JOURNAL OF BACTERIOLOGY, Jan. 1992, p. 130–136
0021-9193/92/010130-07$02.00/0

Vol. 174, No. 1

Escherichia coli purB Gene: Cloning, Nucleotide Sequence, and Regulation by *purR*†

BIN HE,[1] JOHN M. SMITH,[2] AND HOWARD ZALKIN[1]*

Department of Biochemistry, Purdue University, West Lafayette, Indiana 47907,[1] and Seattle Biomedical Research Institute, Seattle, Washington 98109[2]

Received 26 August 1991/Accepted 1 November 1991

Escherichia coli purB **encodes adenylosuccinate lyase (ASL), the enzyme that catalyzes step 8 in the pathway for de novo synthesis of IMP and also the final reaction in the two-step sequence from IMP to AMP. Gene *purB* was cloned and found to encode an ASL protein of 435 amino acids having a calculated molecular weight of 49,225. *E. coli* ASL is homologous to the corresponding enzymes from *Bacillus subtilis* and chickens and also to fumarase from *B. subtilis*. Gene *phoP* is 232 bp downstream of *purB*. Gene *purB* is regulated threefold by the purine pool and *purR*. Transcriptional regulation of *purB* involves binding of the purine repressor to the 16-bp conserved *pur* regulon operator. The *purB* operator is 224 bp downstream of the transcription start site and overlaps codons 62 to 67 in the protein-coding sequence.**

Adenylosuccinate lyase (ASL), the product of gene *purB* in *Escherichia coli*, catalyzes two reactions in de novo purine nucleotide biosynthesis. In addition to catalyzing step 8 in the pathway to IMP, conversion of phosphoribosylsuccinocarboxamideaminoimidazole to phosphoribosylaminoimidazole carboxamide, ASL also converts adenylosuccinate to AMP (Fig. 1). Gene *purB* has been mapped to 25 min on the *E. coli* chromosome (2), unlinked to the other purine genes.

The genes for de novo purine nucleotide synthesis in *E. coli*, *purF*, *purMN*, *purEK*, *purL*, *purC*, *purHD*, and *guaBA*, are scattered around the chromosome as monocistronic and small polycistronic operons and are negatively controlled by *purR* (6, 15). These genes are part of a *pur* regulon. Nucleotide sequence analysis of *pur* regulon genes has identified a conserved 16-bp sequence generally around the −35 promoter region which serves as a site for binding of the *purR*-encoded repressor. Gene *purB* from *E. coli* has not been cloned, and there are conflicting reports concerning regulation by *purR*. Meng et al. (15) reported 2.5-fold regulation of ASL by *purR*, and He et al. (6) found no regulation by *purR* of a *purB-lacZ* fusion made by insertion of Mud1(Ap *lac*).

Here we report the nucleotide sequence of cloned *E. coli purB*. A binding site for the purine repressor in the protein-coding region is required for threefold regulation by *purR*.

MATERIALS AND METHODS

Bacterial strains, bacteriophage, and plasmids. The strains and plasmids used are listed in Table 1. A λSE6 library (4) of *E. coli* K-12 DNA fragments was obtained from Graham Walker, Massachusetts Institute of Technology, Cambridge. Phage λRZ5 has been described (19).

Media. LB (17) and 2XYT (17) were used as rich media. The minimal growth medium used has been described (27).

Cloning of the *purB* gene. Gene *purB* was isolated by complementation of *purB* strain TX570 by using a λSE6 *E. coli* library (4). Phage DNA prepared (24) from a representative isolate was partially digested with *Sau*3AI and electrophoresed on a 0.5% agarose gel. DNA fragments ranging in size from 1.5 to 6 kb were excised, electroeluted, and ligated into the *Bam*HI site of pUC118. The ligation mixture was used to transform *purB* mutant TX530 to *purB*$^+$.

To clone the upstream region of *purB*, a λ phage lysate was prepared from phage 7F9 of the library of Kohara et al. (8), and DNA was purified (24) as described above. Phage DNA was digested with *Eco*RI and electrophoresed on a 0.8% agarose gel. A 3-kb *Eco*RI fragment was identified by Southern blotting (13) by using a ^{32}P-labeled DNA probe corresponding to the 5′ end of the *purB* gene. The 3-kb *Eco*RI fragment was cloned into pUC118 to give plasmid pBH111.

DNA sequence analysis. A series of overlapping exonuclease III deletions was prepared (7), and nucleotide sequences were determined by the dideoxynucleotide chain termination method of Sanger et al. (22).

Enzyme assay. Cells were grown to the mid-log phase in minimal medium with and without adenine. Cells were disrupted (20) in 0.1 M sodium phosphate, pH 7.0, and β-galactosidase activity was determined by the method of Miller (17). Protein concentration was determined by the method of Lowry (10).

Repressor-operator binding. For gel retardation assays, 10 fmol of a 570-bp *Hin*dIII-*Eco*RI fragment from plasmid pBH113 containing the *purB* operatorlike sequence was incubated with the purified Pur repressor in buffer system II (21) with the hypoxanthine corepressor at 50 μM and samples were resolved by electrophoresis. For DNase I footprinting, a DNA fragment was isolated from plasmid pBH113, labeled at one end with [γ-^{32}P]ATP and T4 polynucleotide kinase, and then purified by electrophoresis and electroeluted from a 5% polyacrylamide gel slice. Repressor-operator complexes were prepared in buffer system I (21) by using 400 ng of the Pur repressor. The conditions used for treatment with DNase I and for electrophoresis have been described (20).

