JOURNAL OF BACTERIOLOGY, Nov. 1993, p. 7066–7073
0021-9193/93/217066-08$02.00/0

Vol. 175, No. 21

purU, a Source of Formate for *purT*-Dependent Phosphoribosyl-*N*-Formylglycinamide Synthesis†

PETER L. NAGY,[1] GEORGE M. McCORKLE,[2]‡ AND HOWARD ZALKIN[1]*

Department of Biochemistry, Purdue University, West Lafayette, Indiana 47907,[1] and Department of Biology, Yale University, New Haven, Connecticut 06520[2]

Received 2 July 1993/Accepted 1 September 1993

A gene designated *purU* has been identified and characterized. *purU* is adjacent to *tyrT* at min 27.7 on the *Escherichia coli* chromosome. The gene codes for a 280-amino-acid protein. The C-terminal segment of PurU from residues 84 to 280 exhibits 27% identity with 5′-phosphoribosylglycinamide (GAR) transformylase, the product of *purN*. Primer extension mapping and assays of *lacZ* in a promoter probe vector identified two promoters giving mono- and bi-cistronic *purU* mRNA. Neither mRNA was regulated by purines. Mutations in either of two pairs of genes are required to block synthesis of 5′-phosphoribosyl-*N*-formylglycinamide (FGAR) from GAR: *purN purT* (*purT* encodes an alternative formate-dependent GAR transformylase) or *purN purU*. On the basis of the growth of *purU*, *purN*, and *purU purN* mutants, it appears that PurU provides the major source of formate for the *purT*-dependent synthesis of FGAR.

There are two transformylation reactions in the de novo purine nucleotide biosynthetic pathway. In step 3, the *purN* gene product 5′-phosphoribosylglycinamide (GAR) transformylase catalyzes the transfer of a formyl group to GAR, resulting in synthesis of 5′-phosphoribosyl-*N*-formylglycinamide (FGAR). In step 9, the *purH* gene product 5′-phosphoribosyl-4-carboxamide-5-aminoimidazole (AICAR) transformylase formylates AICAR, resulting in 5′-phosphoribosyl-4-carboxamide-5-formylaminoimidazole (FAICAR) (18). Both enzymes use 10-formyltetrahydrofolate (formyl-FH_4) as the one-carbon donor. There are two major sources of formyl-FH_4 in *Escherichia coli* (for a review, see reference 32). The main one is the serine hydroxymethyltransferase-catalyzed conversion of serine to glycine plus 5,10-methylene FH_4. The *glyA* gene encodes serine hydroxymethyltransferase. 5,10-Methylene FH_4 is converted in two steps into formyl-FH_4. The glycine cleavage enzyme system provides a second source of 5,10-methylene FH_4 in the breakdown of glycine to NH_3 plus CO_2. *E. coli* also can use formate for the synthesis of FGAR, although without formyl-FH_4 synthetase activity there can be no conversion of formate to formyl-FH_4 (4). This observation and the fact that *purN* purine auxotrophs have not been isolated suggest that there is an additional enzyme in *E. coli* that uses formate directly for synthesis of FGAR.

During the preparation of this report, a paper describing the identification of this predicted formate-dependent GAR transformylase (20) appeared. This enzyme, named GAR transformylase T, is encoded by *purT*. Nygaard and Smith (20) noted that pyruvate formate lyase (26) could provide formate for this reaction during anaerobic growth. A source of formate in aerobic *E. coli* is not known. In this paper we describe a gene, *purU*, that appears to provide the major source of formate used by GAR transformylase T.

* Corresponding author. Electronic mail address: BITNET Zalkin@PURCCVM.

† Journal paper 13926 of the Purdue University Agricultural Experiment Station.

‡ Deceased May 6, 1993. We dedicate this paper to his memory.

MATERIALS AND METHODS

Bacterial strains. Strains, bacteriophages, and plasmids used in this study are listed in Table 1.

Media and growth conditions. LB and 2× YT were used as rich media (16). Minimal growth medium contained salts (37) supplemented with 0.5% glucose and 1 μg of thiamine per ml. Solid media contained 1.5% agar. Supplements of adenine at 100 μg/ml, amino acids at 0.5 mM, and sodium formate at 1 mM were used as noted. Ampicillin was added at 50 μg/ml for plasmid selection and 20 μg/ml for lysogen selection. Kanamycin and chloramphenicol were used at 25 μg/ml where indicated.

Construction of translational fusions for promoter analysis. Promoter fragments were amplified by polymerase chain reaction (PCR) from *E. coli* genomic DNA with primers that created *Eco*RI and *Bam*HI adapters at the 5′ and 3′ ends, respectively, of the amplified product. In this paper, restriction sites in primers are underlined and nucleotides corresponding to *E. coli* DNA sequences are in boldface type. All primers were named with the number of the 5′ nucleotide in the *E. coli* sequence. The following primers were used for PCR cloning of fragments A to D (see Fig. 3).

Primer 462 (5′TCATGAATTC**CACAGCGGCCGTTACGTGCC**3′) was the 5′ primer used for isolation of fragments A and B, primer 1 (5′G**GAATTCCAGCGCAGATAATCGATCGCATG**3′) was the 5′ primer used for isolation of fragments C and D, primer 1092 (5′TGATGGATCC**CCAAGGCAATGCGCTTC**3′) was the 3′ primer used for isolation of fragment A, primer 810 (5′TGATGGATCC**TGGAGTGAATGCATTGCTGG**3′) was the 3′ primer used for isolation of fragments B and C, and primer 564 (5′TGATGGATCC**CGATATTGTCGGCGTCCTGC**3′) was the 3′ primer used for isolation of fragment D. Fragments A, B, C, and D were cloned into the *Eco*RI and *Bam*HI polylinker sites of pMLB1034 (30) to create in-frame fusions with the ninth codon of *'lacZ*. Since primer 564 was designed from the *Shigella flexneri* sequence before a sequencing mistake resulting in a frameshift mutation was discovered, the initial fusion plasmid had no β-galactosidase activity. To restore the correct reading frame, the *Bam*HI site was cut and religated after the cohesive ends were filled in.

