

# Genetic Analysis of a Chromosomal Region Containing Genes Required for Assimilation of Allantoin Nitrogen and Linked Glyoxylate Metabolism in *Escherichia coli*

EVA CUSA, NURIA OBRADORS, LAURA BALDOMÀ, JOSEFA BADÍA, AND JUAN AGUILAR\*

*Department of Biochemistry, Faculty of Pharmacy, University of Barcelona, 08028 Barcelona, Spain*

Received 8 July 1999/Accepted 29 September 1999

**Growth experiments with *Escherichia coli* have shown that this organism is able to use allantoin as a sole nitrogen source but not as a sole carbon source. Nitrogen assimilation from this compound was possible only under anaerobic conditions, in which all the enzyme activities involved in allantoin metabolism were detected. Of the nine genes encoding proteins required for allantoin degradation, only the one encoding glyoxylate carboligase (*gcl*), the first enzyme of the pathway leading to glycerate, had been identified and mapped at centisome 12 on the chromosome map. Phenotypic complementation of mutations in the other two genes of the glycerate pathway, encoding tartronic semialdehyde reductase (*glxR*) and glycerate kinase (*glxK*), allowed us to clone and map them closely linked to *gcl*. Complete sequencing of a 15.8-kb fragment encompassing these genes defined a regulon with 12 open reading frames (ORFs). Due to the high similarity of the products of two of these ORFs with yeast allantoinase and yeast allantoate amidohydrolase, a systematic analysis of the gene cluster was undertaken to identify genes involved in allantoin utilization. A BLASTP search predicted four of the genes that we sequenced to encode allantoinase (*allB*), allantoate amidohydrolase (*allC*), ureidoglycolate hydrolase (*allA*), and ureidoglycolate dehydrogenase (*allD*). The products of these genes were overexpressed and shown to have the predicted corresponding enzyme activities. Transcriptional fusions to *lacZ* permitted the identification of three functional promoters corresponding to three transcriptional units for the structural genes and another promoter for the regulatory gene *allR*. Deletion of this regulatory gene led to constitutive expression of the regulon, indicating a negatively acting function.**

Utilization of allantoin as a nitrogen source by *Escherichia coli* and other *Enterobacteriaceae* was first suggested by Vogels and van der Drift in a review of the degradation of purines and pyrimidines by microorganisms (26). The proposed pathway (Fig. 1) opens the allantoin ring, *R*-(-) or *S*-(+) stereoisomer, yielding allantoate by the action of the hydrolase *RS*-allantoinase. Allantoate is converted to ureidoglycine,  $\text{NH}_4^+$ , and  $\text{CO}_2$ , and ureidoglycine is converted to ureidoglycolate and  $\text{NH}_4^+$  by the sequential action of the enzyme allantoate amidohydrolase (28). Ureidoglycolate is a substrate of two different enzymes, ureidoglycolate hydrolase, which converts it to urea and glyoxylate, and ureidoglycolate dehydrogenase, which oxidizes it to oxaluric acid. Glyoxylate is incorporated into the general metabolism through the glycerate pathway. This central pathway requires the action of glyoxylate carboligase, which forms tartronic semialdehyde; tartronic semialdehyde reductase, which converts the semialdehyde into glycerate; and glycerate kinase, which forms the glycolytic intermediate 3-phosphoglycerate (17). The conversion of oxaluric acid into oxamate and carbamoyl phosphate by the action of oxamate transcarbamoylase activity has been reported for *E. coli* (23). Carbamoyl phosphate is subsequently converted to  $\text{CO}_2$ ,  $\text{NH}_4^+$ , and ATP. The values of the fermentation balance obtained by Barker (2) are consistent with the proposed pathway (Fig. 1).

Chang et al. (6), in their examination of enzymes of the acetohydroxy acid synthase family, have studied glyoxylate carboligase and reported the sequence and location of the corre-

sponding gene at centisome 12 on the *E. coli* map. However, the locations of the genes for the other two enzymes of the glycerate pathway remain unknown.

In this report, we characterize and map the genes encoding tartronic semialdehyde reductase and glycerate kinase and show that they are clustered with genes involved in the anaerobic utilization of allantoin. The gene products of this cluster are identified, as is a related regulator.

## MATERIALS AND METHODS

**Bacterial strains.** All the strains used were *E. coli* K-12 derivatives. The relevant genotypes and sources of these bacterial strains are given in Table 1.

**Cell growth.** Cells were grown and harvested as described previously (4). For aerobic growth, carbon sources were added to a basal inorganic medium (4) at a 60 mM carbon concentration; for anaerobic growth, 120 mM was used. Casein acid hydrolysate was prepared at 1% for the aerobic growth of transformed cells. In these minimal media, the nitrogen source was ammonium sulfate (4), except when the nitrogen source was allantoin, which was added at a 60 mM concentration. To avoid hydrolysis, the allantoin medium was freshly prepared by dissolving allantoin in mineral medium and sterilizing the solution by filtration. When required, the following antibiotics were used at the indicated concentrations: ampicillin, 100  $\mu\text{g/ml}$ ; tetracycline, 12.5  $\mu\text{g/ml}$ ; chloramphenicol, 30  $\mu\text{g/ml}$ ; and kanamycin, 50  $\mu\text{g/ml}$ . 5-Bromo-4-chloro-3-indolyl- $\beta$ -D-galactoside (X-Gal) and isopropyl- $\beta$ -D-thiogalactoside were used at 30 and 10  $\mu\text{g/ml}$ , respectively. For  $\beta$ -galactosidase assays, cultures were allowed to double five or six times during mid-log phase to an optical density at 600 nm of 0.5 for aerobic cultures or 0.25 for anaerobic cultures.