Primer extension mapping. *E. coli* R320(pBH111) was grown in 50 ml of minimal medium without adenine. At a Klett reading of 70, using a 66 filter, the cells were poured onto ice chilled to −20°C and collected by centrifugation for

* Corresponding author.

† Journal paper 13209 from the Purdue University Agricultural Experiment Station.

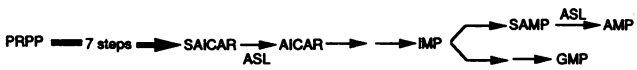

FIG. 1. Outline of the pathway for de novo synthesis of purine nucleotides in *E. coli*. Abbreviations: PRPP, 5-phosphoribosyl-1-pyrophosphate; SAICAR, 5'-phosphoribosylsuccinocarboxamide-aminoimidazole; AICAR, 5'-phosphoribosylaminoimidazole carboxamide; SAMP, 6-succinyladenosine-5'-monophosphate.

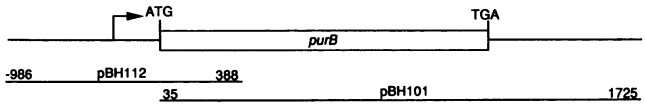

FIG. 2. Scheme showing plasmid clones used to determine the *purB* nucleotide sequence. The transcription start site (+1) is represented by the arrow. Numbering is relative to the transcription start site as given in Fig. 3.

4 min at 3,000 × *g* at 4°C. The cells were then suspended in ice-cold TE (10 mM Tris [pH 7.6], 0.1 mM EDTA), and RNA was isolated by hot-phenol extraction (23). The final RNA pellet was dissolved in water and frozen at −70°C.

To determine the 5' end of *purB* mRNA, two 19-mer oligonucleotide primers, complementary to nucleotides 120 to 138 and 154 to 172, were labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase and annealed to the mRNA. The conditions used for primer extension mapping with avian myeloblastosis virus reverse transcriptase have been described (30).

Plasmid constructions. Plasmid pBH113 was constructed as a *purB* fragment source for gel retardation assay and DNase I footprinting. First, plasmid pBH111 was digested with *Bgl*I and treated with T4 DNA polymerase to make a blunt end. Following this step, the plasmid was digested with *Eco*RI. A 570-bp *Bgl*I-*Eco*RI fragment containing the *purB* operatorlike sequence was subcloned into the *Hin*cII and *Eco*RI sites of pUC118 to yield pBH113. Plasmid pBH112 was constructed by subcloning a 1.3-kb *Kpn*I-*Eco*RI *purB* fragment from pBH111 into pUC118 (Table 1).

For construction of translational and transcriptional *purB-lacZ* fusions, *purB* DNA was obtained from plasmid pBH112 by polymerase chain reaction amplification (18). For translational fusions, the 5' end of *purB* was at nucleotide −986 and contained an *Sma*I site. The 3' end with a *Bam*HI site was at nucleotide 172 for pBLG1 (*purBo*$^-$) or 316 for pBLG2 (*purBo*$^+$). Fragments were ligated into plasmid pMLB1034 to give *purB-lacZ* translational fusions pBLG1 (*purBo*$^-$) and pBLG2 (*purBo*$^+$). In pBLG1 (*purBo*$^-$) and pBLG2 (*purBo*$^+$), *purB* codons 45 and 93, respectively, were fused to codon 8 of *lacZ*. For transcriptional fusions, fragments were ligated into the *Sma*I-*Bam*HI sites of pRS415. Plasmid pLT1 (*purBo*$^-$) contains *purB* nucleotides −986 to −172, pLT2 (*purBo*$^-$) contains nucleotides −986 to 172, pLT3 (*purBo*$^+$) contains nucleotides −986 to 316, and pLT4 (*purBo*$^-$) contains nucleotides −172 to 172. *purB-lacZ* transcriptional and translational fusions were recombined into λRZ5 for insertion into the chromosome in single-copy form (19).

TABLE 1. Strains and plasmids used

E. coli strain	Description	Source or reference
MC4100	Δ(*argF-lac*)*169* Lac$^-$	19
R320	MC4100 *purR320*	19
BH100	MC4100 (λBLG1) Lac$^+$ Ap^r	This work
BH101	MC4100 (λBLG2) Lac$^+$ Ap^r	This work
BH102	MC4100 (λLT2) Lac$^+$ Ap^r	This work
BH103	MC4100 (λLT3) Lac$^+$ Ap^r	This work
BH200	R320 (λBLG1) Lac$^+$ Ap^r	This work
BH201	R320 (λBLG2) Lac$^+$ Ap^r	This work
BH202	R320 (λLT2) Lac$^+$ Ap^r	This work
BH203	R320 (λLT3) Lac$^+$ Ap^r	This work
TX570	*ara* Δ(*lac*) *purB*::Mu Δ*c*I857 *S7*	This work
TX530	*ara* Δ(*lac*) Φ(*purB'-lacZ*$^+$*Y*$^+$::λp1 (*209*, *205*)	28
pMLB1034	'*lacZ* translational fusion vector; Ap^r	24
pRS415	*lacZ* transcriptional fusion vector; Ap^r	25
pUC118	Phagemid cloning vector	26
pBH101	1.7-kb *Sau*3A fragment in *Bam*HI site of pUC118; *purB*$^+$	This work
pBH111	3-kb *Eco*RI fragment from *purB*$^+$ phage 7F9 in *Eco*RI site of pUC118	This work
pBH112	1.3-kb *Kpn*I-*Eco*RI fragment from pBH111 in pUC118	This work
pBH113	570-bp *Bgl*I-*Eco*RI fragment from pBH111 in pUC118	This work
pBLG1	1.1-kb *Sma*I-*Bam*HI fragment (nucleotides −986 to 172 in *Sma*I-*Bam*HI sites of pMLB1034	This work
pBLG2	1.3-kb *Sma*I-*Bam*HI fragment (nucleotides −986 to 316) in pMLB1034	This work
pLT1	0.8-kb *Sma*I-*Bam*HI fragment (nucleotides −986 to −172) in *Sma*I-*Bam*HI sites of PRS415	This work
pLT2	1.1-kb *Sma*I-*Bam*HI fragment (nucleotides −986 to 172) in *Sma*I-*Bam*HI sites of PRS415	This work
pLT3	1.3-kb *Sma*I-*Bam*HI fragment (nucleotides −986 to 316) in PRS415	This work
pLT4	350-bp *Sma*I-*Bam*HI fragment (nucleotides −172 to 172) in pRS415	This work

