JOURNAL OF BACTERIOLOGY, Aug. 1990, p. 4555–4562
0021-9193/90/084555-08$02.00/0

Vol. 172, No. 8

Genes of the *Escherichia coli pur* Regulon Are Negatively Controlled by a Repressor-Operator Interaction†

BIN HE,[1] ALLAN SHIAU,[2] KANG YELL CHOI,[1] HOWARD ZALKIN,[1]* AND JOHN M. SMITH[2]

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

Received 1 March 1990/Accepted 22 May 1990

Fusions of *lacZ* were constructed to genes in each of the loci involved in de novo synthesis of IMP. The expression of each *pur-lacZ* fusion was determined in isogenic *purR* and *purR*$^+$ strains. These measurements indicated 5- to 17-fold coregulation of genes *purF*, *purHD*, *purC*, *purMN*, *purL*, and *purEK* and thus confirm the existence of a *pur* regulon. Gene *purB*, which encodes an enzyme involved in synthesis of IMP and in the AMP branch of the pathway, was not regulated by *purR*. Each locus of the *pur* regulon contains a 16-base-pair conserved operator sequence that overlaps with the promoter. The *purR* product, purine repressor, was shown to bind specifically to each operator. Thus, binding of repressor to each operator of *pur* regulon genes negatively coregulates expression.

In all organisms there are 10 steps for de novo synthesis of IMP, the first purine nucleotide intermediate in the pathway. IMP is a branch point metabolite which is converted to adenine and quanine nucleotides (Fig. 1). Although this pathway is invariant, the genetic organization and regulation of expression differ between organisms. In *Escherichia coli* and *Salmonella typhimurium*, the genes encoding these enzymes are scattered around the chromosome as individual loci and small operons (8), whereas in *Bacillus subtilis* the genes for synthesis of IMP are organized as a single large operon (6). In these bacteria, the addition of exogenous purines to defined growth medium causes repression of all genes in the pathway. However, in *E. coli* and *S. typhimurium*, the AMP and GMP branches appear to be under separate regulation from the main pathway leading to IMP (8). The study of the regulation of purine nucleotide biosynthesis has been hindered by the availability of substrates and appropriately sensitive enzyme assays. This has necessitated, in some cases, the utilization of coupled assays instead of direct measurements of individual steps (8).

The first mutations, designated *purR*, shown to affect overall regulation of the de novo pathway arose fortuitously in an *S. typhimurium purA* strain (8). Other *purR* mutants were isolated by using resistance to the inhibitory purine analog 6-mercaptopurine (8, 15) and by exploiting the adenine sensitivity for growth of *pur-lac* fusions with lactose (8, 16, 17, 29). However, in all of these instances, the term *purR* was used only to designate a regulatory phenotype, since the individual mutations were not characterized genetically. The genetic characterization of the PurR phenotype was limited because of the high spontaneous mutation rate to the *purR* phenotype in the strain backgrounds used for study and by the availability of enzyme assays noted above. As a result of these earlier investigations, the de novo purine nucleotide biosynthetic pathway leading to the synthesis of IMP was inferred to be under the control of a common regulatory element. Although this view is widely accepted, the precise mechanism for the regulation of each of the individual loci by the *purR* gene was not experimentally established. Furthermore, the effector molecules that act as coregulators have not been identified. Nucleotides have been assumed to be the effector molecules, but the purine bases hypoxanthine and guanine have been implicated as acting directly without conversion to the nucleotide form (10).

Recently, the *purR* regulatory element from *E. coli* has been cloned, sequenced, and mapped (12, 25) to coordinate kilobase pair (kb) 17755 on the *E. coli* restriction map (13), corresponding to min 36 on the chromosome. The *pur* repressor is a protein of 341 amino acids having homology to *lacI*, *galR*, and *cytR* (25). The PurR binding site in gene *purF*, a 16-base-pair (bp) imperfect palindrome, was identified by mutational analysis and by direct DNA footprinting (24, 25). This PurR binding site in the *purF* operon has been recognized in the control regions of several of the other *pur* genes (1, 7a, 26, 28, 30, 36; A. A. Tiedemann, D. J. DeMarini, J. Parker, and J. M. Smith, submitted for publication) as well as in gene *purR* (25). Another *purR* mutation has been shown by Kilstrup et al. (12) to regulate the expression of the *purD* gene in the *purHD* operon, the *purF* operon, and the gene for cytosine deaminase. In this report, we extend these studies and demonstrate directly that the product of the *purR* gene (purine repressor) binds to a conserved operator sequence and regulates the expression of the other *pur* loci, leading to the synthesis of IMP.