**Preparation of cell extracts and enzyme assay.** Cells were harvested at the end of exponential phase, and a cell extract was prepared as described previously (4) with 20 mM Tris-HCl buffer (pH 7.3). EDTA and  $\beta$ -mercaptoethanol were added to the allantoate amidohydrolase or tartronic semialdehyde reductase assay mixtures to final concentrations of 1 and 10 mM, respectively.

Glyoxylate carboligase was assayed by measuring the production of  $^{14}\text{CO}_2$  from [ $1,2\text{-}^{14}\text{C}$ ]glyoxylate (specific activity, 0.45  $\mu\text{Ci}/\mu\text{mol}$ ) by a modification of the method of Chang et al. (6); phenylethylamine was used instead of methylbenzethonium hydroxide in the central well of the reaction vial to collect the radioactive  $\text{CO}_2$ .

Tartronic semialdehyde reductase was assayed by coupling this activity to purified glyoxylate carboligase and measuring the decrease in NADH with

\* Corresponding author. Mailing address: Departamento de Bioquímica, Facultad de Farmacia, Universidad de Barcelona, Avda. Diagonal 643, 08028 Barcelona, Spain. Phone: 34-93-402 4521. Fax: 34-93-402 1896. E-mail: Jaguilard@farmacia.far.ub.es.

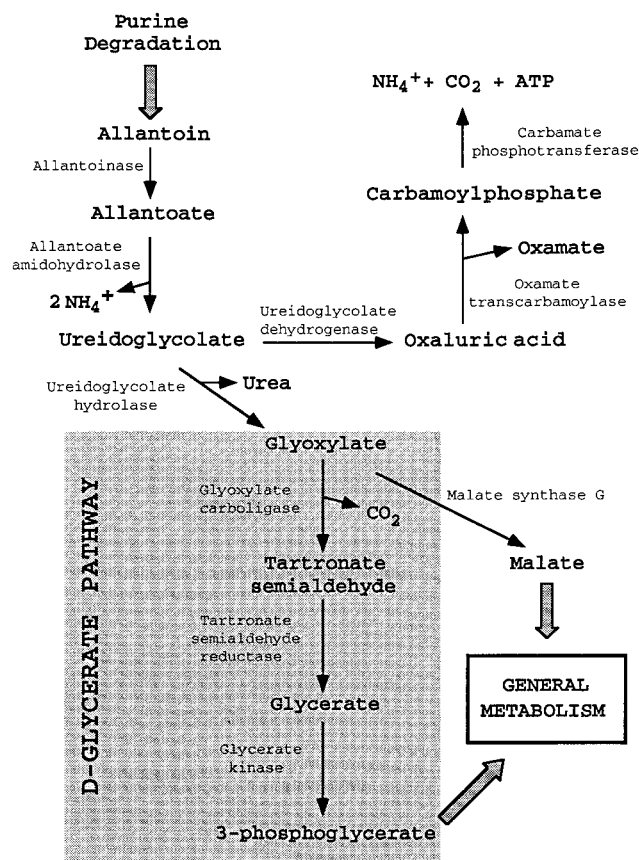


FIG. 1. Metabolic map of the pathway for the degradation of allantoin in *Enterobacteriaceae*.

glyoxylate as a substrate. The reaction mixture (1 ml) contained 100 mM potassium phosphate buffer (pH 8.0), 6 mM glyoxylate, 0.28 mM NADH, 5 mM  $MgCl_2$ , 0.5 mM thiamine diphosphate, and an excess amount (2.5 U) of glyoxylate carboligase purified as described by Chang et al. (6).

Glycerate kinase activity was determined from the rate of NADH oxidation in the presence of 2 mM D-glycerate, 2 mM ATP, 0.35 mM reduced glutathione, 3.5 mM  $MgCl_2$ , 0.28 mM NADH, a pyruvate kinase-L-lactate dehydrogenase mixture (18  $\mu g/ml$ ), and 1.75 mM phosphoenolpyruvate (1).

