JOURNAL OF BACTERIOLOGY, May 1990, p. 2642–2649
0021-9193/90/052642-08$02.00/0

Vol. 172, No. 5

Regulation of the Glyoxylate Bypass Operon: Cloning and Characterization of *iclR*

ALDEN SUNNARBORG, DAVID KLUMPP,† TAEOWAN CHUNG,‡ AND DAVID C. LAPORTE*

Department of Biochemistry, University of Minnesota, Minneapolis, Minnesota 55455

Received 20 November 1989/Accepted 15 February 1990

In *Escherichia coli*, expression of the glyoxylate bypass operon appears to be controlled, in part, by the product of *iclR*$^+$. Mutations in *iclR* have been found to yield constitutive expression of this operon, suggesting that *iclR*$^+$ encodes a repressor protein. We have cloned *iclR*$^+$ by taking advantage of its tight genetic linkage with the glyoxylate bypass operon. The clone complemented a mutant allele of *iclR* in *trans*, restoring an inducible phenotype for this operon. Deletion analysis identified a region of ca. 900 base pairs that was necessary and sufficient for complementation. The nucleotide sequence of the insert was then determined. Translation of this sequence revealed an open reading frame capable of encoding a protein with M_r 29,741 preceded by a potential Shine-Dalgarno ribosome-binding site. The deduced amino acid sequence includes a region at the amino terminus that may form a helix-turn-helix motif, a structure found in many DNA-binding domains.

When *Escherichia coli* adapts to growth on acetate, it induces expression of the enzymes of the glyoxylate bypass: isocitrate lyase and malate synthase (17, 19). This bypass is essential for growth on acetate, since it yields C_4 acids while avoiding the net loss of the acetate carbons as carbon dioxide in the Krebs cycle (Fig. 1). After induction, the flow of isocitrate through the glyoxylate bypass is regulated, in part, by the phosphorylation of isocitrate dehydrogenase (IDH), the Krebs cycle enzyme that competes with isocitrate lyase (8, 12, 24). During growth on acetate, ca. 70% of the IDH is maintained in the inactive phosphorylated form (22, 23, 32), reducing the activity of this enzyme and so forcing isocitrate through the bypass (24, 32). The phosphorylation and dephosphorylation of IDH are catalyzed by a single bifunctional enzyme, IDH kinase/phosphatase (20, 21).

The metabolic and regulatory proteins of the glyoxylate bypass reside in the same operon, which maps at 91 min on the *E. coli* chromosome (4, 5, 23, 26). Isocitrate lyase and malate synthase are encoded by *aceA* and *aceB*; IDH kinase/phosphatase is encoded by aceK. The organization of this operon is:

$$\frac{aceB\ aceA\ aceK}{P}$$

where P indicates the position of the promoter. S1 mapping has indicated that this operon employs a single promoter during growth on acetate (5).

The glyoxylate bypass operon is expressed only during growth on acetate and is repressed if any preferred carbon source (e.g., glucose or pyruvate) is simultaneously present. Expression of this operon appears to be controlled, at least in part, by the product of *iclR*$^+$. Transduction experiments using bacteriophage P1 have demonstrated that *iclR*$^+$ resides downstream of the glyoxylate bypass operon. Although these experiments demonstrated a tight linkage between the glyoxylate bypass operon and the gene that controls its expression, the precise distance between these loci remains uncertain. Recessive mutations in *iclR* have been isolated that yield constitutive expression of the glyoxylate bypass operon. Analysis of these mutations has led to the suggestion that *iclR*$^+$ encodes a repressor protein which acts directly on expression of this operon (4, 17, 18, 26). In this report, we describe the cloning and initial characterization of *iclR*$^+$.

MATERIALS AND METHODS

Materials. Restriction enzymes, other DNA modification enzymes, and linkers were purchased from Bethesda Research Laboratories Inc., or New England BioLabs, Inc. Sequencing primers were synthesized by National Biosciences, Hamel, Minn., or by the Microchemical Facility of the University of Minnesota. Radioactive nucleotides were products of Dupont, NEN Research Products. All other reagents were the purest grades available.

Growth media. The minimal medium used was the morpholinepropanesulfonic acid (MOPS)-based medium described by Neidhardt et al. (31) containing the appropriate amino acids. The carbon source was either 2% acetate or 2% glucose, as indicated. L broth contained 1% Tryptone (Difco Laboratories), 1% NaCl, and 0.5% yeast extract. When indicated, L broth was supplemented with 2% glucose. When required, ampicillin (200 μg/ml) tetracycline (12.5 μg/ml), or kanamycin (50 μg/ml) was included in the growth media.

Measurement of enzymatic activities. Cultures were grown at 37°C in a gyratory incubator to mid-log phase and were then harvested by centrifugation at 4,000 × *g* for 10 min. The cells were suspended in 10 ml of extraction buffer (25 mM *N*-morpholinepropanesulfonate [pH 7.5], 2 mM β-mercaptoethanol, 1 mM EDTA) and then pelleted again by centrifugation. Cell pellets were stored at −80°C. For assay, the samples were thawed, suspended in 5 ml of extraction buffer, and disrupted by sonication. Cellular debris was removed by centrifuged at 22,000 × *g* for 20 min, and the samples were assayed for IDH phosphatase activity. Samples derived from cultures harboring plasmid were also assayed for β-lactamase activity to ensure that the plasmid had not been lost during growth.

The activity of IDH phosphatase was measured by monitoring the release of [^{32}P]phosphate from [^{32}P]phospho-IDH,

* Corresponding author.
† Present address: Northwestern University, Evanston, IL 60602.
‡ Present address: Department of Biochemistry, College of Life Sciences, Yeungnam University, Taegu, South Korea 632.