After submission of this work, a paper by Meng et al. (19) reported evidence that all genes involved in the synthesis of IMP and GMP, except for *purA*, are regulated by *purR*.

MATERIALS AND METHODS

Bacterial strains and media. The strains used in this study were all derivatives of *E. coli* K-12 and are described in Table 1. Strain TX337 was constructed from strain W3110 (2) by the sequential P1-mediated introduction of a *pro-lac* deletion by *o*-nitrophenyl-β-D-thiogalactoside selection (20) and then the introduction of the Δ(*lac*)U169 mutation from strain TX302 by selection for Pro$^+$. The rich medium was LB (20), and the minimal medium contained salts (35), 0.5% glucose, 2 μg of thiamine per ml, 0.2% acid-hydrolyzed casein, and supplements as required. Adenine was added at 100 μg/ml. MacConkey agar was used for isolation of *purR* strains.

Construction of *pur-lacZY* fusion strains. Construction of

* Corresponding author.

† Journal paper no. 12480 from the Purdue University Agricultural Experiment Station.

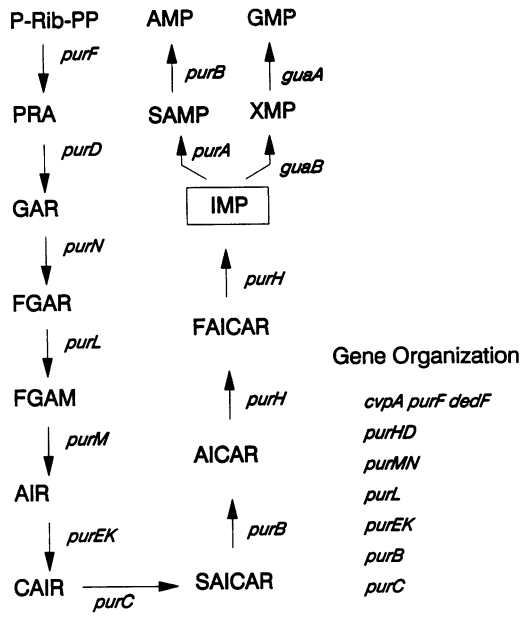

FIG. 1. Pathway for de novo purine nucleotide synthesis, *E. coli* gene designations, and gene organization. Abbreviations: P-Rib-PP, 5-phosphoribosyl-1-pyrophosphate; PRA, phosphoribosylamine; GAR, phosphoribosylglycinamide; FGAR, phosphoribosylformylglycinamide; FGAM, phosphoribosylformylglycinamidine; AIR, phosphoribosylaminoimidazole; CAIR, phosphoribosylcarboxyaminoimidazole; SAICAR, phosphoribosylsuccinocarboxamideaminoimidazole; AICAR, phosphoribosylaminoimidazole carboxamide; FAICAIR, phosophoribosylformimidoimidazole carboxamide.

the *purD-lacZ* (7a), *purE-lacZ* (30), *purC-lacZ* (Tiedeman et al., submitted) fusion strain has been described. The remaining *pur-lac* fusions were constructed by inserting a *lacZY*::Kan^r cassette (31) of the appropriate reading frame into or between unique restriction sites within each structural gene. After verification by restriction digest, each *pur-lac* fusion was recombined onto the *E. coli* chromosome by the procedure of Winans et al. (40) and then transferred by P1 transduction to a common genetic background, strain TX302. A *purM-lacZ* fusion in strain TX709 was constructed by inserting a *Xma*I-digested pLKC481 *lacZY*::Kan^r cassette into the unique *purM Bsp*MII sites of plasmid pJS18 (28). Two separate *purL-lacZY*::Kan^r fusions differing only in the insertion site of the *lacZY*::Kan^r cassette were constructed. The *purL-lacZY*::Kan^r fusion of strain TX701 was constructed by ligating a *Sma*I-digested *lacZY*::Kan^r cassette into the unique *purL Hpa*I restriction site of plasmid pJS157 (26). This fusion was transduced into strain TX302 to yield TX701 and was also transferred into strain TX337, because of its suppressor-free background, to form strain TX705. Strain TX705 was used for the isolation of *purR* mutations by transposon mutagenesis. A second *purL-lacZY*::Kan^r fusion in strain TX768 was constructed by ligating a *Hin*dIII (blunted)-*Sma*I-digested pLKC481 *lacZY*::Kan^r cassette into the unique *purL Pst*I and *Eco*RI sites of plasmid pJS336. Plasmid pJS336 was constructed by subcloning the *Sma*I-*Xba*I restriction fragment from pJS80 (26) into the *Kpn*I (blunted)-*Xba*I sites of Bluescript SK$^-$ (Stratagene, Inc.). The *purF-lacZ* fusion in strain TX771, similar to the one described by Rolfes and Zalkin (25), was constructed by ligating a *Sma*I-digested pLKC480 *lacZY*::Kan^r cassette into the unique *purF Hpa*I and *Nru*I sites of plasmid pJS114. Plasmid pJS114 was constructed by subcloning the 4.3-kb *Pst*I fragment from plasmid pSB5 (32) into the *Pst*I site of plasmid pSB118 (33).