Allantoinase was assayed by the colorimetric method described by Lee and

TABLE 1. Bacterial strains used in this work

Strain	Description	Source or reference
XL1Blue	<i>recA1 lac endA1 gyrA96 thi hsdR17 supE44 relA1 (F' proAB lacI<sup>q</sup> lacZ<math>\Delta</math>M15 Tn10)</i>	Stratagene
ECL1	HfrC <i>phoA8 relA1 tonA22 T2'</i> ( $\lambda$ )	12
JC7623	<i>arg thi thr leu pro his strA recB21 recC22 sbcB15</i>	27
TE2680	$F^- \lambda^- IN(rmD-rmE) \Delta lacX74 rplS galK2 recD::Tn10d-tet tpDC700::putA13033::(Kan^r Cm^r lac)$	7
JA170	ECL1 <i>gbk::Tn5</i>	This work
JA171	ECL1 <i>gbR::Tn5</i>	This work
JA172	ECL1 <i>gcl::Tn5</i>	This work
JA173	ECL1 $\Delta allR$	This work
JA174	ECL1 <i>allA::cat</i>	This work
EcoR26	Human feces isolate	16
EcoR36	Human urine isolate	16
EcoR45	Pig isolate	16

Roush (11) applied to the reaction of the allantoin product with phenylhydrazine HCl and potassium ferricyanide. Allantoate amidohydrolase was assayed by measuring ureidoglycolate production as described by Xu et al. (28). Ureidoglycolate production was determined by differential analyses of glyoxylate derivatives (25).

Ureidoglycolate hydrolase activity was determined by measuring the production of glyoxylate through its reaction with phenylhydrazine as described by Pineda et al. (18). Ureidoglycolate dehydrogenase was in turn assayed by measuring the reduction of NAD associated with the oxidation of the substrate by the procedure described by van der Drift et al. (24).

In all the above indicated determinations, 1 U was defined as the amount of enzyme converting 1  $\mu mol$  of substrate per min.

$\beta$ -Galactosidase activity was measured by the hydrolysis of *o*-nitrophenyl- $\beta$ -D-galactopyranoside and expressed in Miller units (15). Values reported are representative of at least three separate experiments performed in duplicate. Protein concentrations were determined by the method of Lowry et al. (14) with bovine serum albumin as a standard.

**DNA manipulation.** Bacterial genomic DNA was obtained as described by Silhavy et al. (21). Plasmid DNA was routinely prepared by the boiling method (9). For large-scale preparation, a crude DNA sample was passed over a column (Qiagen GmbH, Düsseldorf, Germany). DNA manipulations were performed essentially as described by Sambrook et al. (19). The gene mapping membrane (Takara Shuzo Biomedicals) containing the immobilized DNA of the entire Kohara  $\lambda$  phage bank was hybridized as follows. The membrane was first pre-hybridized at 65°C with salmon DNA in  $6\times$  SSC ( $1\times$  SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–5 $\times$  Denhardt solution–0.5% sodium dodecyl sulfate. After 3 h, 100  $\mu l$  of TE buffer containing  $15\times 10^6$  cpm of the probe,  $^{32}P$  labeled by the random-priming method, was added, and hybridization was maintained for 16 h. The filter was rinsed for three periods of 15 min each with 2 $\times$  SSC, the last rinse also containing 0.1% sodium dodecyl sulfate. The dried filter was used to detect specific hybridization by autoradiography at  $-70^\circ C$ .

The DNA sequence was determined by the dideoxy chain termination procedure of Sanger et al. (20) with double-stranded plasmid DNA as the template.

**Mutagenesis and genetic techniques.** Tn5 insertion mutagenesis was carried out by infection with phage  $\lambda 467$  (*b221 cIs857 rex::Tn5 Oam29 Pam80*) as described by Bruijn and Lupski (5). Phage P1 transduction experiments were performed as described by Miller (15). The chloramphenicol resistance gene cassette CAT19 was used in the gene inactivation experiments (8, 27) by insertion into the restriction sites indicated below. Plasmids carrying inactivated genes were linearized and used to transform strain JC7623 to chloramphenicol resistance ( $Cm^r$ ). This strain efficiently recombinates linear DNA into its chromosome (27). P1 *vir* lysates obtained from the selected  $Cm^r$  recombinants were used to transduce the chloramphenicol acetyltransferase (CAT) insertions into the parental strain ECL1. Precise in-frame tagged deletions were generated by cross-over PCR (13).

To construct the 759-bp deletion of open reading frame (ORF) o271, the following set of oligonucleotide primers was used: o271-No, 5'-GGTGGATCC GATGTTGTCTCTTAAGG-3'; o271-Ni, 5'-CCCATCCACTAAACTAAAC ACCGTCTAACTTCCGTCATAC-3'; o271-Co, 5'-CGCGGATCCGACCAGA ACCATCAGGTG-3'; and o271-Ci, 5'-TGTTAAGTTAGTGGATGGGG ATATCAGCACGGCGTTGGG-3'. The fragment containing the deletion was then cloned into the *Bam*HI site of vector pKO3, a gene replacement vector that contains a temperature-sensitive origin of replication and markers for positive and negative selection for chromosomal integration and excision. The deletion was introduced into the chromosome by use of the pKO3 gene replacement protocol described by Link et al. (13). Chromosomal insertions and deletions were confirmed by PCR.