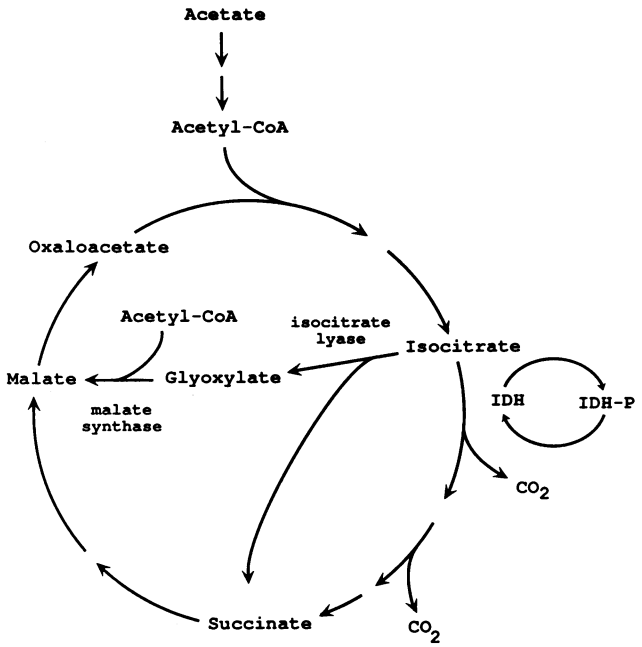

FIG. 1. Krebs cycle and glyoxylate bypass. The glyoxylate bypass is composed of isocitrate lyase and malate synthase. The phosphorylated and dephosphorylated forms of isocitrate dehydrogenase are indicated as IDH-P and IDH, respectively.

as described previously (21). The standard reaction conditions were 25 mM MOPS (pH 7.5), 0.1 μM [^{32}P]phospho-IDH (ca. 30,000 cpm), 1 mM ATP, 5 mM $MgCl_2$, 100 mM NaCl, 2 mM dithiothreitol, 0.5 mM EDTA, 5 mM 3-phosphoglycerate, 1 mM D,L-isocitrate, and 2 mg of bovine serum albumin per ml. One unit of IDH phosphatase activity is defined as the amount of enzyme required to achieve 50% dephosphorylation of the phospho-IDH in 1 min.

β-Lactamase was assayed by the method of O'Callaghan et al. (34), using cephalosporin as the substrate. The reaction mixture contained 100 μM cephalosporin and 100 mM sodium phosphate (pH 7.0). The reaction was performed at 37°C, and its progress was monitored by observing the decrease in A_{252}. The observed activity was corrected for protein concentration.

Protein concentration was determined by the method of Lowry et al. (25), using bovine serum albumin as the standard.

Recombinant DNA techniques. Except where indicated, manipulations of plasmids or phage employed standard techniques (2, 27).

Construction of an *aceB-lacZ* operon fusion. Our initial experiments used an *aceB-lacZ* operon fusion to score for the *iclR* genotype. This fusion was constructed by using the plasmid and lambda phage *lacZ* fusion vectors described by Simons et al. (42). To prevent translational readthrough from interfering with the expression of *lacZ*, a linker with the sequence CTAGCTAGCTAG was inserted in the *Bam*HI site of pRS551, creating plasmid pCL551. Plasmid pCL1000 (which encodes the glyoxylate bypass operon) (5) was digested with *Pvu*II. After the attachment of *Eco*RI linkers, the appropriate fragment was ligated with *Eco*RI-cleaved pCL551, and the sample was transformed into strain W4680. DNA was isolated from transformants that expressed β-galactosidase activity, and the identity of the resulting plasmid, designated pKL4, was confirmed by restriction mapping. The insert from pKL4 was transferred to lambda phage λRS45 by recombination in vivo (42). The resulting phage is designated λKL4. Lysogens harboring this phage can be isolated by selecting for kanamycin resistance, a marker that is transferred from the plasmid to the phage during recombination.

TABLE 1. Principle bacterial strains

Strain	Genotype	Source or reference
W4680	Δ*lacZ39 rpsL melB*	CGSC[a]
TST1	*malE*::Tn*10 araD* Δ(*argF-lac*) *flbB ptsF relA rpsL deoC*	CGSC
PLK831	*iclR7 gal trpE pyrF fnr rpsL trpR*	CGSC
MM294Δ	*recA*	D. E. Koshland
JC10240	*recA39 srl-300*::Tn*10 thr relA ilv spoT thi rpsE*	7
SL1025	Δ*lacZ39 malE*::Tn*10 rpsL melB*	This study
SL1026	Δ*lacZ39 iclR7 malE*::Tn*10 rpsL melB*	This study
SL1027	Δ*lacZ39 recA56 srl-300*::Tn*10 malE rpsL melB*	This study
SL1028	Δ*lacZ39 recA56 iclR7 srl-300*:: Tn*10 malE rpsL melB*	This study

[a] CGSC, *E. coli* Genetic Stock Center.

Bacterial strains. The strains of *E. coli* used are listed in Table 1.

To construct congenic wild-type and *iclR* strains, a Tn*10* transposon was introduced near the *iclR7* allele of strain PLK831 by P1 transduction from strain TST1, with selection for tetracycline resistance. The *iclR7* allele was then transferred to W4680, which had been lysogenized with λKL4, by P1 transduction, with selection for tetracycline resistance. The transduction mixture was plated on L broth supplemented with 12.5 μg of tetracycline per ml, 2% glucose, and 60 μg of X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) per ml. Strains that carried *iclR7* were identified as blue colonies on this medium. The prophage was then eliminated by growth in the absence of kanamycin, followed by testing for resistance to that drug. The genotypes were confirmed by assay for IDH phosphatase in cultures grown on L broth-glucose. Strains obtained by this procedure are designated SL1025 (*iclR*$^+$) and SL1026 (*iclR7*). Tetracycline-sensitive derivatives of these strains were isolated by growth on Bochner plates (3). The *recA52* allele was then introduced into these strains by P1 transduction from strain JC10240, with selection for tetracycline resistance. The *recA* phenotype was scored as enhanced sensitivity to UV light. The resulting strains are designated SL1027 (*iclR*$^+$ *recA*) and SL1028 (*iclR7 recA*).