Transposon mutagenesis. Transposon mutagenesis was with a mini-Tet derivative of Tn*10*, using bacteriophage λ 1098 (37). The resulting *purR* mutant is designated *purR*::Tn*10*.

Isolation and subcloning of DNA fragments containing *pur* regulon 5′ control sites. Fragments containing the 5′ control regions of *pur* genes were isolated by electroelution from 5% polyacrylamide gels and were ligated into phagemid vector pUC118 (34) as follows: *purM*, 383-bp *Eco*RI-*Ava*II DNA from pJS18 (28) made blunt and ligated into the *Hin*cII site of

TABLE 1. Strains used

Strain	Genotype	Reference
TX302	Δ(*lac*)U169 *sup*	39
TX337	Δ(*lac*)U169	This study
TX529	Δ(*gpt-pro-lac*) *ilvB2102 ilvHI2202 rbs221 ara thi srlC300*::Tn*10 recA56*	27
TX530	Δ(*lac*) *ara* Φ(*purB′-lacZ*$^+$ *Y*$^+$::λp1(209))205	41
TX701	Δ(*lac*)U169 *sup* Φ(*purL′-lacZ*$^+$ *Y*$^+$::Kan^r)217 Hyb	This study
TX705	Δ(*lac*)U169 Φ(*purL′-lacZ*$^+$ *Y*$^+$::Kan^r)217 Hyb	This study
TX709	Δ(*lac*)U169 *sup* Φ(*purM′-lacZ*$^+$ *Y*$^+$::Kan^r)218 Hyb	This study
TX717	Δ(*lac*)U169 *sup* Φ(*purC′-lacZ*$^+$ *Y*$^+$::Kan^r)219 Hyb	Tiedeman et al., submitted
TX725	Δ(*lac*)U169 *sup* Φ(*purE′-lacZ*$^+$ *Y*$^+$::Kan^r)214 Hyb	30
TX726	Δ(*lac*)U169 *sup* Φ(*purD′-lacZ*$^+$ *Y*$^+$::Kan^r)216 Hyb	7a
TX729	Δ(*lac*)U169 *sup* Φ(*purE′-lacZ*$^+$ *Y*$^+$::Kan^r)214 Hyb *purR220*::Tn*10*	31
TX764	Δ(*lac*)U169 *sup* Φ(*purD′-lacZ*$^+$ *Y*$^+$::Kan^r)216 Hyb *purR220*::Tn*10*	7a
TX768	Δ(*lac*)U169 *sup* Φ(*purL′-lacZ*$^+$ *Y*$^+$::Kan^r)221 Hyb	This study
TX769	Δ(*lac*)U169 *sup* Φ(*purL′-lacZ*$^+$ *Y*$^+$::Kan^r)221 Hyb *purR220*::Tn*10*	This study
TX771	Δ(*lac*)U169 *sup* Φ(*purF′-lacZ*$^+$ *Y*$^+$::Kan^r)222 Hyb	This study
TX773	Δ(*lac*)U169 *sup* Φ(*purF′-lacZ*$^+$ *Y*$^+$::Kan^r)222 Hyb *purR220*::Tn*10*	This study
TX778	Δ(*lac*)U169 *sup* Φ(*purM′-lacZ*$^+$ *Y*$^+$::Kan^r)218	This study
TX779	Δ(*lac*)U169 *sup* Φ(*purC′-lacZ*$^+$ *Y*$^+$::Kan^r)219 Hyb *purR220*::Tn*10*	Tiedeman et al., submitted
TX780	Δ(*lac*) *ara* Φ([*purB′*−*lacZ*$^+$ *Y*$^+$::gp1209])205 *purR220*::Tn*10*	Tiedeman et al., submitted
JC7623	*recB21 recC22 sbcB15 thr-1 leuB6 hisG4 argE3* λ(*gpt-proA*)*62 thi-1 ara-14 lacY1 tsx-33 supE44 galK2 rpsL31 kdgK51 xyl-5 mtl-1 rfbD1*	40
R303(pRRM127)	MC4100 (λ*purF-lacZ*)*recA* (Mu1$^+$) (*purR*$^+$ Km^r)	25
R320	MC4100 *purR300*	24