**Construction of lacZ fusions for identification of promoter sequences.** To create operon fusions, DNA fragments of the 5' upstream region of each gene (Fig. 2) were cloned into plasmid pRS550 or pRS551 (22). In some experiments, the DNA fragment was extended to several genes to investigate the presence of possible promoters within the coding regions. The plasmids carried a promoterless *lac* operon and genes that confer resistance to both kanamycin and ampicillin. Recombinant plasmids were selected, after transformation of strain XL1Blue, as blue colonies on Luria-Bertani agar plates containing X-Gal, ampicillin, and kanamycin. Plasmid DNA was sequenced by use of an M13 primer to ensure that the desired fragment was inserted in the correct orientation. Merodiploids were obtained by transferring the fusions as single copies into the *trp* operon of *E. coli* TE2680 as described by Elliott (7). Transformants were selected for kanamycin resistance and screened for sensitivity to ampicillin and chloramphenicol. P1 *vir* lysates were made to transduce the fusions into strain ECL1.

**Nucleotide sequence accession numbers.** Sequences determined in this study were deposited in GenBank under accession no. U89024 and U89279.

## RESULTS

**Utilization of allantoin as a nitrogen source.** Cultures of *E. coli* ECL1 placed on minimal medium plus allantoin did not grow either aerobically or anaerobically, indicating that these cells were unable to use this compound as a carbon source. In

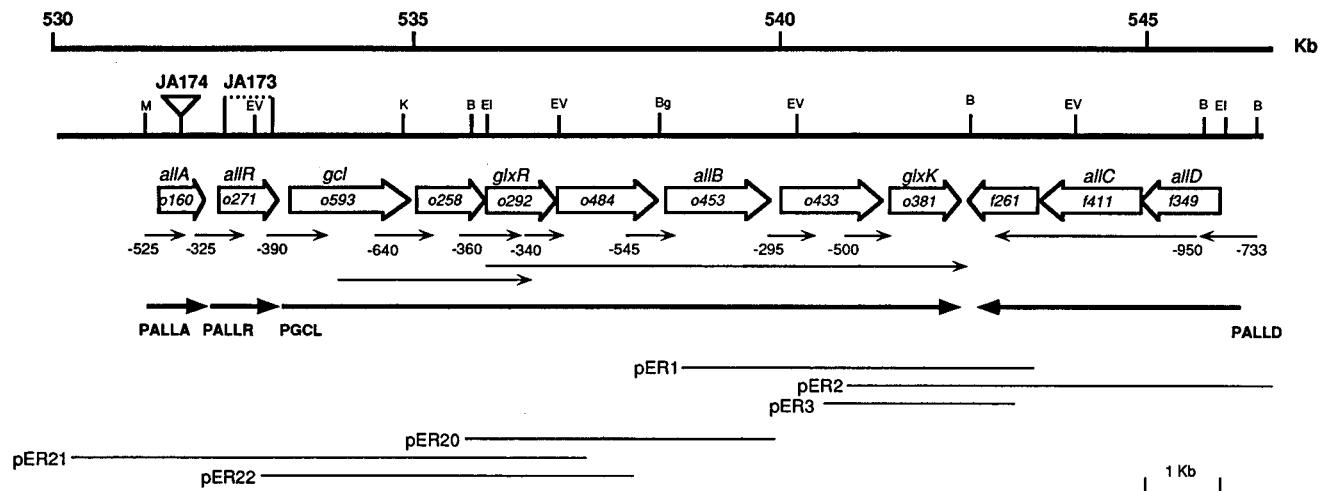


FIG. 2. Physical and genetic map of the region containing the genes involved in glycerate and allantoin metabolism. The thick lines represent the sequenced genomic fragment. The extension and direction of the included ORFs are indicated by open arrows and labeled by proposed gene symbols, according to determined function. Thin arrows correspond to the fragments fused to *lacZ* for testing promoter function and are labeled by numbers that indicate the length in nucleotides upstream of ATG. Thick arrows show the proposed transcriptional units labelled according to the promoter of the first transcribed gene. Thin lines correspond to the clones used for the identification of the glycerate pathway genes. Relevant restriction sites are marked as follows: B, *Bam*HI; Bg, *Bg*III; EI, *Eco*RI; EV, *Eco*RV; K, *Kpn*I; M, *Mlu*I.

contrast, growth on minimal medium with xylose as a carbon source and allantoin instead of ammonium sulfate as a nitrogen source showed that these bacteria used allantoin as a sole nitrogen source only in the absence of oxygen. The doubling time of the anaerobic cultures was 6 h and was about one-third slower than that obtained with ammonium sulfate in the same carbon source. Studies on allantoin utilization were extended to several natural isolates of *E. coli* taken from the Selander collection (16), namely, strain EcoR45 from a pig, strain EcoR36 from human urine, and strain EcoR26 from human feces. All of them were able to utilize allantoin as a nitrogen source but not as a carbon source.

With the exception of oxamate transcarbamoylase (not tested due to the unavailability of a substrate), all activities involved in allantoin utilization, including those of the glycerate pathway, were measured with crude extracts of cells of strain ECL1 grown anaerobically on a xylose-ammonia or xylose-allantoin medium. All the required activities were found to be significantly higher when allantoin replaced ammonia in

the cultures, thus giving support to the functional role of allantoin as a nitrogen source for cell growth (Table 2).