Nucleotide sequencing. The initial nucleotide sequence analysis was performed by the chemical cleavage method of Maxam and Gilbert (29). This partial sequence was then used to design primers so that the remainder of the sequence could be determined by a modification (2) of the dideoxy-chain termination method of Sanger et al. (38).

The DNA sequencing data were compiled and analyzed by using programs developed by IntelliGenetics, Inc.

Plasmids encoding *iclR*. Plasmid pBN1, which was obtained from William Nunn, was described by us previously (5). This plasmid carries the glyoxylate bypass operon and about 15 kilobase pairs (kb) of DNA downstream of this operon and includes *iclR*$^+$ (see Results). Plasmid pKL5, which also encodes *iclR*$^+$, was constructed by subcloning a *Cla*I- *Hin*dIII fragment from pBN1 into pBR322.

TABLE 2. Complementation of *iclR7*

Strain	Relevant genotype	Plasmid	IDH phosphatase activity (mU/mg) during growth on:	
			Glucose	Acetate
SL1027 / SL1028	*iclR*$^+$	None	0.08	3.4
	iclR7	None	1.19	9.0
		pBR322	1.97	8.5
		pKL5	0.03	6.2

The location of *iclR* within the plasmid insert was determined by deletion analysis. Deletions were introduced at the 5′ end of *iclR* by taking advantage of the *Xmn*I and *Mlu*I sites. Deletions were generated at the 3′ end of this gene by digestion with BAL 31. Plasmid pKL5 was cleaved with *Hin*dIII and was then subjected to digestion with BAL 31. Samples were removed at intervals, and the reaction was stopped with EGTA. After this digestion, the pooled samples were extracted with phenol and DNA was isolated by ethanol precipitation. The ends were repaired with the Klenow fragment of DNA polymerase I, and *Hin*dIII linkers were attached. The DNA was then cleaved with *Hin*dIII and ligated. After transformation into MM294Δ, plasmid DNA was isolated from individual colonies. The extent of each deletion was determined by the nucleotide sequencing method of Maxam and Gilbert (29), using an appropriate DNA fragment labeled at the *Hin*dIII site.

RESULTS

Cloning of *iclR*. For routine scoring of the *iclR* genotype, we constructed an *aceB-lacZ* operon fusion in a lambda vector (see Materials and Methods). Strains lysogenized with this construct, designated λKL4, yield blue colonies on L broth-glucose supplemented with X-Gal (repressing conditions) if they are *iclR* and white colonies if they are wild type (not shown).

To clone *iclR*, we took advantage of the genetic linkage between this gene and the glyoxylate bypass operon. We have previously described the isolation and characterization of clones of this operon (5). One of these clones, pBN1, included ca. 15 kb of DNA downstream of the glyoxylate bypass operon, raising the possibility that it included *iclR*$^+$. This plasmid was found to complement the *iclR* mutation in strain SL1026 (Δ*lacZ iclR7*) harboring λKL4. Preliminary deletion mapping indicated that a 2.0-kb *Cla*I-*Hin*dIII fragment was both necessary and sufficient for complementation of *iclR7* (not shown). This fragment was subcloned into pBR322, creating plasmid pKL5.

To confirm that plasmid pKL5 included *iclR*$^+$, we tested its ability to complement an *iclR* mutation. Strain SL1028, which carries *iclR7*, exhibited elevated expression of IDH phosphatase during growth under repressing conditions: glucose minimal medium (Table 2). Although IDH phosphatase activity was clearly elevated under these conditions, expression was not fully constitutive. Growth on acetate (inducing conditions) yielded a further increase of ca. sixfold in IDH phosphatase activity. Plasmid pKL5 complemented this *iclR* mutation, reducing expression of IDH phosphatase to the wild-type level during growth on glucose. Expression of IDH phosphatase activity could be fully induced by growth of SL1028 harboring pKL5 on acetate medium. The ability of plasmid of pKL5 to complement *iclR7* indicates that it does indeed encode *iclR*$^+$. Identification of the cloned gene as *iclR* is also supported by the fact that it occupied the same position relative to the glyoxylate bypass operon on plasmid pBN1 (the original isolate) as it does on the chromosome.

Deletion mapping of *iclR*. To locate *iclR*$^+$ within pKL5, a series of nested deletions was introduced at either end of the insert. The resulting plasmids were then introduced into strain SL1028 to test their ability to complement *iclR7*. A deletion removing the first 430 base pairs (bp) of the insert retained *iclR* function, as indicated by repression of IDH phosphatase expression during growth on glucose (Fig. 2). In contrast, deletion of the first 755 bp eliminated complementation. Similar results identified the right end of *iclR* between bp 1400 and 1560.

One of the deletion mutants, carried by plasmid pKL51, consistently caused elevated levels of IDH phosphatase expression in the *iclR7* host strain. It seems likely that this

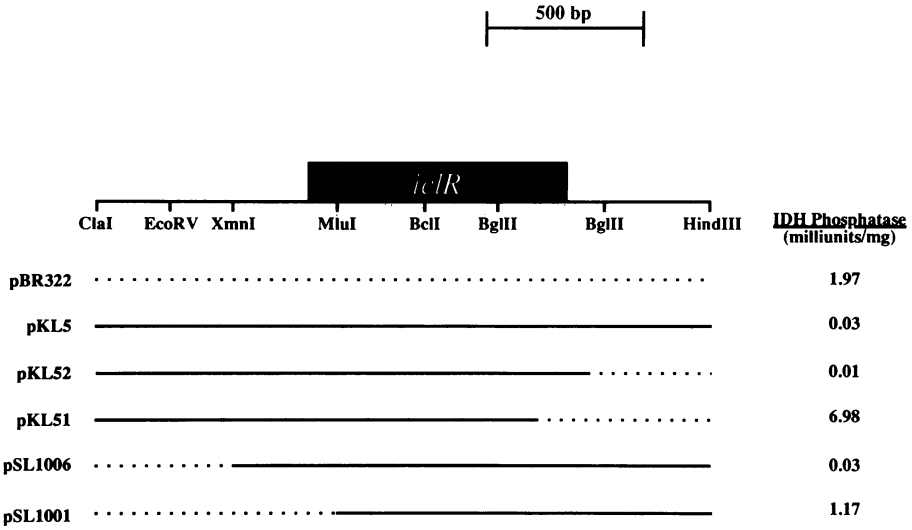

FIG. 2. Deletion mapping of *iclR*. The sequences present in each plasmid are indicated by solid lines. The box indicates the position of the open reading frame identified in Fig. 4.