N-terminal ASL amino acid sequence. An ASL–β-galactosidase chimeric enzyme was purified from *E. coli* R320(pBLG1). The enzyme was isolated from a cell extract by ammonium sulfate precipitation (0 to 45% saturation) followed by immunoaffinity chromatography on an anti-β-galactosidase agarose column (Promega). The eluted enzyme was dialyzed against 10 mM NH_4HCO_3 and then lyophilized. A 50-μg sample of protein was electrophoresed on an 8% sodium dodecyl sulfate-polyacrylamide gel, and proteins were transferred to a polyvinylidene difluoride-type membrane and lightly stained with Coomassie blue (12). The protein band was cut out and subjected to amino acid sequencing by the Purdue Laboratory for Macromolecular Structure.

Nucleotide sequence accession number. The sequence of gene *purB* has been deposited in the GenBank data base under accession no. M74924.

RESULTS

Cloning of *purB*. Gene *purB* was isolated from a library of *E. coli* DNA fragments by complementation of *E. coli purB* mutant TX570 as described in Materials and Methods. One representative plasmid, designated pBH101, containing an insert of 1.7 kb was found to contain the intact *purB* coding sequence but to lack the 5' flanking region (Fig. 2). On the basis of this information, the *purB* upstream region was isolated on a 3-kb *Eco*RI fragment from λ phage 7F9 from the library of Kohara et al. (8). This 3-kb *Eco*RI fragment was

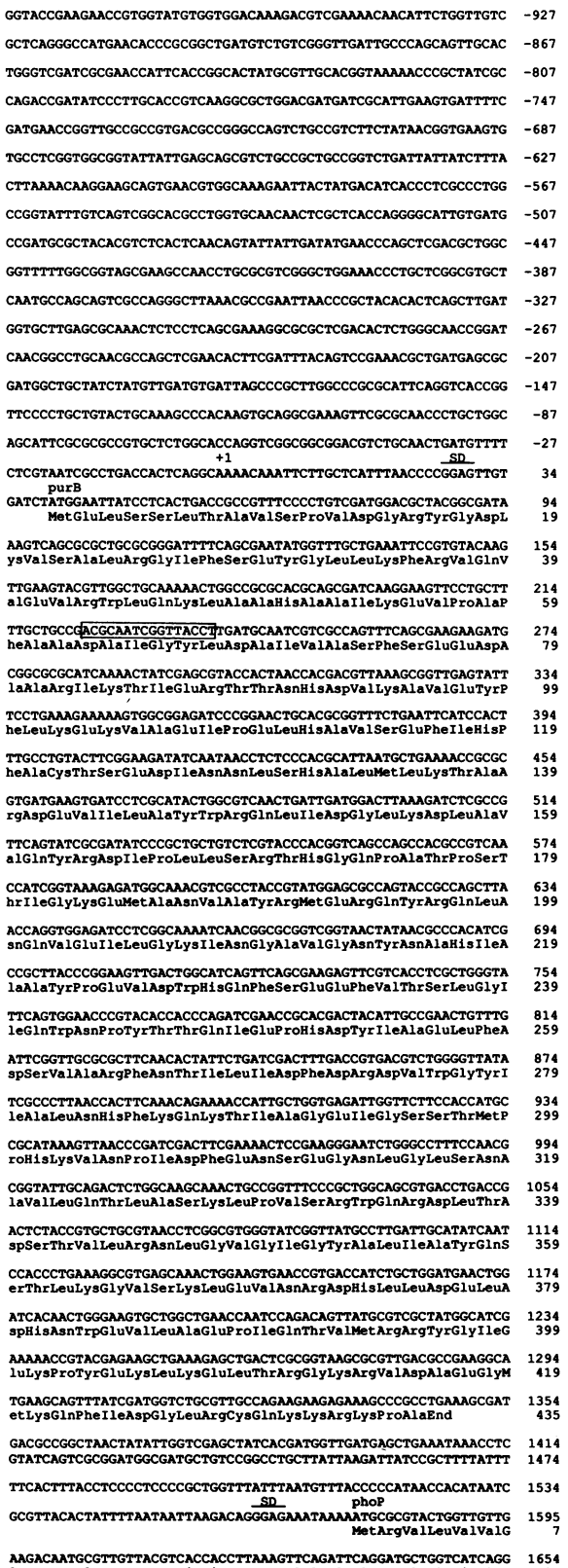

FIG. 3. Nucleotide sequence of *purB* and the derived ASL amino acid sequence. Numbering is from the experimentally determined start of transcription. The 16-bp conserved *pur* regulon operator is boxed. The sequence of *phoP* follows that of *purB*. Shine-Dalgarno (SD) sites are overlined.

initially subcloned into the *Eco*RI site of pUC118, giving pBH111. Next, a 1.3-kb *Kpn*I-*Eco*RI subfragment from pBH111 was transferred to pUC118 to yield pBH112. The 1.3 kb of *E. coli* DNA in pBH112 overlapped with pBH101 and extended the *purB* 5′ flanking region (Fig. 2).