**Mapping of mutations in the glycerate pathway.** The biochemical relationship between allantoin metabolism and the glycerate pathway (Fig. 1), together with the ability to isolate glycerate pathway mutants, enticed us to study the location of the genes encoding the enzymes of this pathway as well as their possible linkage to the allantoin genetic system. The search for mutations in the glycerate pathway was approached by selecting mutants with defects in glyoxylate utilization. To this end, cells of wild-type strain ECL1 were mutagenized with transposon Tn5 and then selected for kanamycin resistance and the inability to utilize glyoxylate. From among 20 independent mutants tested for enzyme activities of the glycerate pathway, the following were selected: strain JA170, lacking glycerate kinase activity; strain JA171, lacking tartronic semialdehyde reductase activity; and strain JA172, lacking glyoxylate carboxylase activity.

Mutations causing defects in tartronic semialdehyde reductase and glycerate kinase were used to map the corresponding genes, for which we propose the gene symbols *glxR* and *glxK*, respectively. Glyoxylate carboxylase, encoded by *gcl*, was previously mapped by Chang et al. (6) at centisome 12 on the chromosome. A *Sau*III A gene library of wild-type strain ECL1 was constructed on pBR322 and used to transform strain JA170. Clones were selected by growth on Luria-Bertani agar-ampicillin-kanamycin plates and replica plating on glyoxylate plates. Among 11 positive clones, 3 were characterized further (pER1 to pER3; Fig. 2). The restriction maps were determined, and plasmid DNA was used to transform strain JA170 to check glycerate kinase expression. All transformants displayed high glycerate kinase activity when grown on casein acid hydrolysate in the presence of 15 mM glyoxylate.

A 1-kb *Sal*I-*Bam*HI fragment (Fig. 2) common to all clones was isolated, <sup>32</sup>P labeled by the random-priming method, and used to hybridize a membrane containing DNA of all clones of the Kohara library (10) (Takara Shuzo Biomedicals). Two clones, 156(9E5) and 157(6E7), hybridized with the *glxK* probe.

TABLE 2. Enzyme activities involved in the anaerobic utilization of allantoin for wild-type strain ECL1 and mutant strain JA173

Enzyme	Activity (mU/mg) under the following growth conditions for the indicated strain:			
	Xylose-NH <sub>4</sub> <sup>+</sup>		Xylose-allantoin	
	ECL1	JA173	ECL1	JA173
Glyoxylate carboxylase	15	515	112	890
Tartronic semialdehyde dehydrogenase	5	175	40	310
Glycerate kinase	0	85	10	70
Allantoinase	0	20	27	520
Allantoate amidohydrolase	0	10	13	320
Ureidoglycolate dehydrogenase	0	10	5	175
Ureidoglycolate hydrolase	0	4	3	9

The 3-kb overlapping fragment between these two clones was, according to the Kohara map, coincident with the physical map of the clones expressing glycerate kinase activity, allowing us to locate the *glxK* gene at centisome 11.7 on the chromosome map.

Likewise, strain JA171 was transformed with the gene library of strain ECL1, and several clones were isolated by phenotypic complementation of the tartronic semialdehyde reductase mutation. All transformants displayed high tartronic semialdehyde reductase activity when grown on casein acid hydrolysate in the presence of 15 mM glyoxylate. The restriction maps of three of the complementing plasmids (pER20 to pER22) showed some overlap with clones of the glycerate kinase complementation collection, indicating linkage between the two markers (Fig. 2). Furthermore, when cells of strain JA172 were transformed with DNA from plasmid pER22, which carried the *glxR* gene, and transformants were grown on casein acid hydrolysate in the presence of glyoxylate, the corresponding extracts also displayed high glyoxylate carboligase activity, thus proving that this clone carried both the *glxR* and the *gcl* genes.

**Sequence of the gene cluster and gene organization.** Clustering of the three genes of the glycerate pathway and the presence between *glxR* and *glxK* of a gene encoding a product displaying high similarity to yeast allantoinase led us to search in this region for other genes encoding enzymes similar to those involved in the utilization of allantoin. A DNA region of 15.8 kb (Fig. 2) was totally sequenced by use of clones of the *Sau*III library, with the exception of an overlapping 2.3-kb internal fragment for glyoxylate carboligase and ORF o258 (GenBank accession no. L03845), the sequence of which had been published before (6). Analyses of the complete sequence permitted us to define the organization of the 12 ORFs indicated in Fig. 2. Subsequent availability of the sequences of this region (3) in the *E. coli* genome (GenBank project accession no. AE000156 and AE000157) allowed confirmation of the sequence identity as well as the deduced gene organization, with the exception of ORF o484. In the genome project, this region was proposed to contain two ORFs yielding proteins of 92 and 437 amino acids. However, on the basis of similarity to other permeases (see below), we assume that ORF o484 is functional for translation *in vivo*.