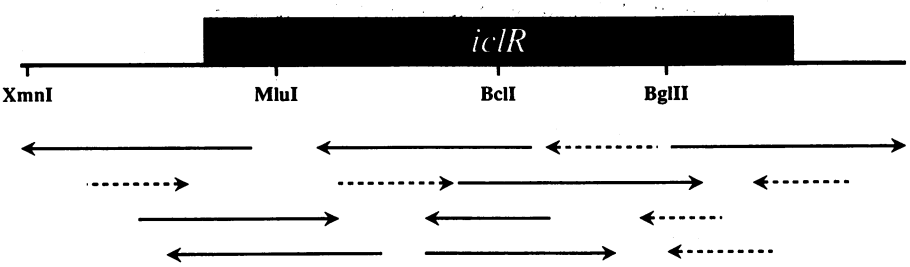

FIG. 3. Sequencing strategy for *iclR*. The box represents the open reading frame of *iclR*. The nucleotide sequence was determined by the dideoxy method of Sanger et al. (2, 38) (solid arrows) and the chemical cleavage method of Maxam and Gilbert (29) (dashed arrows).

deletion, which removed ca. 32 codons from the 3′ end of *iclR* (see below), does not represent a loss-of-function mutation. This phenomenon is discussed below.

Nucleotide sequence of *iclR*. The sequencing strategy for *iclR*$^+$ is presented in Fig. 3. Translation of the nucleotide sequence revealed an open reading frame between nucleotides 202 and 1074 (Fig. 4). A potential AUG initiation codon occurs near the beginning of this reading frame and is

```
    ACC AGA ATA CGT TCA TTT AAC TGC GCA CGC AGT TGT TCC ACT TTG CTG CTC ACA   54

    C TTGCTC CC GAC ACG CTC AAC CCA GA TTTAAT A AAA ATT CAA CAA ACC ATA CTG  108
      -35                              -10

    GCA TAA ACG CAT CTG TGG TAA AAG CGA CCA CCA CGC AAC ATG AGA TTT GTT CAA  162

    CAT TAA CTC ATC GGA TCA GTT CAG TAA CTA TTG CAT TAG CTA ACA ATA AAA ATG  216

    AAA ATG ATT TCC ACG ATA CAG AAA A AGGAG ACT GTC ATG GTC GCA CCC ATT CCC  270
                                   S.D.        MET Val Ala Pro Ile Pro

    GCG AAA CGC GGC AGA AAA CCC GCC GTT GCC ACC GCA CCA GCG ACT GGA CAG GTT  324
  7 Ala Lys Arg Gly Arg Lys Pro Ala Val Ala Thr Ala Pro Ala Thr Gly Gln Val

    CAG TCT TTA ACG CGT GGC CTG AAA TTA CTG GAG TGG ATT GCC GAA TCC AAT GGC  378
 25 Gln Ser Leu Thr Arg Gly Leu Lys Leu Leu Glu Trp Ile Ala Glu Ser Asn Gly

    AGT GTG GCA CTC ACG GAA CTG GCG CAA CAA GCC GGG TTA CCC AAT TCC ACG ACC  432
 43 Ser Val Ala Leu Thr Glu Leu Ala Gln Gln Ala Gly Leu Pro Asn Ser Thr Thr

    CAC CGC CTG CTA ACC ACG ATG CAA CAG CAG GGT TTC GTG CGT CAG GTT GGC GAA  486
 61 His Arg Leu Leu Thr Thr MET Gln Gln Gln Gly Phe Val Arg Gln Val Gly Glu

    CTG GGA CAT TGG GCA ATC GGC GCA CAT GCC TTT ATG GTC GGC AGC AGC TTT CTC  540
 79 Leu Gly His Trp Ala Ile Gly Ala His Ala Phe MET Val Gly Ser Ser Phe Leu

    CAG AGC CGT AAT TTG TTA GCG ATT GTT CAC CCT ATC CTG CGC AAT CTA ATG GAA  594
 97 Gln Ser Arg Asn Leu Leu Ala Ile Val His Pro Ile Leu Arg Asn Leu MET Glu

    GAG TCT GGC GAA ACG GTC AAT ATG GCG GTG CTT GAT CAA AGC GAT CAC GAA GCG  648
115 Glu Ser Gly Glu Thr Val Asn MET Ala Val Leu Asp Gln Ser Asp His Glu Ala

    ATT ATT ATC GAC CAG GTA CAG TGT ACG CAT CTG ATG CGA ATG TCC GCG CCT ATC  702
133 Ile Ile Ile Asp Gln Val Gln Cys Thr His Leu MET Arg MET Ser Ala Pro Ile

    GGC GGT AAA TTG CCG ATG CAC GCT TCC GGT GCG GGT AAA GCC TTT TTA GCC CAA  756
151 Gly Gly Lys Leu Pro MET His Ala Ser Gly Ala Gly Lys Ala Phe Leu Ala Gln

    CTG AGC GAA GAA CAG GTG ACG AAG CTG CTG CAC CGC AAA GGG TTA CAT GCC TAT  810
169 Leu Ser Glu Glu Gln Val Thr Lys Leu Leu His Arg Lys Gly Leu His Ala Tyr

    ACC CAC GCA ACG CTG GTG TCT CCT GTG CAT TTA AAA GAA GAT CTC GCC CAA ACG  864
187 Thr His Ala Thr Leu Val Ser Pro Val His Leu Lys Glu Asp Leu Ala Gln Thr

    CGC AAA CGG GGT TAT TCA TTT GAC GAT GAG GAA CAT GCA CTG GGG CTA CGT TGC  918
205 Arg Lys Arg Gly Tyr Ser Phe Asp Asp Glu Glu His Ala Leu Gly Leu Arg Cys

    CTT GCA GCG TGT ATT TTC GAT GAG CAC CGT GAA CCG TTT GCC GCA ATT TCT ATT  972
223 Leu Ala Ala Cys Ile Phe Asp Glu His Arg Glu Pro Phe Ala Ala Ile Ser Ile

    TCC GGA CCG ATT TCA CGT ATT ACC GAT GAC CGC GTG ACC GAG TTT GGC GCG ATG 1026
241 Ser Gly Pro Ile Ser Arg Ile Thr Asp Asp Arg Val Thr Glu Phe Gly Ala MET

    GTG ATT AAA GCG GCG AAG GAA GTG ACG CTG GCG TAC GGT GGA ATG CGC TGA CTT 1080
259 Val Ile Lys Ala Ala Lys Glu Val Thr Leu Ala Tyr Gly Gly MET Arg

    TTT CTG GCG GGC AGA GGC AAT ATT CTG CCC ATC ATA CCT GAG TGG CAA TAG AAT 1134

    AAG GGT GTC TGT TAA TCG CAT TGA CGC CAA AA
```