Fusions to *'lacZ* were crossed onto λRZ5 for single-copy chromosomal insertion (22).

TABLE 1. Bacteria, bacteriophages, and plasmids used in this work

Strain, phage, or plasmid	Genotype or description	Source or reference
E. coli		
MC4100	F⁻ *araD139* Δ(*argF-lac*)U169 *rpsL150 thiA1 relA1 deoC1 ptsF25 flbB5301*	Laboratory stock
PLN100	MC4100 *purU*::*kan*	This work
PLN101	MC4100 *purN*::*cat*	This work
PLN102	MC4100 *purN*::*cat purU*::*kan*	This work
W3110	F⁻ λ⁻ IN(rrD-rrE)	Laboratory stock
PS2412	W3110 Δ*lac glyA*::Tn5	M. Levinthal[a]
JC7623	F⁻ *thr-1 leuB6 proA2 sbcB15 hisG4 recB21 recC22 argE3 thi-1 ara-14 lacY1 galK2 xyl-5 mtl-1 rpsL31 tsx-33* λ⁻ *supE44*	Laboratory stock
JM101	Δ(*lac-pro*) *thi-1 rpsL hsdR4 endA sbcB supE44* F′ *traD36 proA*⁺*B*⁺ *lacI*^{q} Δ(*lacZ*)M15	Laboratory stock
Bacteriophages		
P1 *vir*	Bacteriophage P1, lytic	Laboratory stock
λRZ5	λ′bla′lacZY⁺, Lac⁻ Ap^s	Laboratory stock
λRZ5A	λRZ5 with fragment A	This work
λRZ5B	λRZ5 with fragment B	This work
λRZ5C	λRZ5 with fragment C	This work
λRZ5D	λRZ5 with fragment D	This work
Plasmids		
pMLB1034	*′lacZY′* fusion vector, Ap^r	30
pPLN1a	pMLB1034 with fragment A	This work
pPLN1b	pMLB1034 with fragment B	This work
pPLN1c	pMLB1034 with fragment C	This work
pPLN1d	pMLB1034 with fragment D	This work
pUC18CML	Plasmid with *cat* cassette	36
pUC4K	Plasmid with *kan* cassette	Pharmacia
pGS1	Plasmid containing *glyA*⁺	33
pTrc99A	Expression vector	3
pPLN2	pTrc 99A containing *glyA*⁺	This work
Bluescript SK⁺	Cloning vector	Laboratory stock
Bluescript SK	Cloning vector	Laboratory stock

[a] Purdue University, West Lafayette, Ind.

Assay of β-galactosidase. Cultures were grown to mid-log phase, and β-galactosidase was assayed in cells with the chloroform-sodium dodecyl sulfate (SDS) lysis procedure (17).

Disruption of *purU* and *purN* genes. A kanamycin cassette was cut out from plasmid pUC-4K with *Pst*I and ligated into the *Nsi*I site of the cloned *purU* fragment A (see Fig. 3). This insertion disrupts the translational start site of *purU* at position 797. The *purN* gene was amplified by PCR from *E. coli* chromosomal DNA with the 5′ primer 1788 (5′G**GAATTC**CGAACAACGCGTGGTTATCG3′) and the 3′ primer 2577 (5′CG**GGATCC**CGAGCAATATTGGCAGATGTCC3′), which have *Eco*RI and *Bam*HI linkers, respectively. *purN* is numbered according to the work of Smith and Daum (31). After being cut with the appropriate enzymes, the fragment was ligated into the polycloning site of Bluescript SK⁺. A chloramphenicol cassette was cut out from plasmid pUC18CML (36) with *Sma*I and *Hin*cII and put into the *Fsp*I site of *purN* at position 2082 of *purMN* (31). Plasmids carrying the disrupted gene fragments were then linearized and transformed into strain JC7623 according to the procedure of Winans et al. (38). Kanamycin- and chloramphenicol-resistant colonies were isolated, and the mutations were transferred to strain MC4100 by P1 transduction (10). A *purU*, a *purN*, and a *purU purN* mutant were obtained. Gene disruptions were checked by Southern blots of DNA isolated from the wild type and from each mutant strain.

Growth rate measurements. Cells were grown overnight in minimal medium supplemented with adenine. After being washed twice with E-salts (37), the cells were diluted 1:20 in minimal medium with appropriate supplements. Cells were grown at 37°C, and growth was monitored with a Klett colorimeter using a 660-nm filter.

Complementations with *glyA*. *glyA* was amplified by PCR from plasmid pGS1 (33) with the 5′ primer 351 (5′**GATGC CCATGG**TAAAGCGT3′) and the 3′ primer 1689 (5′**GTTTTC TAGA**CCGGATAAG3′) and cloned into the *Nco*I-*Xba*I polycloning sites of expression vector pTrc99A (3) to give pPLN2. Expression was checked by complementation of *glyA* strain PS2412. Plasmid pPLN2 was transformed into *purU* mutant PLN100.

Purification of PurU-LacZ fusion proteins. β-Galactosidase immunoaffinity adsorbant was obtained from Promega and was used according to the recommendations of the manufacturer. The protein, partially purified after immunoaffinity chromatography, was further purified by electrophoresis on an SDS–10% polyacrylamide gel, transferred to a polyvinylidene difluoride-type membrane, and lightly stained with Coomassie blue (11). The protein band was cut out and subjected to amino acid sequencing by the Purdue Laboratory for Macromolecular Structure.

DNA sequence analysis. DNA sequences were determined by the method of Sanger et al. (25). Recombinant Bluescript phagemids SK⁺ and SK⁻ were used as templates with T7, T3, and custom-made oligonucleotide primers dictated by the available sequence. The sequence of the *purU* locus from nucleotides 1 to 1636 (see Fig. 1A) was determined on both strands.