The amino acid sequences encoded by the unidentified ORFs were used as the query sequences for a BLASTP search of the GenBank database by use of the National Center for Biotechnology Information BLAST server. In this way, high similarity was found between the o160 product and ureidoglycolate hydrolase of *Saccharomyces cerevisiae*, the o484 product and yeast allantoin permease, the o453 product and yeast allantoinase, the o433 product and xanthine permease of *Bacillus subtilis* or uracil permease of *Bacillus caldolyticus*, and finally between the f349 product and several malate dehydrogenases. Likewise, with the ORF o271 sequence as a query sequence, the BLASTP search revealed similarity to regulatory proteins such as the glycerol activator (*Streptomyces*), the acetate repressor (*E. coli*), and the Kdg repressor involved in pectinolysis in *Erwinia chrysanthemi*, all belonging to the known IclR family. Finally, no similarity to any group of functional proteins was found for the products of ORFs o258 and f261.

**Identification of gene products.** To determine if the functions suggested by the above analyses corresponded to actual protein activities, the following strategies were used. To confirm ureidoglycolate hydrolase as the product encoded by ORF o160 (proposed gene symbol *allA*), the gene was subcloned in plasmid pBlueScript and used to transform strain ECL1. Cell extracts of recombinant ECL1 harboring this plasmid showed

the corresponding activity at about 10 times the level in wild-type *E. coli*, consistent with the suggestion that the protein expressed by *allA* was ureidoglycolate hydrolase. Furthermore, disruption of this gene by a CAT cassette inserted in the *Stu*I restriction site and transfer of the mutant gene to the genome abolished the hydrolase activity in this mutant (strain JA174). Nevertheless, this mutation did not affect the rate of utilization of allantoin, as indicated by the normal growth of this mutant in xylose-allantoin medium. In contrast, the glycerate kinase-negative strain JA170 displayed poor growth when allantoin was used as a nitrogen source.

In parallel experiments, other genes were subcloned in pBlueScript, and their expression was analyzed. The results showed that gene o453 (proposed gene symbol *allB*) expressed high allantoinase activity, gene f411 (proposed gene symbol *allC*) expressed high allantoate amidohydrolase activity, and gene f349 (proposed gene symbol *allD*) expressed high ureidoglycolate dehydrogenase activity. All plasmid-encoded activities were between 10 and 15 times the activities in wild-type *E. coli*.

**Analyses of functional promoters.** Identification of promoters in the gene cluster was performed by fusing to *lacZ* the 5' end of each of the intervening genes. To check the presence of any promoter activity in this region, we tested two additional constructs: one lacking the *gcl* promoter but containing the genes *gcl*, o258, and *glxR* and the other containing the genes o484 to *glxK*. Cells of strain ECL1 containing the different gene fusions indicated in Fig. 2 were grown aerobically and anaerobically on minimal medium containing xylose plus glyoxylate. Expression was also studied by anaerobic growth on xylose with allantoin as a nitrogen source. Analysis of the different *lacZ* fusions revealed four promoters, corresponding to genes *allA*, *allR* (see below), *gcl*, and *allD* (Fig. 2).  $\phi$ (*allR-lacZ*) showed constitutive expression whereas  $\phi$ (*allA-lacZ*),  $\phi$ (*allD-lacZ*), and  $\phi$ (*gcl-lacZ*) showed inducible expression with glyoxylate and allantoin under anaerobic conditions (Fig. 3). Levels of enzyme specific activities under noninducing and inducing conditions matched rather well the pattern of operon expression (Table 2). Aerobically, glyoxylate was also able to induce the expression of  $\phi$ (*allA-lacZ*) and  $\phi$ (*gcl-lacZ*) but not  $\phi$ (*allD-lacZ*).

**Identification of the regulator.** ORF o271 (proposed gene symbol *allR*), whose product displayed similarity to the regulatory proteins indicated above, was deleted by crossover PCR, generating mutant strain JA173. In this mutant, the pattern of induction of the three promoter fusions corresponding to the structural genes was analyzed to determine the involvement and possible function of this gene in the allantoin system. All fusion constructs showed constitutive expression, as indicated by the levels of  $\beta$ -galactosidase activities in the absence of inducers. The lack of *allR* also caused the hyperinducibility of  $\phi$ (*allD-lacZ*) by allantoin under anaerobic conditions. Constitutive levels of enzyme specific activities were also detected in extracts of anaerobically grown cells of strain JA173 under noninducing conditions (Table 2). Thus, the product encoded by *allR* corresponded to a negatively acting regulatory protein.