FIG. 4. Nucleotide sequence of *iclR* (GenBank accession no. M31761) and the deduced amino acid sequence of its product. S.D., Potential Shine-Dalgarno sequence (underlined). Sequences similar to the consensus for *E. coli* promoters are also indicated.

TABLE 3. Codon usage for *iclR*

Amino acid	Codon	*iclR*		Highly expressed genes (%)[a]	Amino acid	Codon	*iclR*		Highly expressed genes (%)
		Total	%				Total	%	
Phe	UUU	6	75	27	Tyr	UAU	2	67	28
	UUC	2	25	73		UAC	1	33	72
Leu	UUA	7	23	3	Ter	UAA	0		
	UUG	2	7	5		UAG	1		
	CUU	2	7	5		UGA	0		
	CUC	3	10	6					
	CUA	3	10	1	His	CAU	6	46	26
	CUG	13	43	81		CAC	7	54	74
Ile	AUU	11	73	26	Gln	CAA	6	40	15
	AUC	4	27	74		CAG	9	60	85
	AUA	0	0	0.3					
					Asn	AAU	5	100	12
Met	AUG	10				AAC	0	0	88
Val	GUU	4	24	46	Lys	AAA	9	82	76
	GUC	3	18	9		AAG	2	18	24
	GUA	1	6	25					
	GUG	9	53	21	Asp	GAU	6	67	40
						GAC	3	33	60
Ser	UCU	4	24	34					
	UCC	5	29	30	Glu	GAA	12	71	75
	UCA	2	12	4		GAG	5	29	25
	UCG	0	0	5					
	AGU	1	6	4	Cys	UGU	2	67	37
	AGC	5	29	22		UGC	1	33	63
Pro	CCU	3	27	10	Trp	UGG	2		
	CCC	4	36	1.4					
	CCA	1	9	12	Arg	CGU	6	38	68
	CCG	3	27	77		CGC	7	44	31
						CGA	1	6	0.7
Thr	ACU	1	6	34		CGG	1	6	0.4
	ACC	6	35	54		AGA	1	6	0.3
	ACA	0	0	3		AGG	0	0	0.3
	ACG	10	59	9					
					Gly	GGU	6	27	56
Ala	GCU	1	3	34		GGC	9	41	39
	GCC	10	30	12		GGA	4	18	1.4
	GCA	9	27	25		GGG	3	14	3
	GCG	13	39	29					

[a] Relative codon usage in highly expressed genes of *E. coli*, determined by averaging the values reported by Sharp and Li (41) for very highly expressed and highly expressed genes.

preceded by a potential Shine-Dalgarno site (9). Two other potential initiation codons are present in the same reading frame just upstream of the one indicated in Fig. 4. Although we cannot rigorously exclude these codons as the site of translational initiation, this possibility seems unlikely since they are not preceded by Shine-Dalgarno sequences. The location and length of this open reading frame are consistent with the results obtained by deletion analysis (see above).

A second potential coding region was found on the same strand beginning at nucleotide 33 with TTG (Leu) or at 123 with GTG (Val) and terminating at 548. The results obtained by deletion mapping make it clear that this reading frame does not correspond to *iclR*. Whether these sequences might encode a different protein is not known. However, the lack of a Shine-Dalgarno sequence associated with either possible start codon makes this possibility seem unlikely.

The codon usage of *iclR* is presented in Table 3. This gene makes extensive use of a variety of codons that are rarely employed by highly expressed genes in *E. coli*. Although genes that are expressed at a high level in *E. coli* exhibit a striking bias in codon usage, genes that are expressed at a low level exhibit a much weaker codon preference (41). The observation that *iclR* employs a significant number of rare codons suggests that its product is maintained at a relatively low cellular level.

The results of the deletion analysis indicated that the promoter for *iclR* lay within the region that had been sequenced. Inspection of the sequences upstream of the coding region revealed a sequence similar to the consensus exhibited by those *E. coli* promoters that are recognized by the primary sigma factor, σ^{70} (10) (Fig. 4). The putative −10 region matches the consensus sequence (TATAAT) at five of six positions, whereas the putative −35 region matches the consensus (TTGACA) at three of six positions. The putative −10 and −35 regions are separated by 19 bp, which is slightly greater than the separation of these regions in the consensus (17 ± 1 bp). It remains to be determined, however, whether these sequences actually function as a promoter.