Nucleotide sequence of *purB*. The nucleotide sequence of *purB* and the 5′ and 3′ flanking regions was determined on both DNA strands from plasmids pBH101 and pBH112 (Fig. 3). In *purB*$^+$ plasmid pBH101, there is a single open reading frame of 1,305 bp that encodes a protein of 435 amino acids with an M_r of 49,225. This open reading frame is ascribed to ASL on the basis of complementation of *purB* strain TX530 and homology with *Bacillus subtilis* (3) and avian (1) ASL. The NH_2-terminal amino acid sequence of ASL from a *purB-lacZ* translational fusion in plasmid pBLG1 was determined and found to correspond exactly to residues 1 to 6 of the amino acid sequence shown in Fig. 3.

A 16-bp sequence with an imperfect dyad symmetry at nucleotides 224 to 239 in the protein-coding region has identity with the *pur* regulon operator in 11 of 16 positions (Fig. 4). All highly conserved bases are found in the *purB* operatorlike sequence, except for the A at left-hand position 2 and the T at right-hand position 2. Evidence that this *pur* operatorlike sequence is involved in *purR* control is presented below.

Approximately 230 nucleotides downstream of *purB*, there is an open reading frame beginning with an ATG at nucleotide 1577. This open reading frame is preceded by a good Shine-Dalgarno sequence and extends for 50 codons to the end of the fragment. A search of the protein data base with FASTA (11) indicated 90% identity with *Salmonella typhimurium phoP* over the first 50 amino acids. We did not investigate *phoP* expression. However, there was close correspondence of nucleotides 1344 to 1349 and 1368 to 1373 to the −35 and −10 promoter elements, respectively.

Upstream of *purB*, at nucleotides −391 to −23, there is an open reading frame with the potential to encode a 123-amino-acid protein chain. This open reading frame is not preceded by a Shine-Dalgarno sequence, and translation is considered unlikely. No significant similarities to this sequence were found in release 28 of the National Biomedical Research Foundation protein data base.

***purB* promoter.** Three transcriptional fusion plasmids were constructed to detect *purB* promoter activity. Fragments of DNA corresponding to nucleotides −986 to −172, −172 to 172, and −986 to 172 were inserted into the *Sma*I-*Bam*HI sites of plasmid pRS415. This plasmid contains an intact *lacZ* coding sequence and ribosome-binding site but lacks a promoter. β-Galactosidase activity was measured from constructs transformed into strain MC4100. The data in Fig. 5 indicate that both subfragments of the upstream region, nucleotides −986 to −172 and −172 to 172, have the capacity for *purB* promoter function. Plasmid pLT2, containing *E. coli* DNA nucleotides −986 to 172, had 1.7- to 2.0-fold more activity than either subfragment.

Primer extension mapping was performed to identify *purB* mRNA 5′ ends. With a primer complementary to nucleotides 154 to 172, a single major cDNA was obtained ending at

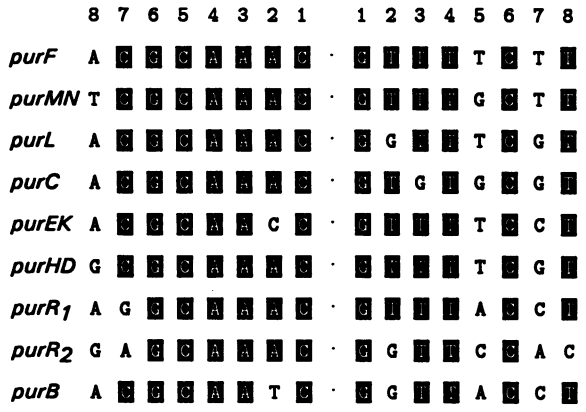

FIG. 4. Alignment of the *purB* operator with other *pur* regulon operators. Consensus positions that are conserved in six or more operators are highlighted.

adenine +1 (Fig. 6). The same mRNA 5′ end was mapped with a second primer complementary to nucleotides 120 to 138 (data not shown). In a separate experiment, the same mRNA 5′ end was mapped by using RNA from cells carrying *purB-lacZ* plasmid pBLG1 (data not shown). The transcription start site is thus designated as adenine +1. This position is not preceded by a typical σ70 promoter sequence. Additional upstream mRNA 5′ ends were not obtained by primer extension mapping with an oligonucleotide corresponding to nucleotides −164 to −147. Therefore, the promoter for *purB* transcription in the upstream region between nucleotides −986 to −172 in plasmid pLT1 remains to be identified.

