>
<td>pBY201</td>
<td>pBY200 ($BsaAI$ - $BglII$)</td>
<td>pBS ($SmaI$ - $BamHI$)</td>
<td>1384 to 2895</td>
</tr>
<tr>
<td>pBY400</td>
<td>pBY201 ($EcoRI$ - $XbaI$)</td>
<td>pUHE1$^*$ ($EcoRI$ - $XbaI$)</td>
<td>1384 to 2895</td>
</tr>
<tr>
<td>pBY415</td>
<td>pBY201 ($EcoRI$ - $NruI$)</td>
<td>pRS415 ($EcoRI$ - $SmaI$)</td>
<td>1384 to 2895</td>
</tr>
<tr>
<td>pBY300</td>
<td>pBY200 ($BsaAI$ - $EcoRV$)</td>
<td>pSP417 ($SmaI$)</td>
<td>1384 to 1885</td>
</tr>
<tr>
<td>pBY301$^t$</td>
<td>pBY300 ($BglII$ - $BamHI$)</td>
<td>pSP417 ($BglII$)</td>
<td>1384 to 1885</td>
</tr>
<tr>
<td>pBY302$^t$</td>
<td>pBY300 ($BglII$ - $BamHI$)</td>
<td>pSP417 ($BglII$)</td>
<td>1885 to 1384</td>
</tr>
<tr>
<td>pBY311</td>
<td>pSH79 ($XbaI$ - $EcoRV$)</td>
<td>pSP417 ($XbaI$ - $SmaI$)</td>
<td>1 to 1885</td>
</tr>
<tr>
<td>pBY312</td>
<td>pBY15 ($XbaI$ - $EcoRV$)</td>
<td>pSP417 ($XbaI$ - $SmaI$)</td>
<td>453 to 1885</td>
</tr>
<tr>
<td>pBY321</td>
<td>253617, 253592 ($EcoRI$ - $XbaI$)</td>
<td>pSP417 ($EcoRI$ - $XbaI$)</td>
<td>366 to 608</td>
</tr>
<tr>
<td>pBY322</td>
<td>253771, 238101 ($EcoRI$ - $XbaI$)</td>
<td>pSP417 ($EcoRI$ - $XbaI$)</td>
<td>591 to 792</td>
</tr>
<tr>
<td>pBY323</td>
<td>253617, 238101 ($EcoRI$ - $XbaI$)</td>
<td>pSP417 ($EcoRI$ - $XbaI$)</td>
<td>367 to 792</td>
</tr>
<tr>
<td>pBY324</td>
<td>253613, 253715 ($EcoRI$ - $XbaI$)</td>
<td>pSP417 ($EcoRI$ - $XbaI$)</td>
<td>992 to 1346</td>
</tr>
<tr>
<td>pBY325</td>
<td>311253, 238253 ($EcoRI$ - $XbaI$)</td>
<td>pSP417 ($EcoRI$ - $XbaI$)</td>
<td>1336 to 1650</td>
</tr>
<tr>
<td>pBY326</td>
<td>253613, 238253 ($EcoRI$ - $XbaI$)</td>
<td>pSP417 ($EcoRI$ - $XbaI$)</td>
<td>992 to 1650</td>
</tr>
<tr>
<td>pBY331</td>
<td>279870, 279484 ($EcoRI$ - $XbaI$)</td>
<td>pSP417 ($EcoRI$ - $XbaI$)</td>
<td>250 to 406</td>
</tr>
<tr>
<td>pBY332</td>
<td>279870, 238101 ($EcoRI$ - $XbaI$)</td>
<td>pSP417 ($EcoRI$ - $XbaI$)</td>
<td>250 to 792</td>
</tr>
<tr>
<td>pBY327</td>
<td>279870, 238253 ($EcoRI$ - $XbaI$)</td>
<td>pSP417 ($EcoRI$ - $XbaI$)</td>
<td>250 to 1650</td>
</tr>
<tr>
<td>pBY328</td>
<td>253771, 238253 ($EcoRI$ - $XbaI$)</td>
<td>pSP417 ($EcoRI$ - $XbaI$)</td>
<td>591 to 1650</td>
</tr>
<tr>
<td>pBY341</td>
<td>301396, 333600 ($EcoRI$ - $XbaI$)</td>
<td>pSP417 ($EcoRI$ - $XbaI$)</td>
<td>582 to 639</td>
</tr>
</tbody>
</table>