RNA analysis. RNA was isolated as described previously (28). The method for primer extension has been described

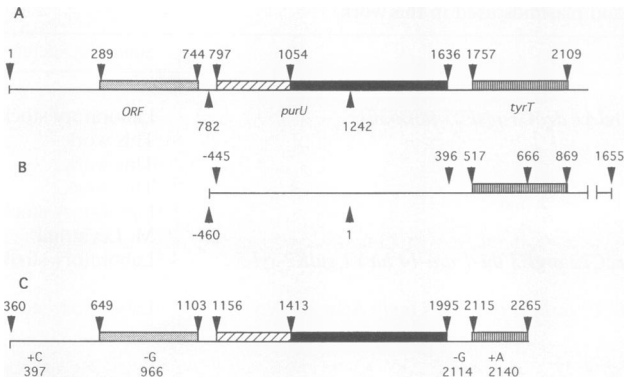

FIG. 1. Organization of the *purU* locus. (A) *E. coli purU* locus. Numbering is arbitrarily begun from the 5′ end of the sequence. Positions 289 to 744 (stippled bar) constitute an unidentified ORF that forms a two-gene operon with *purU*. *purU* extends from position 797 to 1636; the black bar indicates a sequence that is homologous to *purN*. The upstream region of *purU* (vertically hatched bar) is not similar to other known sequences. (B) Clone p9313B (13). Bars are as defined for panel A. The previously sequenced region of p9313B is numbered from 1 to 1665 as given in GenBank (accession number, X04168). The 5′ end of *E. coli* DNA in the insert is assigned −460. The PurU coding sequence starts at −445 and ends at 396. (C) *S. flexneri tyrT* clone (8). Bars are as defined for panel A. The nucleotide sequence, numbered 5′ to 3′ on the lower strand in GenBank (accession number, X66849), has been renumbered 5′ to 3′ on the top strand to permit the alignment shown in the figure. Nucleotides 2115 to 2265 (numbered 1 to 150 on the lower strand) were assigned to *tyrT* in X66849, but other regions were not identified. Insertions (+) and deletions (−) relative to the *E. coli* sequence are indicated below the line. The assignments of *purU* and the ORF are derived from the present analysis.

previously (23). Primers 810 (described above) and 327 (5′**TACTCGACAGCACTACCAC**3′) were used.

DNA manipulations. Procedures for the isolation of DNA, restriction enzyme digestion, ligation, and bacterial transformation have been described previously (12).

Computer analysis. All computer analyses were done with the University of Wisconsin Genetics Computer Group software package (6).

Nucleotide sequence accession number. The nucleotide sequence of the *E. coli purU* locus from positions 1 to 1757 (see Fig. 1A) has been submitted to GenBank (accession number, L20251). Previous submissions include *E. coli tyrT* from positions 1 to 1655 (see Fig. 1B) (accession number, X04168) and *S. flexneri tyrT* from positions 1 to 2265 (see Fig. 1C) (accession number, X66849).

RESULTS

Identification of *purU*. A 517-nucleotide sequence upstream of *tyrT* has been reported (13) and is stored in the GenBank data base under accession number X04168. Computer analysis of this sequence identified a gene related to *purN* (1, 31). An open reading frame (ORF) numbered from nucleotide 1 to 396 (Fig. 1B) showed 31% identity at the amino acid level to the C-terminal portion of *E. coli purN*-encoded GAR transformylase. Accordingly, the nucleotide sequence was extended 460 nucleotides upstream to the 5′ end of the *E. coli* insert in plasmid p9313B (Fig. 1B). An ORF that could encode a protein of 280 amino acids was found to extend from position −445 to +396. The translated amino acid sequence, designated PurU, is given in Fig. 2. This sequence is identical in 53 of 199 positions with GAR transformylase in the area where they overlap.

A data base search found a virtually identical nucleotide sequence upstream of *S. flexneri tyrT* (Fig. 1C) (8). The *S. flexneri* gene assignments shown in Fig. 1C other than *tyrT* were not previously made. Since there was 98% nucleotide sequence identity in the *purU* region, PCR primers based on the *Shigella* sequence were used to clone the homologous upstream segment from *E. coli* DNA. In this way, the *E. coli purU* upstream sequence was extended by 782 bp. A 152-amino-acid ORF was identified upstream of *purU* (Fig. 1A). The amino acid sequence is not similar to that of any other protein in the data base except for a 16-amino-acid identity, residues 152 to 170, with *E. coli secA* protein residues 881 to 899 (27). The function of this segment of SecA is not known. Insertion of a guanine at nucleotide 966 in the *Shigella* sequence results in a homologous ORF in *S. flexneri* (Fig. 1C). Insertions and deletions in the *S. flexneri* sequence compared with the *E. coli* sequence are outlined in Fig. 1C.

The *E. coli* and *S. flexneri* PurU sequences are aligned in Fig. 2 with *E. coli* GAR transformylase. The *E. coli* and *S. flexneri* sequences differ in only 2 of 280 amino acids. According to this alignment with PurN, PurU corresponds to a GAR transfor-