Regulation by *purR*. A series of experiments was conducted to determine whether *purB* is subject to regulation by *purR* and whether the *pur* operatorlike sequence in the *purB* coding region has a regulatory function. Transcriptional and translational fusions to *lacZ* were constructed to investigate in vivo regulation by purines and *purR*. DNA fragments containing nucleotides −986 to 172 and −986 to 316 (Fig. 3) were ligated into vectors to give transcriptional *lacZ* fusion plasmids pLT2 and pLT3 and translational *lacZ* fusions in plasmids pBLG1 and pBLG2 (Fig. 7). Plasmids pLT2 and pBLG1 lack the operatorlike sequence, whereas pLT3 and pBLG2 contain the putative *purB* operator. These *purB-lacZ* fusions were integrated into the chromosome for assay of β-galactosidase. The data summarized in Table 2 demonstrate that with *purB-lacZ* transcriptional fusions in strains BH102, BH202, BH103, and BH203 and with translational fusions in strains BH100, BH200, BH101, and BH201, repression by adenine was dependent upon *purR*$^+$ and the *purB* operatorlike site. The *cis*-acting *purB* control site is

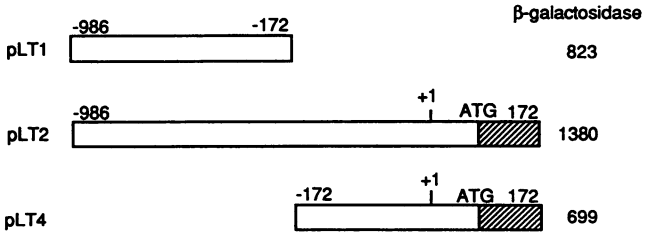

FIG. 5. Detection of *purB* promoter function. The boxes provide a schematic representation of *purB* 5′ flanking DNA cloned into *lacZ* transcriptional fusion vector pRS415. The shaded segment corresponds to the protein-coding region. Numbers refer to nucleotide positions. β-Galactosidase activity (nanomoles per minute per milligram of protein) is given for MC4100 transformants. The β-galactosidase control with the vector alone gave 35 U of β-galactosidase activity.

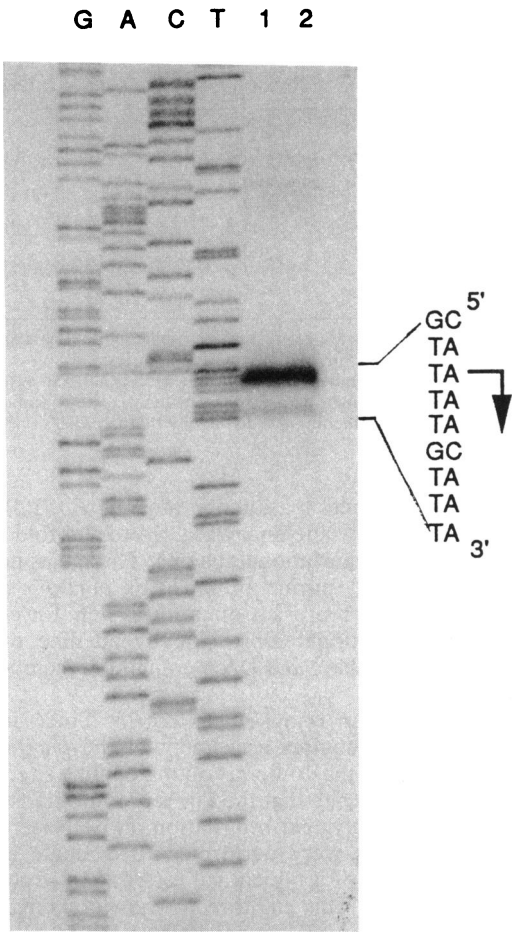

FIG. 6. Primer extension mapping of the 5′ end of *purB* mRNA. RNA was isolated from strain R320(pBH111). The same primer was used for synthesis of cDNA (lanes 1 and 2) and the sequencing ladder (lanes G, A, C, and T). The nucleotide sequence and mRNA 5′ end corresponding to the site for transcription initiation are shown on the right.

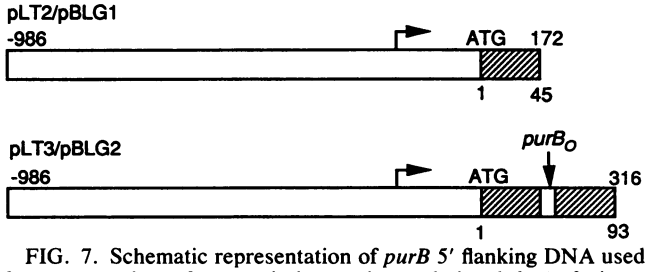

FIG. 7. Schematic representation of *purB* 5′ flanking DNA used for construction of transcription and translational *lacZ* fusions. Transcriptional fusions in pRS415 are pLT2 and pLT3. Translational fusions in pMLB1034 are pBLG1 and pBLG2. Numbers above the diagrams refer to the nucleotide sequence, and those below refer to the amino acid sequence. The horizontal arrows define the transcription start sites. The *purB* operator is designated $purB_o$.

TABLE 2. Regulation of *purB-lacZ*

Strain	*purR*	β-Galactosidase[a] sp act		Fold repression[b]
		+Adenine	−Adenine	
BH102 (*purBo*$^-$)	+	151	158	1.0
BH202 (*purBo*$^-$)	−	150	144	1.0
BH103 (*purBo*$^+$)	+	46.7	152	3.1–3.3
BH203 (*purBo*$^+$)	−	143	146	1.0
BH100 (*purBo*$^-$)	+	63.3	56.4	0.9–1.0
BH200 (*purBo*$^-$)	−	60.0	57.3	1.0
BH101 (*purBo*$^+$)	+	17.8	61.1	3.3–3.4
BH201 (*purBo*$^+$)	−	58.3	59.2	1.0

[a] β-Galactosidase values are expressed in specific activity as averages of three or four experiments with a variation of 10% or less.