104
<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Restriction Sites</th>
<th>Vector</th>
<th>Restriction Sites</th>
<th>Size Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBY342</td>
<td>301411, 333486</td>
<td>(EcoRI-XbaI)</td>
<td>pSP417 (EcoRI-XbaI)</td>
<td>674 to 751</td>
</tr>
<tr>
<td>pBY343</td>
<td>329718, 329865</td>
<td>(EcoRI-XbaI)</td>
<td>pSP417 (EcoRI-XbaI)</td>
<td>1033 to 1092</td>
</tr>
<tr>
<td>pBY344</td>
<td>333605, 336837</td>
<td>(EcoRI-XbaI)</td>
<td>pSP417 (EcoRI-XbaI)</td>
<td>1333 to 1398</td>
</tr>
<tr>
<td>pBY345</td>
<td>253613, Bng4</td>
<td>(EcoRI)</td>
<td>pSP417 (EcoRI-SmaI)</td>
<td>992 to 1134</td>
</tr>
<tr>
<td>pBY346</td>
<td>253613, 311308</td>
<td>(EcoRI)</td>
<td>pSP417 (EcoRI-SmaI)</td>
<td>992 to 1239</td>
</tr>
<tr>
<td>pBY347</td>
<td>311253, 91687</td>
<td>(EcoRI)</td>
<td>pSP417 (EcoRI-SmaI)</td>
<td>1336 to 1474</td>
</tr>
<tr>
<td>pBY348</td>
<td>311253, 69882</td>
<td>(EcoRI)</td>
<td>pSP417 (EcoRI-SmaI)</td>
<td>1336 to 1582</td>
</tr>
</tbody>
</table>

*The template for PCR was plasmid pSH20 (14).*

*Nucleotides of glpEGR cloned; See Fig. 16.*

*The orientation of the glp' EGR genes in this plasmid is the same as that of the glpEGR genes in pSH79.*

*This plasmid was constructed as follows: After digestion of pSH79 with BsaAI (one site is present downstream of glpD in the vector and the other site is at position 1384 of glpEGR), the DNA band containing the glpR gene and the majority of the vector was separated on an agarose gel, purified, and self-ligated.*

*PUHE1 contains the ColE1 origin and confers Amp resistance. The cloned glpR gene is transcribed from the lac promoter, which is inducible by IPTG.*

*pBY301 and pBY302 have the same restriction fragment cloned in opposite orientation in pSP417, with pBY301 having the same orientation as pBY300.*
extension assays, one picomole of $^{32}$P-end-labeled primer (complementary to a region downstream of the possible promoter) was mixed with 5 µg of total RNA in a final volume of 8 µl. The mixture was boiled for 2 min and immediately cooled on ice. The hybridized primer was extended by addition of the four dNTPs (0.7 mM each), reverse transcriptase buffer and 50 U of Moloney murine leukemia virus reverse transcriptase in a total volume of 15 µl, followed by incubation at 42 °C for 30 min. Reactions were stopped by adding 15 µl of stop solution (95% formamide, 20 mM EDTA, 0.05% bromphenol blue and 0.05% xylene cyanol). The samples were analyzed on a 6% polyacrylamide sequencing gel with the appropriate size markers indicated in the figure legend.

**RESULTS**

**Additional promoters after the GlpE start codon are responsible for glpR gene expression.** Although previous experiments suggested a single promoter was responsible for expression of the glpEGR operon, more recent data indicated that glpR might be transcribed independently of glpE. In order to find out if this is the case, plasmids harboring portions of the glpEG region along with the entire glpR gene were constructed using three different cloning vectors (listed in Tables 11 and 12). Plasmids pBY15 and pBY200 contain the glp' EGR or glp'GR region (no glpE promoter) cloned into pBS. Plasmids pSH79 and pNZ80 carried the entire glpEGR sequence in pBS and were used as positive controls. These plasmids were introduced by transformation into NZ45, a glpK-lacZ fusion strain with glpEGR deleted. The glpK-lacZ fusion serves as a
TABLE 12. Repression of Φ(glpg-lacZ) in Δ(glpgEGR) strain NZ45 by plasmids carrying portions of glpgEGR