the vector to yield pMNo; *purL*, 105-bp *Mlu*I (blunt)-*Hin*dIII from pJS188 (2.5-kb *Eco*RI fragment from pJS80 [26] in pUC18) ligated into the *Hin*cII-*Hin*dIII sites of the vector to yield pLo; *purEK*, 240-bp *Rsa*I-*Ava*II (blunt) DNA from pJS131 (30) ligated into the *Hin*cII site to yield pEKo; *purHD*, 440-bp *Bam*HI-*Eco*RI DNA from pJS189 (Flannigan et al., in press) ligated into the *Bam*HI-*Eco*RI sites to yield pHDo; *purC*, 280-bp *Dde*I (blunt) DNA from pJS229 ligated into the *Hin*cII site to yield pCo; *purF*, 220-bp *Stu*I-*Nde*I (blunt) DNA from pRR10 ligated into the *Hin*cII site to yield pFo. DNA fragments were made blunt, when indicated, by filling ends with DNA polymerase Klenow fragment and deoxynucleoside triphosphates.

Preparation of extract containing purine repressor. A 20-ml culture of strain R303 (pRRM127) was grown to late log phase in minimal medium supplemented with adenine (100 μg/ml) and kanamycin (50 μg/ml). Extracts were prepared as described previously (25) and stored in small samples at −70°C.

Repressor-operator binding. DNA fragments containing a *pur* gene control region were labeled on one end with T4 polynucleotide kinase and [γ-^{32}P]ATP and were isolated from a 5% polyacrylamide gel by electroelution. Gel retardation assays were conducted as described previously (25) with 10 fmol of DNA fragment and variable amounts of extract from *purR*$^+$ plasmid-bearing strain R303 (pRRM127) in a volume of 20 μl. Binding specificity was determined by using extract from *purR* strain R320. After electrophoretic separation, bands corresponding to free DNA and protein-DNA complex were excised from the gel and counted for radioactivity. The method for DNase I footprinting has been described (25).

Enzyme assays. All strains were grown in minimal medium supplemented with adenine (100 μg/ml). Cells grown overnight were inoculated into fresh medium and grown to late log phase (Klett 100) at 37°C. Cells were harvested by centrifugation, suspended in 0.1 M sodium phosphate (pH 7.0), and disrupted by two passages through a French pressure cell at 20,000 lb/in^2. The lysate was centrifuged at 12,000 × *g* for 15 min, glycerol was added to 20% (vol/vol), and extracts were stored at −70°C before assay of β-galactosidase activity (20). Protein was determined by the method of Lowry as described by Layne (14).

RESULTS

Isolation and characterization of *purR*::Tn*10* mutations. Because of a high spontaneous mutation rate to PurR$^-$, it is difficult to use spontaneous *purR* mutations for genetic characterizations and manipulations. To overcome this obstacle in the construction of isogenic strains, it was desirable to isolate a transposon-induced *purR* mutation. Accordingly, the *purL-lacZY*::Kan^r strain TX705 was mutagenized with the mini-Tet element derived from the Tn*10* transposon (37), and tetracycline-resistant colonies were selected on MacConkey agar plates supplemented with hypoxanthine (50 μg/ml). On this medium, the wild type *purL-lacZY*::Kan^r fusion strain forms white colonies. Therefore, red colonies, which should represent derepression of the *purL-lacZY*::Kan^r fusion, were selected as putative *purR*::Tn*10* mutants. After initial characterization by P1-mediated backcrosses into strain TX705, three independent mini-Tet-induced regulatory mutants were retained for genetic characterization. All three were initially identified as *purR* mutations by their P1 linkage to the *man* and *pdxH* loci and confirmed by complementation with *purR*$^+$ plasmid pPR1002 (25). One mutation, *purR220*::Tn*10*, was used to test the other *pur* loci for regulation by the *purR* regulatory protein.