DISCUSSION

Expression of the glyoxylate bypass operon responds to the availability of a wide variety of carbon sources and culture conditions. This operon is induced when acetate is present as the sole carbon source. The glyoxylate bypass is essential for growth under these conditions, since it yields C_4 acids while avoiding the quantitative loss of the entering carbon as CO_2. Expression of the operon is repressed, even in the presence of acetate, if a preferred carbon source (e.g., glucose, glycerol, or pyruvate) is available (17). This operon is also repressed under anaerobic conditions (15). The mechanisms that allow expression of the glyoxylate bypass operon to respond to such a wide range of culture conditions remain unclear. However, the available evidence suggests that expression is controlled by multiple regulatory components.

Mutant alleles of *iclR* have been isolated that yield increased expression of the glyoxylate bypass operon under repressing conditions (18, 26). These alleles are recessive to wild-type *iclR*, an observation which suggests that *iclR*$^+$ encodes a repressor protein. The product of *fadR*$^+$, which was originally identified as encoding a regulatory protein for the fatty acid degradation regulon, has also been implicated in the control of the glyoxylate bypass operon (26). Mutations in *fadR* have been found to yield constitutive expression of the glyoxylate bypass operon during growth on succinate, a carbon source that significantly represses expression in wild-type strains.

The metabolic signals that allow the glyoxylate bypass operon to respond to the available carbon source remain elusive. A number of observations suggest that expression does not simply respond to the presence of acetate or its immediate products, acetyl phosphate and acetyl coenzyme A (acetyl-CoA). For example, expression of this operon does not respond to acetate in the presence of a preferred carbon source such as glucose or pyruvate (17). The ability of these alternative carbon sources to prevent induction by acetate does not appear to result from interference with acetate transport or metabolism, since cells growing on glucose convert acetate to acetyl-CoA at a rate that is nearly equivalent to that observe in cultures growing on acetate. Furthermore, addition of glucose to a culture growing on acetate did not affect the level of acetyl-CoA (45). It should also be noted that the glyoxylate bypass operon is induced during growth on fatty acids, a carbon source whose utilization does not proceed through acetate or acetyl phosphate (33). These observations clearly suggest that the glyoxylate bypass operon does not respond directly to the availability of acetate. Rather, it appears that expression of this operon responds to the general metabolic state of the cell by a more subtle mechanism.

Expression of the glyoxylate bypass operon also responds to the presence of molecular oxygen through a mechanism mediated by *arcA* and *arcB* (15). The products of *arcA* and *arcB* participate in the repression of a wide variety of genes when *E. coli* grows under anaerobic conditions. *arcA* exhibits 40% sequence similarity with *ompR* and therefore has been suggested to encode a DNA-binding protein. Genetic analysis has led to the suggestion that *arcB* may regulate *arcA* in much the same way as *envZ* regulates *ompR*: by phosphorylation (13, 14, 37).

Expression of the glyoxylate bypass operon is probably not subject to catabolite repression. This conclusion is suggested by the observation that the expression of isocitrate lyase was indistinguishable during growth of *E. coli* on glucose, glycerol, or pyruvate. A similar result was obtained when an *iclR* strain was examined (17). Furthermore, Wilson and Maloy have presented evidence that the glyoxylate bypass operon of *Salmonella typhimurium* is not subject to catabolite repression (46).

Although *iclR7*, isolated by Vinopal and Fraenkel (44), yielded a significant increase in expression of the glyoxylate bypass operon (20-fold) under repressing conditions, it did not yield full constitutivity: growth on acetate medium produced a further induction of 6-fold. A number of alternative explanations must be considered in accounting for this observation. For example, *iclR7* may encode a partially functional product that yields significant inhibition of expression during growth on repressing media. It is also possible that the repression observed in the *iclR7* strain during growth on glucose results from a different regulatory system that does not depend on the product of *iclR*$^+$. Resolution of this problem will require the generation of a null allele of *iclR*.

The ability of plasmid pKL51, which carries a truncated allele of *iclR*, to increase the level of expression of the glyoxylate bypass operon in the *iclR7* host under repressing conditions was unexpected. This plasmid had a qualitatively similar effect in a wild-type host (unpublished observation). One explanation for this phenomenon might be that the deletion in pKL51 represents a dominant negative mutation. Dominant negative mutations have been proposed to result when the overproduction of a mutant protein inhibits a regulatory pathway by forming an inactive complex with a normal protein (11). For example, derivatives of the *trp* repressor from which the DNA-binding domain has been deleted can inactivate the wild-type protein through the formation of inactive oligomers that contain both mutant and wild-type subunits (16). Dominant negative effects have also been suggested to arise from competition between wild-type and mutant proteins for a second regulatory factor that is present in limiting amounts. In our case, the product of *iclR7* may be partially functional and may be inactivated by formation of mixed oligomers with the truncated protein expressed from pKL51. As mentioned above, the generation of an unambiguous null mutation in *iclR* will be very helpful in resolving these issues.

A preliminary report of the cloning of *iclR* from *E. coli* has recently appeared (6). Surprisingly, the restriction map of that clone is very different from that which we have determined. Unfortunately, these authors did not present evidence that the clone which they isolated affects expression of the glyoxylate bypass operon. In contrast, the ability of our clone to complement *iclR7* indicates that it does include *iclR*$^+$.