## DISCUSSION

The utilization of purine bases and their derivatives as carbon or nitrogen sources in *E. coli* remains undefined, as no consistent information is found in the literature. In their review, Vogels and van der Drift (26) found important discrepancies in reports of the capacity of *E. coli* and other bacteria to use these nutrients for growth. Our results indicate that cells of *E. coli* are able to utilize allantoin as a nitrogen source only

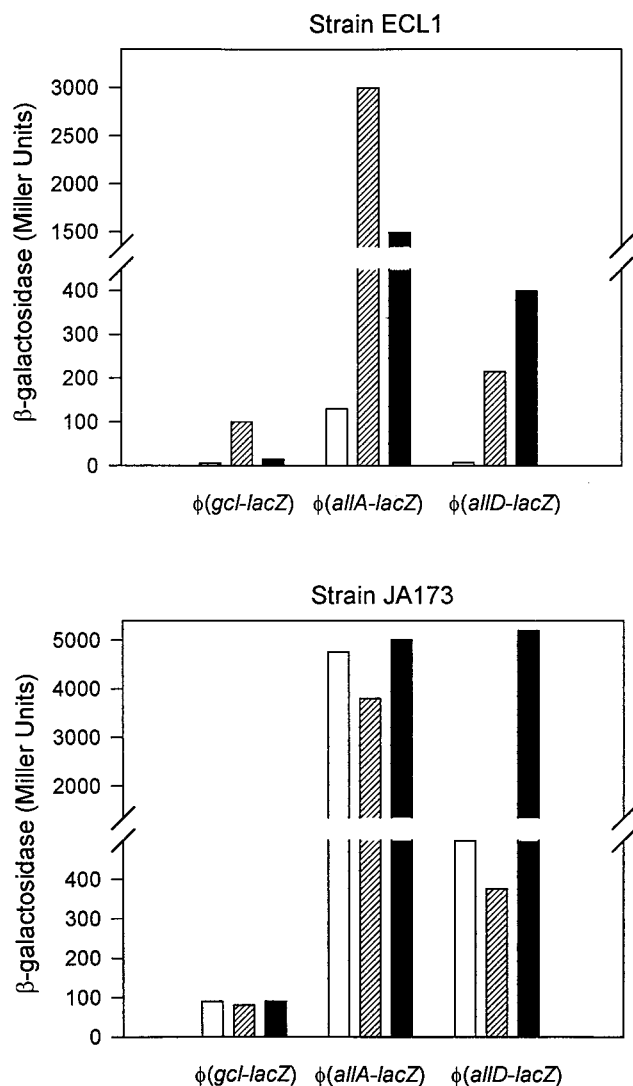


FIG. 3.  $\beta$ -Galactosidase activities of  $\phi(gcl-lacZ)$ ,  $\phi(allA-lacZ)$ , and  $\phi(allD-lacZ)$  transcriptional fusions in cells of strain ECL1 and strain JA173 grown anaerobically on D-xylose-NH<sub>4</sub><sup>+</sup> (open bars), D-xylose-glyoxylate (hatched bars), and D-xylose-allantoin (black bars).

anaerobically. The noninducibility under aerobic conditions of the *allD* promoter, corresponding to the transcriptional unit encoding ureidoglycolate dehydrogenase and allantoate amidohydrolase, accounts for the inability to utilize allantoin as a nitrogen source aerobically. Concerning the metabolic roles shared between the two branches of the allantoin pathway, the observation of a normal utilization of allantoin in the absence of ureidoglycolate hydrolase indicates a predominant channeling of nitrogen through the branch of oxamate formation. Thus, impairment of allantoin utilization in glycerate kinase mutants must be explained by toxic effects of accumulated intermediate metabolites, such as glyoxylate, resulting from minor metabolic flow through the glycerate branch under anaerobic conditions.

Of the nine genes encoding the proteins required to convert allantoin to NH<sub>4</sub><sup>+</sup>, CO<sub>2</sub>, urea, 3-phosphoglycerate, and oxamate, only the one encoding glyoxylate carboligase had been previously identified and mapped (6). There is a slight discrep-

ancy in the kilobase scale between the maps of Kohara et al. (10) and the genome project (3), causing a difference in the locations of the genes of the cluster defined here. In the last and more precise map established by the genome project, the system lies between 531.5 and 546.3 kb, corresponding to centisomes 11.45 and 11.77, respectively. We have found good agreement between our proposal and that of the genome project in the assignment of the 12 ORFs, with two exceptions: (i) an alternative translation of ORF o484 into a protein of 92 amino acids followed by another of 437 amino acids, with the overlapping of 1 nucleotide; and (ii) a start for ORF o433 one codon upstream, converting it to o434, which is inconsistent with the location of the Shine-Dalgarno box. The similarity of the o484 product to other permeases of the family seems to indicate that the full-length translation is more likely and realistic than the reported split of the gene.

In spite of the multiple origins of the central intermediate metabolite glyoxylate, the substrate of the glycerate pathway, it is tempting to speculate that the genes of this pathway have evolved as part of the allantoin utilization system.

Under anaerobic conditions, allantoin activates the promoters at the 5' ends of *allA* and *allD*, leading to the expression of ureidoglycolate hydrolase, allantoate amidohydrolase, and ureidoglycolate dehydrogenase. Allantoin also weakly induces the third operon (*gcl*), expressing, together with the genes of the glycerate pathway, low levels of allantoinase, thus allowing a basal metabolic flow of allantoin degradation. The function of ureidoglycolate hydrolase is to generate the intermediate glyoxylate, which in turn reinforces the weak activation of the *gcl* promoter by allantoin and thus allows efficient utilization of allantoin. Hyperinducibility of the *allD* promoter could correspond to the effects of another regulator of unknown location and which could be under the control of *allR* or simply compete with it for the same promoter sequences.