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Region of glpgEGR(^a)</th>
<th>Genes carried(^b)</th>
<th>β-galactosidase sp act(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBS</td>
<td>none</td>
<td>vector</td>
<td>320</td>
</tr>
<tr>
<td>pSH79</td>
<td>1-3000(^d)</td>
<td>glpD’EGR</td>
<td>1</td>
</tr>
<tr>
<td>pNZ80</td>
<td>272-3000(^d)</td>
<td>glpEGR</td>
<td>1</td>
</tr>
<tr>
<td>pBY15</td>
<td>453-2692(^d)</td>
<td>glp EGR</td>
<td>1</td>
</tr>
<tr>
<td>pBY200</td>
<td>1384-2895(^d)</td>
<td>glp’GR</td>
<td>1</td>
</tr>
<tr>
<td>pRS415</td>
<td>none</td>
<td>vector</td>
<td>600</td>
</tr>
<tr>
<td>pBY415</td>
<td>1384-2898(^e)</td>
<td>glp’GR</td>
<td>15</td>
</tr>
<tr>
<td>pUHE1</td>
<td>none</td>
<td>vector</td>
<td>500</td>
</tr>
<tr>
<td>pBY400</td>
<td>1384-2895(^f)</td>
<td>glp’GR</td>
<td>3</td>
</tr>
</tbody>
</table>

\(^a\) The regions cloned above are indicated using the coordinates of Fig. 16.

\(^b\) A prime located to the left of a gene indicates a portion of the 5' end of the gene is missing. A prime located to the right of a gene indicates a portion of the 3' end of the gene is missing.

\(^c\) Growth medium: A and B salts were supplemented with 2 μg of thiamine per ml and 0.2% each of maltose and Casamino Acids. The enzyme specific activity is in Miller units.

\(^d\) The indicated region (see Fig. 16) was cloned into pBS\(^c\).

\(^e\) The indicated region (see Fig. 16) was cloned into pRS415.

\(^f\) The indicated region (see Fig. 16) was cloned into pUHE1.
reporter for the level of GlpR expressed from the transformed plasmids, where the β-galactosidase activity is expected to be inversely proportional to the level of GlpR. Finding of repressed lacZ expression would indicate the existence of promoter(s) within the cloned region yielding expression of GlpR. The results shown in Table 12 indicate that the glp’GR region produces sufficient GlpR for repression, even though these regions do not contain the glpE promoter (see Table 11). These results suggest that glpG and/or glpR might have their own promoters that express GlpG/GlpR.

It is possible, however, that pBS contains a promoter that transcribes into glpR in pBY15 or pBY200. Plasmid pBY415 was constructed to rule out this possibility. This plasmid was constructed by replacing the lacZ gene of the vector (pRS415) with the glp’GR region (position 1384 to 2895, see Fig. 16). Thus, pBY415 has no lacZ expression. It has 4 strong transcriptional termination sites positioned upstream of glp’GR. The fact that this plasmid still can express sufficient GlpR for repression is the strongest evidence that the glp’GR sequence (1384-2898) cloned into the vector must contain its own promoter(s) for glpR expression.

Plasmid pBY400 contains the same glp’GR sequence as pBY200. This plasmid was used because the expression of genes cloned into pUHE1 are inducible by IPTG. As shown in Table 12, pBY400 repressed NZ45 very well (compare 3 to 500) even without IPTG induction, again suggesting that glpR has its own promoter.

Since neither the strain (NZ45) nor the plasmid (pBY415) contains the glpE or intact glpG gene, it is clear that repression conferred by multicopy GlpR is not dependent
on GlpE/G. The same conclusion was reached by another researcher in this lab (145).

**Contribution of the newly discovered internal promoters to transcription of the glpG and glpR genes.** In order to locate and characterize potential promoters responsible for expression of glpR, different regions upstream of glpEG through the EcoRV site of glpR were fused to the lacZ gene of the vector pSP417 (124). The vector contains four strong transcriptional termination sites upstream of the multiple cloning site and the promoterless lacZ gene. Therefore, the strength of promoters cloned in front of lacZ determines the level of lacZ expression. The results are shown in Table 13. The authors who developed the vector indicated that the weakest promoter should give about 30 fold elevated lacZ expression compared with vector alone (129). Insertion of the DNA fragment encoding the 3' portion of glpG (pBY300) yielded 25 fold higher lacZ expression compared with the vector pSP417, indicating that the glp 'GR' region contains a very weak promoter. Including all of glpG, but only the 3' portion of glpE (pBY312), increased lacZ expression by a factor of 8 compared with pBY300, indicating that the sequence after the glpE promoter but before the end of glpG contains other promoter(s). An additional threefold elevation of lacZ expression was found when the glpE promoter was included (pBY311). These results indicate that the glpEGR transcriptional pattern is complex, and thus warrant further characterization as described below.