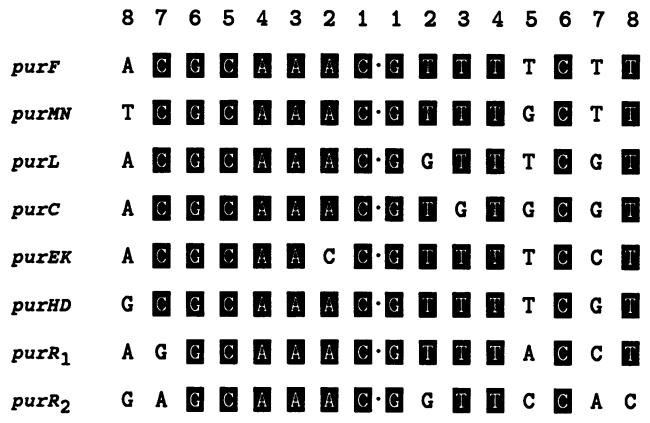

FIG. 2. Alignment of predicted *pur* regulon control sites. Each dot represents the center of an imperfect dyad symmetry. Sequences for *purF* (18, 24), *purMN* (28), *purL* (26), *purEK* (30, 36), *purHD* (1), *purC* (Tiedeman et al., submitted), and *purR* (25) control sites have been reported. Consensus positions that are conserved in six or more of the operators are highlighted.

Common regulation of the *pur* genes by *purR*. The DNA sequences of the genes required for the synthesis of IMP have been determined (1, 7a, 18, 26, 28, 30, 36; Tiedeman et al., submitted) except for *purB*, and a conserved segment with dyad symmetry is found in each control region (Fig. 2). This conserved region, which is also present in *purR* (25), has a consensus sequence 5′-NCGCAAACGTTTNCNT. This sequence in the *purF* control region (18) has been shown by mutational analysis (24) and DNase I footprinting (25) to be the binding site for the *purR* regulatory protein. Thus, the conserved sequences in the control regions of the other *pur* loci were also inferred to be binding sites for the *purR* regulatory protein. To determine the precise role of the *purR* regulatory protein in the control of expression of these different *pur* loci, we undertook to measure the effect of a *purR*::Tn*10* mutation on the expression of these genes as well as investigate protein-DNA binding.

Coregulation of *pur* genes by *purR* was quantitated by measurement of β-galactosidase activity from *pur-lacZ* fusions. Each locus in the common pathway to IMP, with the exception of *purB*, was regulated by *purR* (Table 2). The effect of *purR* was to repress gene expression between 4.6- and 17-fold, with *purF*, which encodes the first enzyme in the pathway, exhibiting the greatest regulation. The 17-fold regulation of *purF* obtained with this *purR*::Tn*10* mutation is similar to the 18- to 21-fold regulation of a different *purF-lacZ* construct with two different *purR* alleles (24). Gene *purB* encodes adenylsuccinate lyase, an enzyme that catalyzes reaction 8 in the pathway to IMP and also a reaction in the branch to AMP (Fig. 1). In agreement with suggestions from earlier studies (29, 41), *purB* expression was not subject to regulation by *purR*.

Interaction of purine repressor with a conserved control site. We conducted gel retardation and DNase I protection assays to evaluate binding of purine repressor to the inferred control sites. End-labeled DNA fragments containing each of the putative control sites were used for DNA binding studies with crude repressor protein. Typical results for titration of *pur* regulon control sites with purine repressor by gel retardation are shown in Fig. 3. In every case there were

TABLE 2. Regulation of *pur-lacZ* by PurR

Strain	Gene fusion	*purR*	β-Galactosidase activity[a]	Fold regulation
TX771	*purF*	+	3.29	17
TX773	*purF*	−	56.7	
TX726	*purHD*	+	39.6	12
TX764	*purHD*	−	479	
TX768	*purL*	+	53.1	6.7
TX769	*purL*	−	357	
TX709	*purM*	+	39.7	4.6
TX778	*purM*	−	183	
TX725	*purE*	+	16.2	15
TX729	*purE*	−	246	
TX717	*purC*	+	24.8	11
TX779	*purC*	−	280	
TX530	*purB*	+	52.5	1.2
TX780	*purB*	−	64.3	

[a] β-Galactosidase specific activity is given as nanomoles of *o*-nitrophenol formed per minute per milligram of protein at 28°C. Values are the average of two to three determinations that differed by less than 15%.

single, clearly separated bands corresponding to free DNA and DNA-protein complex. Binding curves obtained from these data are shown in Fig. 4. Under the conditions used, binding of repressor to the six control sites was similar, with approximately 2.0 to 3.0 μg of extract protein required for 50% binding. Evaluation of more precise binding constants awaits measurements with pure repressor protein. Control experiments using extracts from *purR* mutant R320 established the specificity for repressor-operator binding. In no case was a protein-DNA complex obtained with use of 10 μg of extract protein from the *purR* mutant (not shown).