#### ACKNOWLEDGMENTS

We are grateful to Eva Román for help in the characterization of the glycerate kinase mutants.

This work was supported by grant PB97-0920 from the Dirección General de Enseñanza Superior e Investigación Científica, Madrid, Spain, and by the help of the Comissionat per Universitats i Recerca de la Generalitat de Catalunya.

#### REFERENCES

- Anderson, R. L., and W. A. Wood. 1962. Pathway of xylose and l-xylose degradation in *Aerobacter aerogenes*. *J. Biol. Chem.* **237**:296-303.
- Barker, H. A. 1943. *Streptococcus allantoicus* and the fermentation of allantoin. *J. Bacteriol.* **46**:251-259.
- Blattner, F. R., G. Plunkett III, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* **277**:1453-1474.
- Boronat, A., and J. Aguilar. 1979. Rhamnose-induced propanediol oxidoreductase in *Escherichia coli*: purification, properties, and comparison with the fucose-induced enzyme. *J. Bacteriol.* **140**:320-326.
- Bruijn, F. J., and J. R. Lupski. 1984. The use of transposon Tn5 mutagenesis in the rapid generation of correlated physical and genetic maps of DNA segments cloned into multicopy plasmids. *Gene* **27**:131-149.
- Chang, Y. Y., A. Y. Wang, and J. E. Cronan, Jr. 1993. Molecular cloning, DNA sequencing, and biochemical analyses of *Escherichia coli* glyoxylate carboligase. *J. Biol. Chem.* **268**:3911-3919.
- Elliot, T. 1992. A method for constructing single-copy *lac* fusions in *Salmonella typhimurium* and its application to the *hemA-prfA* operon. *J. Bacteriol.* **174**:245-253.
- Fuqua, W. C. 1992. An improved chloramphenicol resistance gene cassette for site-directed marker replacement mutagenesis. *BioTechniques* **12**:223-225.
- Holmes, D. S., and M. Quigley. 1981. A rapid boiling method for the preparation of bacterial plasmids. *Anal. Biochem.* **114**:193-197.
- Kohara, Y., K. Akiyama, and K. Isono. 1987. The physical map of the whole

- E. coli* chromosome: application for a new strategy for rapid analysis and sorting of a large genomic library. *Cell* **50**:495–508.
11. **Lee, K. W., and A. H. Roush.** 1964. Allantoinase assays and their application to yeast and soybean allantoinases. *Arch. Biochem. Biophys.* **108**:460–467.
  12. **Lin, E. C. C.** 1976. Glycerol dissimilation and its regulation in bacteria. *Annu. Rev. Microbiol.* **30**:535–578.
  13. **Link, A. J., D. Phillips, and G. M. Church.** 1997. Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. *J. Bacteriol.* **179**:6228–6237.
  14. **Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall.** 1951. Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193**:265–275.
  15. **Miller, J. H.** 1992. A short course in bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
  16. **Ochman, H., and R. K. Selander.** 1984. Standard reference strains of *Escherichia coli* from natural populations. *J. Bacteriol.* **157**:690–693.
  17. **Ornston, L. N., and M. K. Ornston.** 1969. Regulation of glyoxylate metabolism in *Escherichia coli* K-12. *J. Bacteriol.* **98**:1098–1108.
  18. **Pineda, M., P. Piedras, and J. Cardenas.** 1994. A continuous spectrophotometric assay for ureidoglycolase activity with lactate dehydrogenase or glyoxylate reductase as coupling enzyme. *Anal. Biochem.* **222**:450–455.
  19. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
  20. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
  21. **Silhavy, T. J., M. L. Berman, and L. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
  22. **Simons, R. W., F. Houtman, and N. Kleckner.** 1987. Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. *Gene* **53**:85–96.
  23. **Tigier, H., and S. Grisolia.** 1965. Induction of carbamyl-P specific oxamate transcarbamylase by parabanic acid in a Streptococcus. *Biochem. Biophys. Res. Commun.* **19**:209–214.
  24. **van der Drift, C., P. E. M. van Helvoort, and G. D. Vogels.** 1971. S-Ureidoglycolate dehydrogenase: purification and properties. *Arch. Biochem. Biophys.* **145**:465–469.
  25. **Vogels, G. D., and C. van der Drift.** 1970. Differential analyses of glyoxylate derivatives. *Anal. Biochem.* **33**:143–157.
  26. **Vogels, G. D., and C. van der Drift.** 1976. Degradation of purines and pyrimidines by microorganisms. *Bacteriol. Rev.* **40**:403–468.
  27. **Winans, S. C., S. J. Elledge, J. H. Krueger, and G. C. Walker.** 1985. Site-directed insertion and deletion mutagenesis with cloned fragments in *Escherichia coli*. *J. Bacteriol.* **161**:1219–1221.
  28. **Xu, Z., F. E. de Windt, and C. van der Drift.** 1995. Purification and characterization of allantoinase from *Bacillus fastidiosus*. *Arch. Biochem. Biophys.* **324**:99–104.