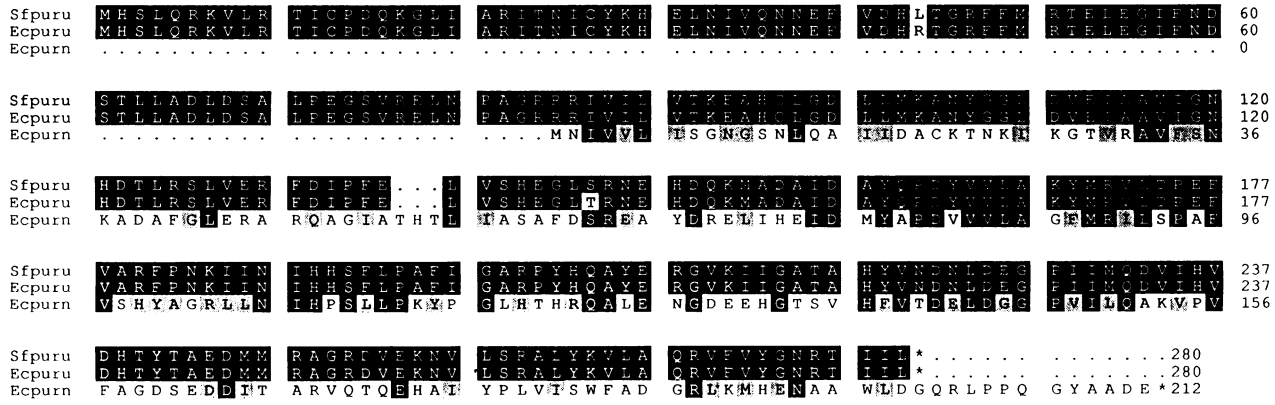

FIG. 2. Amino acid sequence alignments of *E. coli* and *S. flexneri* PurU with *E. coli purN*-encoded GAR transformylase. Sfpuru, *S. flexneri* PurU; Ecpuru, *E. coli* PurU; Ecpurn, *E. coli* GAR transformylase. Amino acids identical to (white on black), similar to (black on grey), and different from (black on white) those of Sfpuru are indicated.

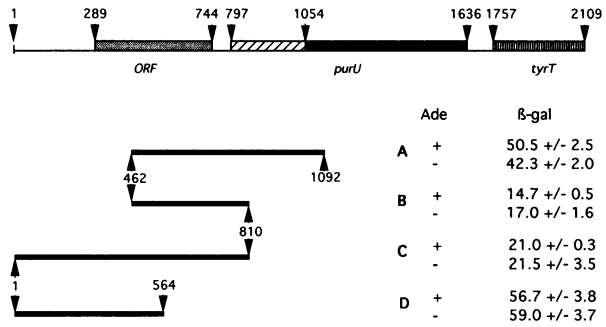

FIG. 3. Expression of *purU*- and ORF-'*lacZ* fusions. Bars are as defined for Fig. 1A. DNA fragments, designated A to D, were ligated into a *lacZ* reporter plasmid and integrated into the chromosome. Cells were grown with or without 100 μg of adenine (Ade) per ml. β-Galactosidase (β-gal) activity from permeabilized cells is given in Miller units. All values are means of triplicate assays ± standard deviations.

mylase with an 84-amino-acid N-terminal extension. In the region of overlap, there are 27% identity and 48% similarity between the PurU and GAR transformylase sequences. This similarity suggests that *purU* encodes an enzyme with at least some of the functions of GAR transformylase N. There was no similarity of the N-terminal 84-amino-acid sequence to other proteins in the data base.

In order to verify the PurU translation start site inferred from the nucleotide sequence, a *purU'-'lacZ* translational fusion was constructed. Fragment A, corresponding to nucleotides 462 to 1092 (Fig. 3), was ligated into a *lacZ* reporter plasmid giving a fusion to *lacZ* codon 9. The PurU–β-galactosidase chimeric protein was purified by immunoaffinity chromatography and SDS-polyacrylamide gel electrophoresis. N-terminal analysis gave the sequence Met-His-Ser-Leu-Gln-Arg, which corresponds to residues 1 to 6 derived from the gene (Fig. 2).

***purU* promoter and gene expression.** A series of DNA fragments from the *purU* upstream region were ligated into the *lacZ* reporter vector to give in-frame *purU'-'lacZ* and ORF-'*lacZ* translational fusions with codon 9 of β-galactosidase as outlined in Fig. 3. Plasmid pPLN1a has the coding sequence for the first 99 amino acids of PurU, whereas pPLN1b and pPLN1c have the coding sequence only for the first 5 amino acids. In pPLN1d, the first 91 amino acids of the ORF are fused to β-galactosidase. Assays of these fusions from multicopy plasmids gave the following values for β-galactosidase (in Miller units): A, 1,580; B, 415; C, 680; D, 1,890; vector control, 50. These results indicate promoter function in the *purU* upstream region and in the region upstream from the ORF.

To localize the *purU* promoter more precisely, primer extension mapping was performed on RNA isolated from wild-type strain W3110 and from W3110 bearing multicopy *purU'-'lacZ* plasmid pPLN1a. By using primer 810, two transcript 5′ ends were identified at nucleotide 667 and at a position near the top of the gel in the vicinity of nucleotides 240 to 250; they are marked P1 and P2, respectively (Fig. 4B). Transcript P1 was obtained from reverse transcription of RNA from cells bearing plasmid pPLN1a (Fig. 4B, lane 1), as well as from reverse transcription of RNA from chromosomal *purU* (Fig. 4B, lane 2). There were four closely spaced bands in the P1 region. The most intense band, corresponding to an endpoint at G-667, was arbitrarily chosen as the transcription start site. The P2 transcript can be seen only in lane 2, of Fig. 4B, since fragment A of pPLN1a ends at nucleotide 462 and all