Comparison of the deduced amino acid sequence of IclR with the sequences on file in the Protein Identification Resource failed to identify significant homology to other proteins. However, prediction of the secondary structure of this protein by the Chou-Fasman algorithm revealed the presence of a possible helix-turn-helix motif at the amino terminus (Fig. 5). The helix-turn-helix motif has been found in a variety of procaryotic and eucaryotic DNA-binding proteins and has been shown, in some cases, to constitute an essential part of the DNA-binding domain (28, 39, 47). The location of these sequences very near the amino terminus of IclR is reminiscent of most of these DNA-binding proteins. To optimize the alignment of the second helix shown in Fig. 5, we have assumed that the β turn includes four residues. The resulting alignment suggests that the sequences from IclR include the amino acids that are highly conserved in

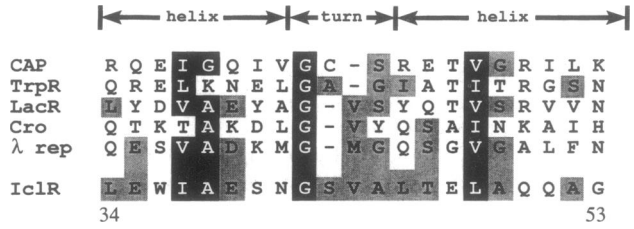

FIG. 5. Possible helix-turn-helix motif in the product of *iclR*. The Chou-Fasman algorithm identified a region of the *iclR* product that may form a helix-turn-helix motif. This region of IclR is compared with the DNA-binding domains of the catabolite activator protein (CAP) (30, 43), the *trp* repressor (TrpR) (40), the *lac* repressor (LacR) (28, 48), lambda Cro (1), and the lambda repressor (λ rep) (35, 36). Positions at which helix-turn-helix domains exhibit a strong preference are indicated with black boxes. Gray boxes indicate positions at which the IclR sequence is similar to one or more of the other sequences.

helix-turn-helix domains. Helix-turn-helix domains exhibit a strong preference for hydrophobic residues at positions 4 and 15, for glycine or alanine at position 5, and for glycine at position 9. Of course, it remains to be directly determined whether the corresponding sequences in IclR actually constitute a DNA-binding domain. The availability of clones of *iclR* should greatly facilitate future studies on the structure and function of this protein.

ACKNOWLEDGMENTS

We thank Howard Towle for a critical reading of the manuscript and Leonard Banaszak for advice on the prediction of protein secondary structure.

This work was funded by Public Health Service research grant DK40486 and research career development award DK01722, both from the National Institutes of Health.

LITERATURE CITED

1. **Anderson, W. F., D. H. Ohlendorf, Y. Takeda, and B. W. Matthews.** 1981. Structure of the cro repressor from bacteriophage lambda and its interaction with DNA. Nature **290**:754–758.
2. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl.** 1987. Current protocols in molecular biology. John Wiley & Sons, New York.
3. **Bochner, B. R., H. C. Huang, G. L. Schieven, and B. N. Ames.** 1980. Positive selection for loss of tetracycline resistance. J. Bacteriol. **143**:926–933.
4. **Brice, C. G., and H. L. Kornberg.** 1968. Genetic control of isocitrate lyase activity in *Escherichia coli*. J. Bacteriol. **96**: 2185–2186.
5. **Chung, T., D. J. Klumpp, and D. C. LaPorte.** 1988. Glyoxylate bypass operon of *Escherichia coli*: cloning and determination of the functional map. J. Bacteriol. **170**:386–392.
6. **Cortay, J. C., F. Bleicher, B. Duclos, Y. Cenatiempo, C. Gautier, J. L. Prato, and A. J. Cozzone.** 1989. Utilization of acetate in Escherichia coli: structural organization and differential expression of the ace operon. Biochimie **71**:1043–1049.
7. **Csonka, L. N., and A. J. Clark.** 1980. Construction of an Hfr strain useful for transferring *recA* mutations between *Escherichia coli* strains. J. Bacteriol. **143**:529–530.
8. **Garnak, M., and H. C. Reeves.** 1979. Phosphorylation of isocitrate dehydrogenase of Escherichia coli. Science **203**:1111–1112.
9. **Gold, L., D. Pribnow, T. Schneider, S. Shinedling, B. S. Singer, and G. Stormo.** 1981. Translational initiation in prokaryotes. Annu. Rev. Microbiol. **35**:365–403.
10. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of Escherichia coli promoter sequences. Nucleic Acids Res. **11**:2237–2255.
11. **Herskowitz, I.** 1987. Functional inactivation of genes by dominant negative mutations. Nature (London) **329**:219–222.
12. **Holms, W. H., and P. M. Bennett.** 1971. Regulation of isocitrate dehydrogenase activity in Escherichia coli on adaptation to acetate. J. Gen. Microbiol. **65**:57–68.
13. **Igo, M. M., A. J. Ninfa, and T. J. Silhavy.** 1988. A bacterial environmental sensor that functions as a protein kinase and stimulates transcriptional activation. Genes Dev. **3**:598–605.
14. **Iuchi, S., D. C. Cameron, and E. C. C. Lin.** 1989. A second global regulator gene (*arcB*) mediating repression of enzymes in aerobic pathways in *Escherichia coli*. J. Bacteriol. **171**:868–873.
15. **Iuchi, S., and E. C. C. Linn.** 1988. *arcA* (*dye*), a global regulatory gene in *Escherichia coli* mediating repression of enzymes in aerobic pathways. J. Bacteriol. **85**:1888–1892.
16. **Kelley, R. L., and C. Yanofsky.** 1985. Mutational studies with the trp repressor of Escherichia coli support the helix-turn-helix model of repressor recognition of operator DNA. Proc. Natl. Acad. Sci. USA **82**:483–487.
17. **Kornberg, H. L.** 1966. The role and control of the glyoxylate cycle in Escherichia coli. Biochem. J. **99**:1–11.
18. **Kornberg, H. L.** 1970. British biochemistry—past and present. Biochem. Soc. Symp. **30**:155–171.
19. **Kornberg, H. L., and N. B. Madsen.** 1951. Synthesis of C4-dicarboxylic acids from acetate by a "glyoxylate bypass" of the tricarboxylic acid cycle. Biochim. Biophys. Acta **24**:651–653.
20. **LaPorte, D. C., and T. Chung.** 1985. A single gene codes for the kinase and phosphatase which regulate isocitrate dehydrogenase. J. Biol. Chem. **260**:15291–15297.
21. **LaPorte, D. C., and D. E. Koshland, Jr.** 1982. A protein with kinase and phosphatase activities involved in the regulation of the tricarboxylic acid cycle. Nature **300**:458–460.
22. **LaPorte, D. C., and D. E. Koshland, Jr.** 1983. Phosphorylation of isocitrate dehydrogenase as a demonstration of enhanced sensitivity in covalent regulation. Nature (London) **305**:286–290.
23. **LaPorte, D. C., P. E. Thorsness, and D. E. Koshland, Jr.** 1985. Compensatory phosphorylation of isocitrate dehydrogenase: a mechanism for adaptation to the intracellular environment. J. Biol. Chem. **260**:10563–10568.
24. **LaPorte, D. C., K. Walsh, and D. E. Koshland, Jr.** 1984. The branch point effect: ultrasensitivity and subsensitivity to metabolic control. J. Biol. Chem. **259**:14068–14075.
25. **Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall.** 1951. Protein measurement with the Folin reagent. J. Biol. Chem. **193**:265–275.
26. **Maloy, S. R., and W. D. Nunn.** 1982. Genetic regulation of the glyoxylate shunt in *Escherichia coli* K-12. J. Bacteriol. **149**: 173–180.
27. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
28. **Matthews, B. W., D. H. Ohlendorf, W. F. Anderson, and Y. Takeda.** 1982. Structure of the DNA-binding region of lac repressor inferred from its homology with cro repressor. Proc. Natl. Acad. Sci. USA **79**:1428–1432.
29. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
30. **McKay, D. B., and T. A. Steitz.** 1981. Structure of catabolite gene activator protein at 2.9 Å resolution suggest binding to left-handed B-DNA. Nature (London) **290**:744–749.
31. **Neidhardt, F. C., P. L. Bloch, and D. F. Smith.** 1974. Culture medium for enterobacteria. J. Bacteriol. **119**:736–747.
32. **Nimmo, G. A., and H. G. Nimmo.** 1984. The regulatory properties of isocitrate dehydrogenase kinase and isocitrate dehydrogenase phosphatase from Escherichia coli ML308 and the roles of these activities in the control of isocitrate dehydrogenase. Eur. J. Biochem. **141**:409–414.
33. **Nunn, W. D.** 1986. A molecular view of fatty acid catabolism in Escherichia coli. Microbiol. Rev. **50**:179–192.
34. **O'Callaghan, C. H., P. W. Muggleton, and G. W. Ross.** 1969. Effects of β-lactamase from gram-negative organisms on ceph-