[b] Fold repression was calculated two ways and expressed as a range: enzyme activity with adenine divided by that without adenine in *purR*$^+$ and enzyme activity with adenine in *purR* divided by that in *purR*$^+$.

thus confirmed to function as an operator (*purBo*$^+$). Repression of *purB-lacZ* by adenine was 3.1- to 3.4-fold in the transcriptional and translational fusions. There was no *purB* repression by added purine in *purR* or *purBo*$^-$ strains, indicative of a single repressor-operator system for control. Therefore, transcriptional repression by adenine requires not only *purR*$^+$ but also *purB* DNA containing the operator-like sequence.

Binding of the Pur repressor to *purBo*. Since in vivo regulation of *purB* requires a segment of the *purB* coding sequence containing the 16-bp *cis*-acting control site, it was important to demonstrate that the Pur repressor binds to this sequence. Accordingly, gel retardation and DNase I footprinting experiments were performed to evaluate specific binding. The results of a typical gel retardation experiment are shown in Fig. 8. This experiment shows binding of the purine repressor to a 570-bp *purB* DNA fragment corresponding to nucleotides 221 to 242. Approximately 80 ng of the repressor was required for 50% binding. Binding of the repressor was dependent upon the purine corepressor and was specific for *purB* DNA containing the 16-bp *pur* operator (data not shown).

To define the site of repressor-DNA interaction, DNase I footprinting was conducted. The results of a typical experiment are presented in Fig. 9, which shows that the Pur repressor bound specifically to the 16-bp *pur* operator in the 570-bp DNA fragment that was retarded in the experiment shown in Fig. 8. No other positions in the DNA fragment were protected against DNase I cleavage. This confirms that the operatorlike sequence in the *purB* coding region functions as a site for repressor binding and this repressor-DNA interaction is required for *purB* regulation.

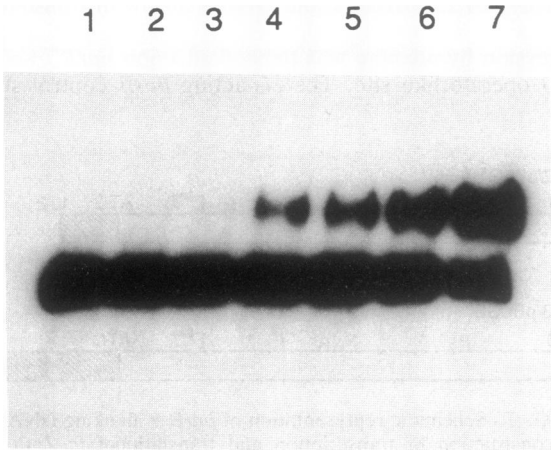

FIG. 8. Gel retardation assay for binding of the purine repressor to *purBo*. All lanes contained 10 fmol of the 570-bp *Hin*dIII-*Eco*RI DNA fragment from pBH113. Lanes 1 to 7 contained 0, 5, 10, 20, 40, 80, and 160 ng of the purine repressor. Free DNA and DNA-protein complexes were resolved on a 4% polyacrylamide gel.

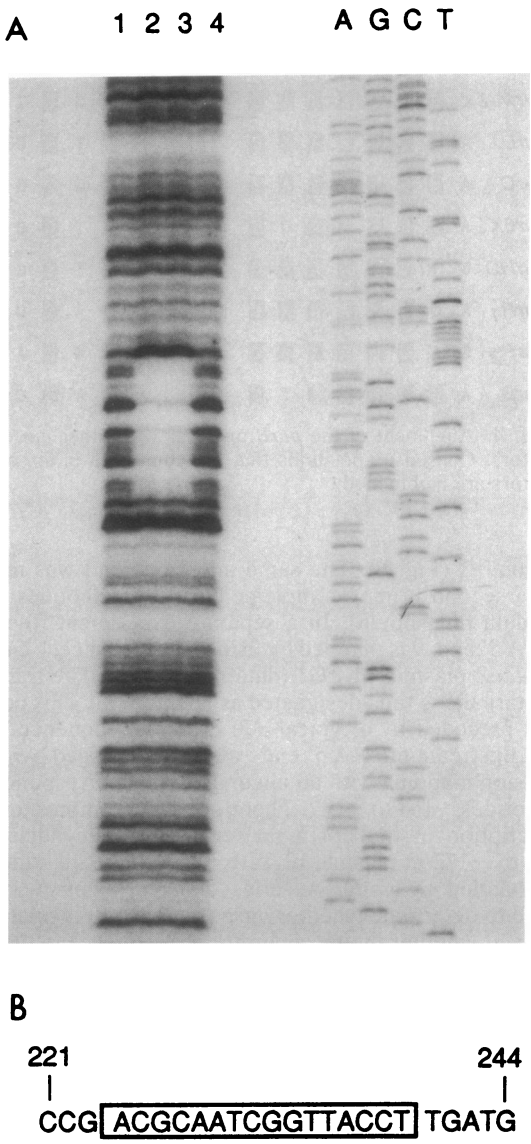

FIG. 9. DNase I footprint of the *purB*-repressor complex. (A) Lanes: 1 and 4, 570-bp *Hin*dIII-*Eco*RI DNA from pBH113 labeled at the *Eco*RI site; 2 and 3, the same DNA plus 0.4 μg of the purine repressor; A, G, C, and T, sequencing ladder from pBH113. (B) Nucleotide sequence of the region protected by the repressor. The 16-bp *purB* operator is boxed.