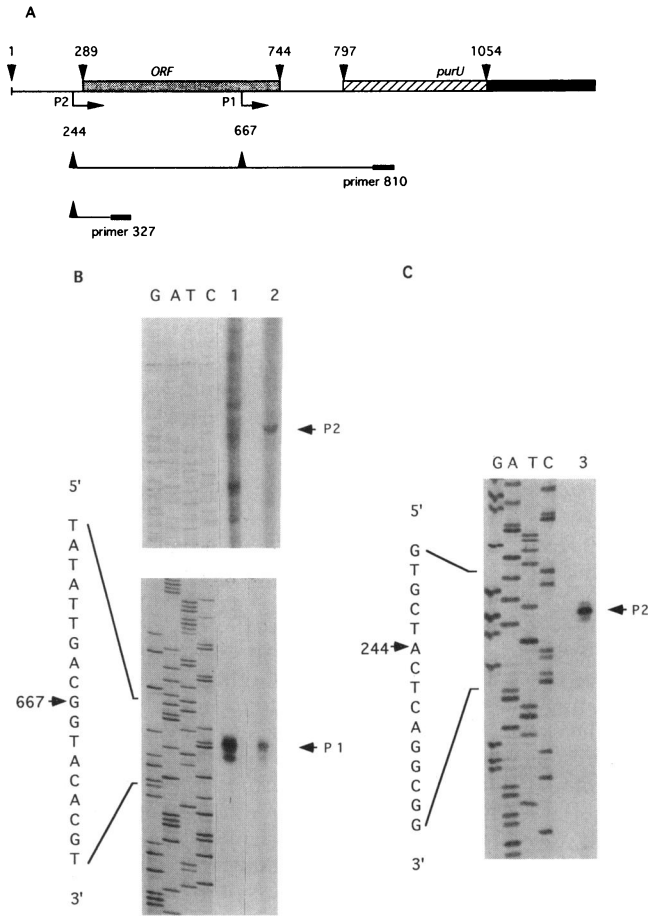

FIG. 4. Localization of promoters by primer extension mapping. (A) Diagram showing sites for annealing of primers and location of promoters P1 and P2. Bars are as defined for Fig. 1A. (B) Reverse transcription from primer 810. Lane 1, RNA from strain W3110/pPLN1a; lane 2, RNA from strain W3110. In the upper section, all lanes are from the upper part of one autoradiogram. In the lower part, the original lane 1 was replaced by a shorter exposure of the same gel. (Lane 2 is from the same film in both panels). A sequencing ladder of the template strand from plasmid pPLN1a using the same primer indicates that transcription starts at G-667 and around positions 250 to 240. The upper section of the sequencing ladder shows the vector sequence following nucleotide 462 in plasmid pPLN1a. The sequence written to the left is complementary to the template strand. (C) Reverse transcription from primer 327. RNA was prepared from strain W3110/pPLN1d (lane 3). The sequencing ladder is for the template strand, and the sequence written to the left is for the complementary strand.

bands in the upper section of the gel in lane 1 are nonspecific and result from vector sequences. The intensities of bands are similar for P1 and P2 in Fig. 4B, lane 2. To more precisely map the 5′ end of the mRNA transcribed from upstream promoter P2 in the region from nucleotide 240 to 250, primer 327 (Fig. 4A) was used for reverse transcription of RNA from cells bearing plasmid pPLN1d. The major primer extension band P2 aligns between A-244 and C-245 (Fig. 4C). The 5′ end has been assigned to A-244, since most *E. coli* transcripts start with a purine (24). These results thus support the notion of two *purU* transcripts, a bicistronic ORF-*purU* mRNA and a monocistronic *purU* mRNA, synthesized from promoters marked P2

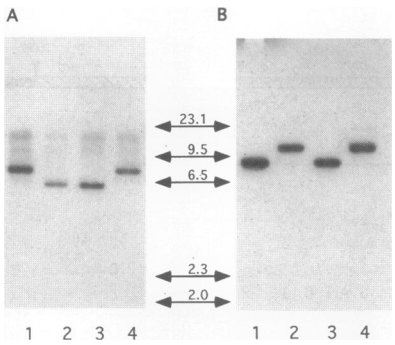

FIG. 5. Southern blots of chromosomal *purU* and *purN* disruptions. (A) *Pst*I-digested chromosomal DNA probed with *purU*$^+$ fragment A; (B) *Sph*I-digested DNA probed with a *purN* fragment. Lanes: 1, *purU*; 2, *purN*; 3, wild type; 4, *purU purN*. Size markers in kilobases are shown.

and P1, respectively, in Fig. 4A. A potential sigma 70 promoter sequence, (−33) TTGAAA-N_{18}-TATATT (−4), is present upstream of start site P1. There is no obvious sigma 70 promoter sequence upstream of start site P2.

In order to determine if *purU* expression is regulated by purines in parallel with expression of other *pur* regulon genes, DNA fragments A to D (Fig. 3) fused to *lacZ* were integrated into the chromosome and β-galactosidase activity was assayed. The results shown in Fig. 3 support the assignments of promoters P1 and P2. *purU'-'lacZ* expression from fragments A and B was driven by promoter P1. Although higher reporter activity was obtained from P1 in fragment A than from the same promoter in fragment B, the different PurU junctions preclude direct comparison of expression from these fusions. Fragment D provides evidence for promoter P2 function in an *orf'-'lacZ* fusion. Fragment C results in a *purU'-'lacZ* fusion in which expression is driven by promoters P1 and P2. Expression from P1 plus P2 in fragment C cannot be directly compared with that from P2 in fragment D because of the different junctions to *'lacZ*. Since the *purU'-'lacZ* junctions are identical in fusions with fragments B and C, it should be reliable to compare expression from promoter P1 in fragment B with that from P1 plus P2 in fragment C. The results in Fig. 3 show marginally higher expression from P1 plus P2 (fragment C) than from P1 (fragment B). This small increment in expression due to P2 is at variance with the primer extension mapping results in Fig. 4B, lane 2, which show that comparable amounts of mRNA are derived from P2 and from P1. The basis for the disparity between the strengths of promoters P1 and P2 in the two experiments is not known. Finally, data in Fig. 3 show that addition of adenine to cells did not repress expression of *purU*, in contrast to the results for other *pur* regulon genes (7). This is in agreement with the absence of operator sites for PurR near the P1 and P2 promoters.

Construction and analysis of *purN* and *purU* mutants. *purN* and *purU* mutants were constructed in order to investigate the role of PurU. Cloned *purN* was disrupted with a gene for chloramphenicol resistance, and *purU* was disrupted with a kanamycin marker. Each of the disrupted genes was used to replace the corresponding chromosomal copy. Southern blots of chromosomal DNA were used to verify the gene disruptions (Fig. 5). The Southern blot in Fig. 5A shows the disruption by the 1.3-kb kanamycin cassette in the single *purU* mutant (lane 1) and the double *purU purN* mutant (lane 4). Figure 5B shows disruption of *purN* in the single mutant (lane 2) and the double *purU purN* mutant (lane 4).