alosporins and penicillins, p. 57–63. Antimicrob. Agents Chemother. 1968.
35. **Ohlendorf, D. H., W. F. Anderson, M. Lewis, C. O. Pabo, and B. W. Matthews.** 1983. Comparison of the structures of cro and lambda repressor proteins from bacteriophage lambda. J. Mol. Biol. **169**:757–769.
36. **Pabo, C. O., and M. Lewis.** 1982. The operator-binding domain of lambda repressor: structure and DNA recognition. Nature (London) **298**:443–447.
37. **Ronson, C. W., B. T. Nixon, and F. M. Ausubel.** 1987. Conserved domains in bacterial regulatory proteins that respond to environmental stimuli. Cell **49**:579–581.
38. **Sanger, F., and A. R. Coulson.** 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J. Mol. Biol. **94**:441–448.
39. **Sauer, R. T., R. R. Yocum, R. F. Doolittle, M. Lewis, and C. O. Pabo.** 1982. Homology among DNA-binding proteins suggests use of a conserved super-secondary structure. Nature (London) **298**:447–451.
40. **Schevitz, R. W., Z. Otwinowski, A. Joachimiak, C. L. Lawson, and P. B. Sigler.** 1985. The three-dimensional structure of trp repressor. Nature (London) **317**:782–786.
41. **Sharp, P. M., and W.-H. Li.** 1986. Codon usage in regulatory genes in Escherichia coli does not reflect selection for "rare" codons. Nucleic Acids Res. **14**:7737–7749.
42. **Simons, R. W., F. Houman, and N. Kleckner.** 1987. Improved single and multicopy lac-based cloning vectors for protein and operon fusions. Gene **53**:85–96.
43. **Steitz, T. A., D. H. Ohlendorf, D. B. McKay, W. F. Anderson, and B. W. Matthews.** 1982. Structural similarity in the DNA-binding domains of catabolite gene activator and cro repressor proteins. Proc. Natl. Acad. Sci. USA **79**:3097–3100.
44. **Vinopal, R. T., and D. G. Fraenkel.** 1974. Phenotypic suppression of phosphofructokinase mutations in *Escherichia coli* by constitutive expression of the glyoxylate shunt. J. Bacteriol. **118**:1090–1100.
45. **Walsh, K., and D. E. Koshland.** 1985. Branch point control by the phosphorylation state of isocitrate dehydrogenase: a quantitative examination of fluxes during a regulatory transition. J. Biol. Chem. **260**:8430–8437.
46. **Wilson, R. B., and S. R. Maloy.** 1987. Isolation and characterization of *Salmonella typhimurium* glyoxylate shunt mutants. J. Bacteriol. **169**:3029–34.
47. **Yudkin, M. D.** 1988. The prediction of helix-turn-helix DNA-binding regions in proteins. Protein Eng. **1**:371–372.
48. **Zuiderweg, E. R., R. Kaptein, and K. Wuthrich.** 1983. Secondary structure of the lac repressor DNA-binding domain by two-dimensional 1H nuclear magnetic resonance in solution. Proc. Natl. Acad. Sci. USA **80**:5837–5841.