**Localization of newly discovered promoters within the glpEGR operon.**

By computer analysis of the glpEGR sequence, several potential -10 and -35 promoter elements that weakly resemble the consensus promoter sequence were found upstream of
TABLE 13. Mapping of promoters within the glpEGR genes

<table>
<thead>
<tr>
<th>Plasmid*(glpEGR region)</th>
<th>Gene(s) carriedb</th>
<th>β-galactosidasec</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSP417 (none)</td>
<td>none</td>
<td>32 ± 1</td>
</tr>
<tr>
<td>pBY300 (1384-1885)</td>
<td>glp 'GR'</td>
<td>760 ± 20</td>
</tr>
<tr>
<td>pBY312 (453-1885)</td>
<td>glp'EGR'</td>
<td>6000 ± 700</td>
</tr>
<tr>
<td>pBY311 (1-1885)</td>
<td>glpD' EGR'</td>
<td>17,000 ± 600</td>
</tr>
</tbody>
</table>

* The host strain used was DH5αF. The regions cloned are indicated using the coordinates of Fig. 16.

b Genetic designations are defined in Table 12.

c The enzyme specific activities are given in Miller units with cells grown in LB medium.
*glpG* and *glpR*. Six different regions of *glpEGR* containing these sequences were amplified by PCR and cloned into pSP417 and analyzed for promoter activity as described above. The results are shown in Fig 17. The region upstream of *glpG* but lacking the *glpE* promoter (pBY322 and pBY323) gave about 100-fold elevated *lacZ* expression compared with the vector. Thus, it is very likely that this region contains at least one promoter. Two separate regions upstream of *glpR* (pBY324 and pBY325) each gave about 25-fold elevated expression compared with the vector. It is possible that each of them contains a weak *glpR* promoter. When the two regions were combined (pBY326), the sum of the activities encoded by pBY324 and pBY325 was obtained, indicating there is no apparent transcriptional termination in the *glpEGR* sequence cloned in pBY326. A small DNA fragment containing most of *glpE* without its upstream promoter (pBY321) expressed only 240 (8 times the vector) Miller units and probably doesn’t contain a promoter (124).

Since *glpEGR* apparently forms a single operon, it would be interesting to know the actual transcriptional expression of each gene. Therefore, different regions of the *glpEGR* operon were cloned in front of *lacZ* using pSP417 as described above. By comparison of *lacZ* expression in Fig.17, the contribution of the internal promoters and the *glpE* promoter previously described (142,146) to *glpG* or *glpR* expression was clarified. pBY331 contains only the *glpE* promoter and expressed about 2500 Miller units when introduced into strain DH5α F+. pBY332 contains the *glpE* promoter and the internal *glpG* promoter and expressed 6500 Miller units, which is about the same as the
<table>
<thead>
<tr>
<th>$P_{glpC}$</th>
<th>$P_{glpE}$</th>
<th>$glpE$</th>
<th>$glpG$</th>
<th>$glpR'$</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBY327</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>9500±70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBY332</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6500±400</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBY331</td>
<td>pBY328</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2500±80</td>
<td>5000±50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBY323</td>
<td>pBY326</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4000±200</td>
<td>1780±30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBY321</td>
<td>pBY324</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>240±3</td>
<td>800±70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBY322</td>
<td>pBY325</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3000±50</td>
<td>750±3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBY341</td>
<td>pBY342</td>
<td>pBY345</td>
<td>pBY347</td>
<td></td>
</tr>
<tr>
<td>335±2</td>
<td>4000±100</td>
<td>96±1</td>
<td>670±8</td>
<td></td>
</tr>
<tr>
<td>pBY346</td>
<td>pBY348</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>114±3</td>
<td>430±10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBY343</td>
<td>pBY344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.52±1</td>
<td>.45±1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 17. Localizations and contributions of the $glpEGR$ promoters. The host strain used was DH5αF’ grown in LB. The regions cloned are indicated by solid lines, with coordinates provided in Table 10. The enzyme activity is expressed in Miller units. DH5αF’ harboring pSP417 expressed 32±3 Miller units.
sum of the activities expressed by pBY331 and pBY322 or pBY323. pBY328 contains 
the glpR and glpG promoters and expressed 5000 Miller units which is close to the sum 
of the activities expressed by pBY326 and pBY322 or pBY323. pBY327 contains all of 
the promoters within the glpEGR region and gave lacZ expression of 9500 Miller units, 
which is close to the sum of all the individual promoter activities. Overall there seems to 
be no transcriptional termination occurring within the glpEGR region. Thus it seems that 
the relative levels of transcription of the glpEGR genes is glpR > glpG > glpE.