Growth phenotypes of the mutants were determined on minimal agar medium and are given in Table 2. The *purN* and *purU* single mutants were prototrophic, whereas the *purN purU* double mutant required purine or formate supplementation for growth. This establishes a role for *purU* in de novo purine nucleotide synthesis. Addition of amino acids and adenine had different effects on the mutants. Growth of the *purN* mutant strain was stimulated by adenine, formate, methionine, or histidine, while growth in the absence of adenine was strongly inhibited by glycine or threonine. The inhibitory effect of threonine is likely due to its conversion to glycine (19). This inhibition was completely overcome by formate. Of the 20 amino acids, only glycine, threonine, histidine, and methionine affected the growth rate of the *purN* mutant.

Growth of the *purU* mutant was retarded by adenine and inhibited by adenine plus either methionine or histidine. The *purU purN* double mutant, which required adenine for growth, was likewise inhibited by methionine or histidine. This growth inhibition was reversed by glycine. Since the same inhibition and reversal of inhibition were observed in both the *purU* mutant and the *purU purN* double mutant, the effects are clearly due to the lack of PurU. Since growth inhibition of the *purU* mutant by adenine and methionine was possibly due to repression of *glyA* (5, 32, 34) with resultant starvation for glycine, we examined the effect of *glyA* under the control of the *lac* promoter on a multicopy plasmid. Overexpression of *glyA* largely prevented growth inhibition of the *purU* mutant by adenine and methionine (data not shown). Other than the responses shown in Table 2, all other amino acids were without effect.

TABLE 2. Comparison of growth rates on agar plates

Mutant gene(s)	Adenine concn (μg/ml)	Growth rate[a] in medium supplemented with:							
		No supplement	Gly	Formate	Gly + formate	Met	Gly + Met	His	Gly + His
purN	100	++	++	++	++	++	++	++	++
	0	+	−	++	++	++	−	++	−
purU	100	+	++	+	++	−	++	−	++
	0	++	++	++	++	++	++	+	+
purU purN	100	+	++	+	++	−	++	−	++
	0	−	−	+	++	−	−	−	−

[a] Growth rate, compared with that of the wild-type strain MC4100, was estimated after 24 h according to the colony size: ++, same rate as the wild type; +, approximately half the rate of the wild type; −, little or no growth.

TABLE 3. Growth rates of *E. coli* strains[a]

Supplement to culture	Doubling time (min) for strain		
	PLN100 (*purU*)	PLN101 (*purN*)	PLN102 (*purN purU*)
None	130	65	>600
Adenine	180	65	220
Formate	180	65	220
Glycine	100	>600	>600
Formate + glycine	80	65	80
Adenine + glycine	80	65	80
Adenine + methionine	>600	65	>600
Adenine + methionine + glycine	70	65	70

[a] Cells were grown in minimal medium. The doubling time of the MC4100 parental strain was approximately 60 min under all growth conditions used here.

Growth responses in liquid medium were similar to those on agar plates. Data in Table 3 are representative for growth in liquid medium. They show the growth inhibition of the *purN* mutant by glycine and the reversal of this inhibition by formate. The growth rate with glycine plus formate was similar to that with adenine. The inhibitory effects of glycine and *purU* disruption on the *purN* mutant were similar, and in both cases growth was restored by formate to the rate seen in the presence of adenine. These data and those in Table 2 lead to the conclusions that formate is required as the one-carbon donor for PurT-dependent synthesis of FGAR and that formate is produced in a PurU-dependent reaction. Glycine in some way interferes with this production of formate and inhibits growth of the *purN* mutant.

Similar growth phenotypes were reported for a *purN* mutant by Nygaard and Smith (20). The reversal by formate of growth inhibition by glycine was attributed by these authors to an effect on the *purT*-dependent conversion of GAR to FGAR.

DISCUSSION

Formyl-FH_4 is used for formylation of GAR to FGAR in step 3 of de novo purine nucleotide synthesis and for synthesis of FAICAR from AICAR at step 9. These reactions account for incorporation of one-carbon units into C-8 and C-2 of the purine ring, respectively. In wild-type *E. coli* grown in minimal medium, the C-2 and C-8 carbons of the purine ring are derived entirely from serine and glycine via formyl-FH_4 (4). However, formate added to the medium contributes approximately 14% of the purine ring C-8 carbon. The utilization of formate as a one-carbon donor for purine synthesis does not depend on FH_4. Since there is no known route for attachment of formate to FH_4 in *E. coli* (4), there must be an enzyme able to catalyze a formate-dependent synthesis of FGAR from GAR. This enzyme, GAR transformylase T, was recently identified (20). Surprisingly, PurT was reported to have 27% amino acid sequence similarity with PurK but none with PurN (14).

In this paper, we describe the identification of a gene, *purU*, that is required for production of formate for GAR transformylase T under conditions of aerobic growth. This identification is based on mutant isolation and growth analysis. Whereas *purN* and *purU* single mutants are prototrophic, a *purN purU* double mutant has a purine growth requirement that can be satisfied by formate. The *purU*-dependent production of formate appears to be inhibited by glycine in aerobic cells, thus explaining the growth inhibition of the *purN* mutant by glycine and the reversal of this inhibition by formate (Tables 2 and 3) (20). Glycine does not inhibit the growth of a *purN* mutant under anaerobic conditions (20), indicating an alternative anaerobic source of formate for purine biosynthesis.