DISCUSSION

E. coli purB was cloned by functional complementation of a *purB* mutant. The DNA sequence predicts an ASL protein of 435 amino acids having a molecular weight of 49,225. *E.*

TABLE 3. Summary of *E. coli* ASL alignments[a]

Enzyme	EcASL	BsASL	GgASL	EcFUM	BsFUM	EcASP	HuRSL
EcASL	100						
BsASL	26 (29)	100					
GgASL	22 (15)	26	100				
EcFUM	19 (6.0)	22	24	100			
BsFUM	21 (12)	22	21	63	100		
EcASP	20 (6.5)	21	20	27	32	100	
HuRSL	19 (3.4)	21	22	22	23	20	100

[a] The numbers are percentages of identical amino acids for pairwise comparisons of *E. coli* ASL (EcASL), *B. subtilis* ASL (BsASL) (3), chicken ASL (GgASL) (1), *E. coli* fumarase (EcFUM) (29), *B. subtilis* fumarase (BsFUM) (16), human argininosuccinate lyase (HuRSL) (14), and *E. coli* aspartase (EcASP) (29). The numbers in parentheses are *Z* values (standard deviations above the alignment of randomized sequences).

coli ASL was aligned with ASL sequences from *B. subtilis* and chickens, as well as sequences of related enzymes that catalyze β-elimination reactions. These results are summarized in Table 3. There was 26% identity with the enzyme from *B. subtilis* and 22% identity with avian ASL. Both comparisons had *Z* values of >10 standard deviations above random, indicating a highly significant probability of relatedness. Interestingly, the data in Table 3 indicate significant similarity between ASL and *B. subtilis* fumarase. The *Z* value of 12 suggests an evolutionary relationship for these enzymes. *E. coli* ASL, *E. coli* fumarase, and human argininosuccinate lyase have several blocks of identical amino acids, although the *Z* values are insufficient to support homology. The most highly conserved region in these enzymes is between residues 294 and 308, on the basis of *E. coli* ASL numbering (Fig. 10). This block of residues has been proposed to represent a "signature sequence" for enzymes that catalyze β-elimination reactions that generate fumarate as a product (1). The previously suggested mechanistic similarity among fumarase, ASL, and argininosuccinate lyase (5) thus reflects a similarity in primary structure and a likely evolutionary relationship.

The data summarized in Table 2 indicate that *purB* expression is regulated by the purine pool and *purR*. Guanine and hypoxanthine are corepressors, needed for Pur repressor binding to the DNA control site (21). With single-copy *purB-lacZ* chromosomal integrants, there was threefold regulation by *purR*. This transcriptional regulation requires the 16-bp *purB* operator sequence located 224 nucleotides downstream from the transcription start site in the protein-coding region. Whereas this location for the *cis*-acting control site relative to the promoter is different from that in the other six operons involved in IMP synthesis, it resembles that in *purR*. For the other structural genes required for IMP synthesis, the *pur* operator is located between positions −46 and +10 relative to the start of transcription (6). However, in *purR* there are dual operators at positions 96 to 111 and 184 to 199 relative to the start of transcription (20). The downstream operator, *purRo2*, overlaps codons 10 to 15 in *purR*. In *purR*, operators O_1 and O_2 contribute 2.5- to 2.7-fold transcriptional repression, whereas *purRo2* by itself contributes only about 1.5-fold regulation by the repressor. Thus repressor control at *purBo* is about twofold more effective than the regulation at *purRo2*, even though *purBo* is 41 nucleotides farther downstream from the site for transcription initiation and is situated deeper in the protein-coding region that overlaps codons 62 to 67. As shown in Fig. 4, there are only minor nucleotide sequence differences between *purBo* and *purRo2*. There is one further strong similarity between *purB* and *purR*. Neither gene has a good consensus −35 and −10 σ70 promoter sequence next to the site mapped for transcription initiation.

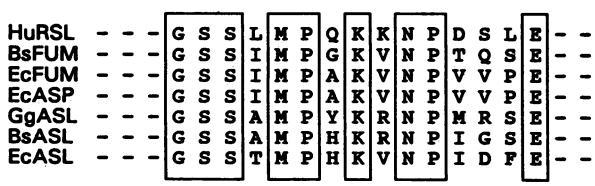

FIG. 10. Sequence alignment of a conserved region implicated as a signature sequence for β-elimination reactions that generate fumarate. Sequences are ASL from *E. coli* (Ec), *B. subtilis* (Bs) (3), and chickens (Gg) (1); fumarase (FUM) from *E. coli* (29) and *B. subtilis* (16); and argininosuccinate lyase (RSL) from humans (Hu) (14).

The regulation of the other *pur* genes is over a 5- to 17-fold range (6) and is likely determined by the competing rates of repressor-operator and RNA polymerase-promoter complex formation. Large differences in *lac* operon repression have been found to result from placement of the operator in different positions relative to the promoter, and it was concluded that the mechanism of repression of transcription initiation is dependent upon relative operator-promoter positioning (9). The mechanism by which a repressor that binds to a DNA site greater than 200 bp downstream from the promoter can regulate transcription is not known. It would appear to be necessary for the bound repressor either to interfere with transcription initiation or to act as a road block and obstruct the progress of the transcribing polymerase.