Primer extension analysis. Since computer analysis of the glpEGR sequence 
revealed no sequence closely matching the bacterial consensus promoter elements, 
primer extension analysis was necessary in order to attempt to identify the 5' ends of 
mRNAs and thereby locate the the possible promoter regions. The strategy is outlined in 
Fig. 18. To localize the glpG promoter(s), primers 11360 (708 to 682, at the 3' end of 
glpE, Table 10) and 198093 (complementary to the 5'-end of lacZ, Table 2) were used in 
primer extension assays containing RNA purified from strain DH5α F'(pBY322). Using 
primer 11360, one band was found that corresponds to the base C at position 630 of glpE 
(Fig. 16) (data not shown). Primer extension using the same RNA with primer 198093 
failed to yield any bands.

Primer 238101 (806 to 776 in glpG, Table 9) was hybridized to RNA purified 
from strain DH5α F'(pBY328) in a primer extension assay. As shown in Fig. 19, four 
bands were seen (positions 725 and 726 (GG) and positions 741 and 742 (AC), Fig. 16). 
There are no good consensus -10 and -35 sequences upstream of these positions. One
Fig. 18. Primer extension analysis of the *glpEGR* promoters. The single lines represent the sequences in *glpEGR* and the black boxes represent the vector sequences of pSP417 (*lacZ*). The oligonucleotides used for primer extension are indicated as (←).
Fig. 19. Primer extension analysis of the glpG promoter. Total RNA was purified from DH5α F' (pBY323). ³²P-end-labeled primer 11360 was used in reactions. The size markers (ACGT) are sequencing standards using end-labeled primer 11360 with template pBY323.
possibility is that these bands represent the 5' ends of RNAs resulting from digestion by RNase.

To locate the possible glpR promoter(s) within glpG, primers Bing4 (1134 to 1114, Table 10), 311308 (1239 to 1219, Table 10) and 198093 (complementary to the 5'-end of lacZ) were used in primer extension assays using RNA purified from strain DH5αF'(pBY324) and primers 91687 (1474 to 1455, Table 10), primers 69882 (1582 to 1554, Table 10) and 198093 (complementary to the 5'-end of lacZ) were used in similar assays using RNA purified from the strain DH5αF'(pBY325). These primers covered all the possible 5'-ends of the mRNAs initiated upstream of glpR. No strong and clear bands were found. Some very weak bands were found after several week's exposure and these bands were between positions 1034 to 1093 or 1336 to 1399 in glpEGR (data not shown).

**Further localization of the glpGR promoters.** According to the results of primer extension analysis of the glpG promoter, there is one band that corresponds to position 630 (Fig. 16), and four bands that correspond to positions 725/726 and 741/742 (Fig. 16). However, no sequences resembling the consensus promoter elements were found upstream the sequences represented by these bands. In order to find out if these bands actually represent the 5' ends of mRNAs expressed from promoters, plasmids pBY341 and pBY342 were constructed by cloning relatively small DNA fragments that cover these two possible transcriptional start regions into pSP417 (Table 11). From the levels of β-galactosidase expression, it seems that the insert of pBY342 has a strong
promoter as expected (expression of 4000 Miller units, Fig. 17), confirming that these
double bands at 725/726 or 741/742 are derived from the expression of a strong
promoter. However, pBY341 expressed only 335 Miller units (Fig. 17), only tenfold
higher than the expression of the vector, suggesting it doesn’t have a significant promoter
(129). Therefore, the band corresponding to position 630 may have been generated by
processing of a longer mRNA initiated from the upstream glpE promoter. pBY343 and
pBY344 contain the sequence of the observed primer extension bands at positions 1033
to 1092, and 1333 to 1398 (Fig. 16), respectively. However, they have no promoter
activity with low β-galactosidase activity of about 50 Miller units each (Fig. 17).