Pyruvate formate lyase, encoded by *pfl*, is a central enzyme of anaerobic glucose metabolism (9), and catalyzes a nonoxidative cleavage of pyruvate to acetyl coenzyme and formate (26). The *pfl* gene is subject to a 10- to 12-fold anaerobic induction. Although aerobically grown cells contain an inactive form of pyruvate formate lyase, activity has not been reported and is unlikely in the presence of air. Activation of pyruvate formate lyase requires an activase-dependent generation of a stable glycyl free radical at position 734, a strictly anaerobic process.

How does PurU produce formate? Although the enzymatic activity was not characterized, structural evidence suggests that PurU converts a one-carbon FH_4 metabolite, presumably formyl-FH_4, to formate and FH_4. First, PurU exhibits 27% amino acid sequence identity with *purN*-encoded GAR transformylase. This implies structural similarity of at least some of the functional domains in the two enzymes. Second, comparison of PurU with the X-ray structural model of an *E. coli* GAR transformylase–GAR-5-deaza-5,6,7,8-tetrahydrofolate (5dTHF) ternary complex (2) suggests that PurU has retained a binding site for formyl-FH_4 but not for GAR. In the X-ray structure of the ternary complex, there are six H bonds to the bicyclic ring of the FH_4 analog (2). These interactions are with the α-chain backbone of Arg-90, Leu-92, Thr-140, Asp-141, and Asp-144. In addition, five amino acid side chains are thought to be close enough to the formyl group of formyl-FH_4 to participate in catalysis. These are Asn-106, His-108, Ser-110, His-137, and Asp-144. All of these residues except for Thr-140 are conserved in PurU. On the other hand, there are several GAR transformylase N hydrogen bond interactions with GAR. These include Gly-11, Ser-12, Asn-13, Glu-170, and Glu-173. Of these residues that interact with GAR in the GAR transformylase N structure, only Glu-173 is conserved in PurU. These considerations thus lead to the suggestion that PurU produces formate from formyl-FH_4. Other routes of formate production during aerobic growth, if they exist, must be minimal, because a mutation in *purU* severely limits growth in the absence of *purN* function.

We have considered three problems with the proposal that PurU produces formate from formyl-FH_4 for use by GAR transformylase T. First, it is unexpected that PurU should require an N-terminal 84-amino-acid addition relative to GAR transformylase N for the simple steps of formyl-FH_4 binding and hydrolysis. We therefore consider that our present understanding of PurU may be quite limited. Second, if formate is derived from formyl-FH_4, why is the GAR transformylase step of the purine pathway less sensitive than the AICAR transformylase step to sulfonamide inhibition (21, 29, 35)? Previous explanations have assumed an alternative FH_4-independent GAR transformylase (see reference 20, for example). Although the answer to this question is not known and must await characterization of the enzyme, one possibility is that PurU has a higher affinity for formyl-FH_4 than does AICAR transformylase. Third, such a reaction is wasteful and could deplete the cell of an important one-carbon pool needed for biosynthesis. In the discussion that follows, we consider the possibility that PurU functions to regulate the C_1FH_4 pool.

In considering additional roles for *purU*, we are struck by the sensitivity of *purU* single and *purU purN* double mutants to growth inhibition by adenine plus methionine, or histidine (Tables 2 and 3). Earlier, Michelsen et al. (15) inadvertently made a *purU tyrT* double mutation although they thought that only *tyrT* was mutated. They observed the same growth properties in their *tyrT purU* strain as those seen in our *purU* mutant

and demonstrated that growth inhibition by adenine plus methionine was due to insufficient glycine for protein synthesis. The fact that multicopy *glyA* largely prevented growth inhibition of the *purU* mutant by adenine plus methionine confirms that the growth inhibition results from starvation for glycine. Since this growth inhibition is observed only in the *purU* mutants, PurU may be necessary to consume formyl-FH_4 and balance the C_1FH_4-to-glycine ratio. Conversely, excess glycine can inhibit the production of formate (Tables 2 and 3), thus conserving formyl-FH_4, and balance the glycine-to-C_1FH_4 ratio. These considerations suggest a possible role for *purU* in regulating the one-carbon pool. Purification and characterization of the enzyme is required to help clarify some of these issues.

ACKNOWLEDGMENTS

We thank Barbara Bachmann for advice on *E. coli* purine gene nomenclature. We also thank Gary Sawers, Rowena Matthews, and Laszlo Csonka for helpful discussions on one-carbon metabolism. We especially thank Ariane Marlewski, Pennsylvania State University, for reading an earlier version of the manuscript and for communicating unpublished information about GAR transformylase T. This work was initiated following discussions with our colleague George McCorkle, who noted an incomplete ORF upstream of *tyrT* having similarity to *purN*. We are saddened that Dr. McCorkle did not live to enjoy our present understanding of *purU*.

This research was supported by Public Health Service grant GM24658. Computer facilities were supported by NIH grant AI27713. The Purdue Laboratory for Macromolecular Structure received support from the Diabetes Research and Training Center (NIH P60 DK20542).