The threefold regulation of *purB* by *purR* confirms the earlier observation by Meng et al. for repression of ASL by *purR* (15). Previously, we (6) found no regulation of ϕ(*purB-lacZ*)*205* made in vivo by Mud1(Ap *lac*) insertion (28), although two- to threefold regulation had been noted in the original report (28). The position of the *purB* fusion to *lacZ* in ϕ(*purB-lacZ*)*205* is not known, and therefore we do not know whether *purBo* is included in the fusion.

ACKNOWLEDGMENTS

We thank our colleagues Ronald Somerville and Barry Wanner for providing phage 7F9 and plasmid pRS415, respectively, Kang Yell Choi for providing the pure PurR repressor, and Gaochao Zhou for important suggestions and discussion.

This work was supported by Public Health Service grants GM24658 (H.Z.) and AI20068 (J.M.S.) from the National Institutes of Health. Computer facilities were supported by NIH grant AI27713.

REFERENCES

1. **Aimi, J., J. Badylak, J. Williams, Z. D. Chen, H. Zalkin, and J. E. Dixon.** 1990. Cloning of a cDNA encoding adenylosuccinate lyase by functional complementation in *Escherichia coli*. J. Biol. Chem. **265**:9011–9014.
2. **Bachmann, B.** 1983. Linkage map of *Escherichia coli* K-12, edition 7. Microbiol. Rev. **47**:180–230.
3. **Ebbole, D. J., and H. Zalkin.** 1987. Cloning and characterization of a 12-gene cluster from *Bacillus subtilis* encoding nine enzymes for *de novo* purine nucleotide synthesis. J. Biol. Chem. **262**:8274–8287.
4. **Elledge, S. J., and G. C. Walker.** 1985. Phasmid vectors for identification of genes by complementation of *Escherichia coli* mutants. J. Bacteriol. **162**:777–783.
5. **Havir, E., and K. Hanson.** 1973. The enzymic elimination of ammonia, p. 75–166. *In* P. D. Boyer (ed.), The enzymes, 3rd ed., vol. 7. Academic Press, Inc., New York.
6. **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.
7. **Henikoff, S.** 1987. Unidirectional digestion with exonuclease III in DNA sequence analysis. Methods Enzymol. **155:**156–165.
8. **Kohara, Y., K. Akiyama, and K. Isono.** 1987. The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell **50:**495–508.
9. **Lanzer, M., and H. Bujard.** 1988. Promoters largely determine the efficiency of repressor action. Proc. Natl. Acad. Sci. USA **85:**8973–8977.
10. **Layne, E.** 1957. Spectrophotometric and turbidometric methods for measuring proteins. Methods Enzymol. **3:**448–451.
11. **Lipman, D. J., and W. R. Pearson.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85:**2444–2448.
12. **Madsudaira, P.** 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membrane. J. Biol. Chem. **262:**10035–10038.
13. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 383–389. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
14. **Matuo, S., M. Tatsuno, K. Kobayashi, T. Saheki, T. Miyata, S. Iwanaga, Y. Amaya, and M. Mori.** 1988. Isolation of cDNA clones of human argininosuccinate lyase and corrected amino acid sequence. FEBS Lett. **234:**395–399.
15. **Meng, L. M., M. Kilstrup, and P. Nygaard.** 1990. Autoregulation of *purR* repressor synthesis and involvement of *purR* in the regulation of *purB*, *purC*, *purL*, *purMN* and *guaBA* expression in *Escherichia coli*. Eur. J. Biochem. **187:**373–379.
16. **Miles, J. S., and J. R. Guest.** 1985. Complete nucleotide sequence of the fumarase gene (*citG*) of *Bacillus subtilis* 168. Nucleic Acids Res. **13:**131–140.
17. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 403–404. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. **Mullis, K. B., and F. A. Faloona.** 1987. Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. Methods Enzymol. **155:**335–350.
19. **Rolfes, R. J., and H. Zalkin.** 1988. Regulation of *Escherichia coli purF*. Mutations that define the promoter, operator, and purine repressor gene. J. Biol. Chem. **263:**19649–19652.
20. **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.
21. **Rolfes, R. J., and H. Zalkin.** 1990. Purification of the *Escherichia coli* purine regulon repressor and identification of corepressors. J. Bacteriol. **172:**5637–5642.
22. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
23. **Shimotsu, H., M. I. Kuroda, C. Yanofsky, and D. J. Henner.** 1986. Novel form of transcriptional attenuation regulates expression of the *Bacillus subtilis* tryptophan operon. J. Bacteriol. **166:**461–471.
24. **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.
25. **Simons, R. W., F. Houman, and N. Klechner.** 1987. Improved single and multicopy Lac-based cloning vector for protein and operon fusions. Gene **53:**85–96.
26. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19:**259–268.
27. **Vogel, H. T., and D. M. Bonner.** 1956. Acetylornithinase of *Escherichia coli*: partial purification and some properties. J. Biol. Chem. **218:**97–106.
28. **Wolfe, S. A., and J. M. Smith.** 1985. Separate regulation of *purA* and *purB* loci of *Escherichia coli* K-12. J. Bacteriol. **162:**822–825.
29. **Woods, S. A., J. S. Miles, R. E. Roberts, and J. R. Guest.** 1986. Structural and functional relationships between fumarase and aspartase. Biochem. J. **237:**547–557.
30. **Zheng, L. M., P. C. Andrews, M. A. Hermodson, J. E. Dixon, and H. Zalkin.** 1990. Cloning and structural characterization of porcine heart aconitase. J. Biol. Chem. **265:**2814–2821.