In order to localize the glpR promoters, pBY345 and pBY346 were constructed
and served as deletion derivatives of pBY324; pBY347 and pBY348 were constructed and
served as deletion derivatives of pBY325. From the results shown in Fig. 17, pBY345
and pBY346 contained insignificant promoter activity. So if any promoter exists in
pBY324, it must be located between positions 1239 and 1346 (Figs. 16 and 17). pBY347
and pBY348 all expressed weak promoter activity (670 and 430, respectively). Thus the
the possible promoter must be between positions 1336 to 1474 (Figs. 16 and 17). Since
the expression of these two promoters is very low, primer extension analysis did not
reveal any 5'-ends of mRNAs in that region (data not shown). The reason that pBY348
has lower expression compared with pBY347 is unknown.

*glpR is not autoregulated.* A previous study (143) suggested that a possible
operator located at position 390 to 409 (Fig. 16) might negatively control glpE
expression by binding of GlpR. However, in my work, GlpR was unable to protect this putative operator site, as determined by DNase I footprinting (data not shown). Computer analysis of the glpEGR sequence revealed two other possible operator sites, located at positions 1480 to 1499 and 1554 to 1573 (Fig. 16). It is possible that these two operators could cooperatively bind GlpR and negatively control the expression of the glpR promoter. The distance between these two operators is about 70 bp, suitable for cooperative binding of GlpR via a DNA loop (Part I). To reinvestigate the possibility that glpR is autoregulated, pBY300, pBY325 and pBY327 were introduced by transformation into strain TL73 (glpR2 relA1) harboring either pSH58 (encoding GlpR (40)) or pACYC184 (vector control). If there is control by GlpR, the expression of lacZ in the strain with pSH58 should be lower than that in the strain with pACYC184. pBY300 contains one glpR promoter just upstream of the GlpR start codon, pBY325 contains both of the weak glpR promoters, and pBY327 contains all of the glpEGR promoters (see Fig. 16). As the results shown in Table 14 indicate, none of the promoters were subject to negative control by GlpR. Therefore, the glpEGR genes are not controlled by the glp repressor. A glpD-lacZ fusion plasmid (pDA904, (143)) containing the glpD promoter and operators was used as the positive control in a similar assay with pSH58 or pACYC184. Strong repression (70-fold) was observed in TL73 with pSH58 but not pACYC184, confirming the efficacy of repressor produced by pSH58.
TABLE 14. Expression of \( \text{glpEGR} \) is not controlled by GlpR

<table>
<thead>
<tr>
<th>Plasmids</th>
<th>Promoter(^a)</th>
<th>( \beta )-Galactosidase specific activity in TL73/pSH58</th>
<th>TL73/pACYC184</th>
<th>Repression</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSP417</td>
<td>none</td>
<td>35 ± 1</td>
<td>33 ± 2</td>
<td>1</td>
</tr>
<tr>
<td>pBY300</td>
<td>( P_R )</td>
<td>1000 ± 100</td>
<td>1110 ± 10</td>
<td>1</td>
</tr>
<tr>
<td>pBY325</td>
<td>( P_C P_R )</td>
<td>1100 ± 50</td>
<td>1180 ± 9</td>
<td>1</td>
</tr>
<tr>
<td>pBY327</td>
<td>( P_E P_C P_R )</td>
<td>9300 ± 50</td>
<td>9690 ± 30</td>
<td>1</td>
</tr>
<tr>
<td>pDA904(^b)</td>
<td>( P_D )</td>
<td>5</td>
<td>350</td>
<td>70</td>
</tr>
</tbody>
</table>

\(^a\)The \( \text{glp} \) promoters carried by each plasmid were indicated.

\(^b\)pDA904 carries a \( \text{glpD-lacZ} \) transcriptional fusion (143).
DISCUSSION

It was previously believed that the glpE, glpG and glpR genes were cotranscribed from a single promoter upstream of glpE. My discovery of the promoters responsible for the expression of the glpG and glpR genes provides new insight regarding the transcriptional organization of the glpEGR operon. It seems that the relative level of the transcriptional expression of the glpEGR genes is glpR > glpG > glpE. Since the translational efficiency of each of the glpEGR genes is unknown, the level of the expression of each protein is undetermined. However, since the stop codon of glpG and the start codon of glpR are very close together (Fig. 16), translational coupling of glpG and glpR is possible. Since the functions of GlpE and GlpG are still unknown, the advantage or rationale of this type of expression of GlpR from multiple promoters is unknown.