REFERENCES

1. **Aimi, J., H. Qui, J. Williams, H. Zalkin, and J. E. Dixon.** 1990. *De novo* purine nucleotide biosynthesis: cloning of human and avian cDNAs encoding the trifunctional glycinamide ribonucleotide synthetase-aminoimidazole ribonucleotide synthetase-glycinamide ribonucleotide transformylase by functional complementation in *E. coli*. Nucleic Acids Res. **18:**6665–6672.
2. **Almassy, R. J., C. A. Janson, C. C. Kan, and Z. Hostomska.** 1992. Structures of apo and complexed *Escherichia coli* glycinamide ribonucleotide transformylase. Proc. Natl. Acad. Sci. USA **89:** 6114–6118.
3. **Amann, E., B. Ochs, and K. J. Abel.** 1988. Tightly regulated *tac* promoter vectors useful for the expression of unfused and fused proteins in *E. coli*. Gene **69:**301–315.
4. **Dev, I. K., and R. J. Harvey.** 1978. Sources of one-carbon units in the folate pathway of *Escherichia coli*. J. Biol. Chem. **257:**1980–1986.
5. **Dev, I. K., and R. J. Harvey.** 1984. Role of methionine in the regulation of the synthesis of serine hydroxymethyltransferase in *Escherichia coli*. J. Biol. Chem. **259:**8402–8406.
6. **Genetics Computer Group.** 1992. Program manual for the GCG package, version 7.2, November 1992. Genetics Computer Group, Madison, Wis.
7. **He, B., A. Shiau, K. Y. Choi, H. Zalkin, and J. M. Smith.** 1990. Genes of the *Escherichia coli pur* regulon are negatively controlled by a repressor-operator interaction. J. Bacteriol. **172:**4555–4562.
8. **Hromockyj, A. E., S. C. Tucker, and A. T. Maurelli.** 1992. Temperature regulation of Shigella virulence: identification of the repressor gene *virR*, an analogue of *hns*, and partial complementation by tyrosyl transfer RNA ($tRNA_1^{tyr}$). Mol. Microbiol. **6:**2113–2124.
9. **Knappe, J.** 1987. Anaerobic dissimilation of pyruvate, p. 151–155. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
10. **Lennox, E. S.** 1955. Transduction of linked genetic characters of the host by bacteriophage P1. Virology **1:**190–206.
11. **Madsudaira, P.** 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membrane. J. Biol. Chem. **262:**10035–10038.
12. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
13. **McCorkle, G. M., and S. Altman.** 1982. Large deletion mutants of *E. coli* $tRNA_1^{tyr}$. J. Mol. Biol. **155:**83–103.
14. **Meyer, E., N. J. Leonard, B. Chat, J. Stubbe, and J. M. Smith.** 1992. Purification and characterization of the *purE*, *purK*, and *purC* gene products: identification of a previously unrecognized energy requirement in the purine biosynthetic pathway. Biochemistry **31:**5022–5032.
15. **Michelsen, U., M. Bösl, T. Dingermann, and H. Kersten.** 1989. The *tyrT* locus of *Escherichia coli* exhibits a regulatory function for glycine metabolism. J. Bacteriol. **171:**5987–5994.
16. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
17. **Miller, J. H.** 1992. A short course in bacterial genetics. A laboratory manual and handbook for *Escherichia coli* and related bacteria, p. 72–74. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. **Neuhard, J., and P. Nygaard.** 1987. Purines and pyrimidines, p. 445–473. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
19. **Newman, E. B., V. Kapoor, and R. Potter.** 1976. Role of L-threonine dehydrogenase in the catabolism of threonine and synthesis of glycine by *Escherichia coli*. J. Bacteriol. **126:**1245–1249.
20. **Nygaard, P., and J. M. Smith.** 1993. Evidence for a novel glycinamide ribonucleotide transformylase in *Escherichia coli*. J. Bacteriol. **175:**3591–3597.
21. **Rohlman, C. E., and R. G. Matthews.** 1990. Role of purine biosynthetic intermediates in response to folate stress in *Escherichia coli*. J. Bacteriol. **172:**7200–7210.
22. **Rolfes, R. J., and H. Zalkin.** 1988. Regulation of *Escherichia coli purF*. J. Biol. Chem. **263:**19649–19652.
23. **Rolfes, R. J., and H. Zalkin.** 1988. *Escherichia coli* gene *purR* encoding a repressor protein for purine nucleotide synthesis. Cloning, nucleotide sequence and interaction with the *purF* operator. J. Biol. Chem. **263:**19653–19661.
24. **Rosenberg, M., and D. Court.** 1979. Regulator sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13:**319–353.
25. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
26. **Sawers, G., and A. Böck.** 1988. Anaerobic regulation of pyruvate formate-lyase from *Escherichia coli* K-12. J. Bacteriol. **170:**5330–5336.
27. **Schmidt, M. G., E. E. Rollo, J. Grodberg, and D. B. Oliver.** 1988. Nucleotide sequence of the *secA* gene and *secA*(Ts) mutations preventing protein export in *Escherichia coli*. J. Bacteriol. **170:** 3404–3414.
28. **Shimotsu, H., M. I. Kuroda, C. Yanofsky, and D. J. Henner.** 1986. Novel form of transcription attenuation regulates expression of the *Bacillus subtilis* tryptophan operon. J. Bacteriol. **166:**461–471.
29. **Shive, W., W. W. Ackermann, M. Gordon, M. E. Getzendaner, and R. E. Eakin.** 1947. 5(4)-Amino-4(5)-imidazole carboxamide, a precursor of purines. J. Am. Chem. Soc. **69:**725–726.
30. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions, p. 95–96. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
31. **Smith, J. M., and H. A. Daum III.** 1987. Identification and nucleotide sequence of a gene encoding 5′-phosphoribosylglycinamide transformylase in *Escherichia coli* K12. J. Biol. Chem. **262:**10565–10569.
32. **Stauffer, G. V.** 1987. Biosynthesis of serine and glycine, p. 412–418. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.

33. **Stauffer, G. V., M. D. Plamann, and L. T. Stauffer.** 1981. Construction and expression of hybrid plasmids containing the *Escherichia coli glyA* gene. Gene **14:**63–72.
34. **Steiert, J. G., R. J. Rolfes, H. Zalkin, and G. V. Stauffer.** 1990. Regulation of the *Escherichia coli glyA* gene by the *purR* gene product. J. Bacteriol. **172:**3799–3803.
35. **Stetten, M. R., and C. L. Fox, Jr.** 1945. An amine formed by bacteria during sulfonamide bacteriostasis. J. Biol. Chem. **161:** 333–349.
36. **Sweitzer, H. P.** 1990. The pUC18CM plasmids: a chloramphenicol resistance gene cassette for site directed insertion and deletion mutagenesis in *E. coli*. BioTechniques **8:**612.
37. **Vogel, H. T., and D. M. Bonner.** 1956. Acetylornithinase of *Escherichia coli*: partial purification and some properties. J. Biol. Chem. **218:**97–106.
38. **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.