The expression of glpR is not autoregulated as is the case with many other prokaryotic repressors. In case of positive regulation by CRP, the potential CRP site at position 719 to 740 seems to overlap the regions corresponding to the 5' ends identified by primer extension (positions 725/726 and 741/742) of the glpG promoter. If the bands at positions 725/726 correspond to the 5'-ends of the mRNA initiated at the glpG promoter, this CRP site would overlap the -35 and -10 region of the promoter. If the bands corresponding to positions 725/726 are the results of the RNase cleavage of a longer mRNA, the actual essential promoter elements will be located further upstream. Therefore this putative CRP site is not likely to contribute to the activation of the glpG
promoter. The other potential CRP site at position 1254 to 1275 overlaps or is upstream of one of the weak glpR promoters (between 1336 to 1476). However, the very low expression of this promoter in the presence of CRP raises the question of its physiological relevance.

Recently it was reported that FIS represses the expression of glpR (84). This was concluded after comparing the levels of glpR mRNA expression in a Δfis and an isogenic wild type strain. Several FIS binding sites are found overlapping or close to the CRP site upstream of the glpE promoter (Fig. 16). If binding of FIS competes with the binding of CRP-cAMP at the CRP site, FIS would then negatively control the glpR expression. Future work needs to be done to clarify the mechanism.

This work is not complete as the exact transcriptional start point of the glpG promoter is not defined. The region containing glpG promoter in pBY342 will be further characterized. For example, a primer complementary to the sequence covering the four bands identified by previous primer extension analysis could be used to perform a similar primer extension analysis using a RNA source from plasmid pBY342, assuming that the previous bands resulted from RNAse cleavage of an uncut nascent RNA. Hopefully this would result in identification of the actual 5' end of the mRNA initiated from the glpG promoter. If this is successful, the possible -10 and -35 promoter elements could be suggested. Further subcloning of a shortened DNA fragment of pBY342 with only the -10 or -35 promoter element into pSP417 should result in no promoter activity. This will confirm the transcriptional initiation site revealed by primer extension assay. Since the
$glpR$ promoters are so weak, the identification of the actual transcriptional initiation sites seems unlikely by using currently available molecular techniques.
LITERATURE CITED


64. van Rooijen, R.J., K.J. Decherling, C. Niek, J. Wilmink, and W.M. de Vos. 1993. Lysines 72, 80 and 213 and aspartic acid 210 of the *Lactococcus lactis* LacR repressor are involved in the response to the inducer tagatose-6-phosphate leading to induction of lac operon expression. *Protein Engineering* 6:201-206.


70. Reznikoff, W.S., R.B. Winter, and C.K. Hurley. 1974. The location of the


    USA 82:3776-3780.

114. Thompson, J.F., L. Moitoso de Vargas, C. Koch, R. Kahmann, and A. Landy.
    of a new component in the lambda site-specific recombination pathway. Cell
    50:901-908.

    174:8043-8056.

    subject to stringent control and autoregulation. EMBO J. 11:1075-1083.

    requires two 26 bp recombination sites and a 60 bp recombinational enhancer.
    Cell 41:781-791.

    bacteriophage Mu DNA is stimulated by a site within the invertase gene and a
    host factor. Cell 41:771-780.


120. Gille, H., J.B. Egan, A. Roth, and W. Messer. 1991. The FIS protein binds and
    bends the origin of chromosomal replication, oriC, of Escherichia coli. Nucleic
    Acids Res. 19:4167-4172.

    16:356.

    Current Inovations. H.G. Griffin and A.M. Griffin, editors. CRC Press, Inc. Boca
    Raton, FL. 69-83.


CURRICULUM VITAE

Name: Bing Yang
Home Address: 307 Hunt Club Road #6400F, Blacksburg, VA 24060

Date of Birth: March 28, 1969

Education:

   Ph. D. 1993-1996 Virginia Polytechnic Institute and State University
          Blacksburg, VA 24061-0308
          Major: Biochemistry

   B. A. 1986-1990 FuDan University
          Shanghai, P.R. of China
          Major: Genetics and Genetic Engineering

Honor:
   Wu Travel Fellowship, Department of Biochemistry, Virginia Tech

Publication:


Abstract:


Bing Yang