Gel retardation assays provide evidence that purine repressor binds to 5′-flanking sequences of *pur* genes that contain a control site. DNase I footprinting experiments were conducted to define the site of protein-DNA interaction. Figure 5 shows representative DNase I footprints. For each DNA fragment, there was a single region protected from digestion by DNase I. The DNase I footprinting results are summarized in Fig. 6. Control regions in *purF*, *purL*, *purMN*, *purHD*, *purEK*, and *purC* are numbered from +1, the start of transcription. For each operon, purine repressor bound to the control region and protected approximately 20 to 24 bp against digestion by DNase I. Although the exact boundaries were difficult to determine because not every base is subject to digestion, in each case bound repressor protected the entire operator and protection usually extended approximately two to five bases beyond the operator boundaries in the 5′ and 3′ directions. The two exceptions were *purL* and *purMN*. There was no protection beyond the 3′ boundary of the *purMN* operator, and only one base on the 5′ end of the operator in *purL* was protected. In the *purF*, *purL*, *purMN*, and *purEK* control regions, the operator is seen to overlap the sequence corresponding to the −35 region of the promoter. However, in the *purHD* and *purC* operons, the operator abuts the −10 promoter region.

DISCUSSION

For 7 of the 10 steps in the pathway to IMP, each enzyme is encoded by a single gene. There are three steps that are more complex. Phosphoribosylaminoimindazole carboxylase, the enzyme catalyzing step 6, is a heterodimer encoded by genes *purEK*. Phosphoribosylaminoimidazole carboximide transformylase (*purH*) is a bifunctional enzyme that catalyzes steps 9 and 10. Adenylosuccinate lyase (*purB*) is a

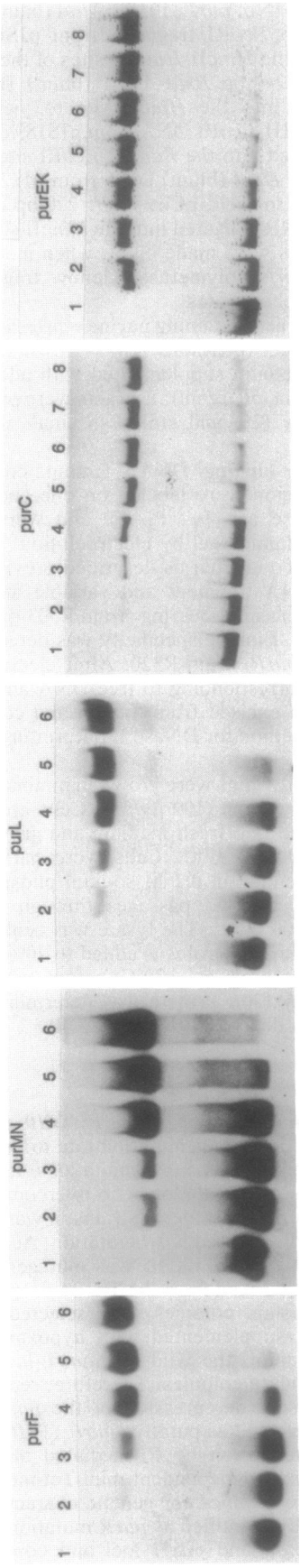

FIG. 3. Gel retardation assay for protein-DNA binding. (*purF*) 271-bp DNA fragment, 0, 0.63, 1.25, 2.5, 5.0, and 10 μg protein for lanes 1 to 6, respectively; (*purMN*) 434-bp fragment, 0, 0.63, 1.25, 2.5, 5.0, and 10 μg of protein for lanes 1 to 6, respectively; (*purL*) 137-bp fragment, 0, 0.63, 1.25, 2.5, 5.0, and 10 μg of protein for lanes 1 to 6, respectively; (*purC*) 330-bp fragment, 0, 0.45, 0.9, 1.8, 3.6, 5.4, 7.2, and 9.0 μg of protein for lanes 1 to 8, respectively; (*purEK*) 290-bp fragment, 0, 1.8, 3.6, 5.4, 7.2, 9.0, 13, and 18 μg of protein for lanes 1 to 8, respectively.

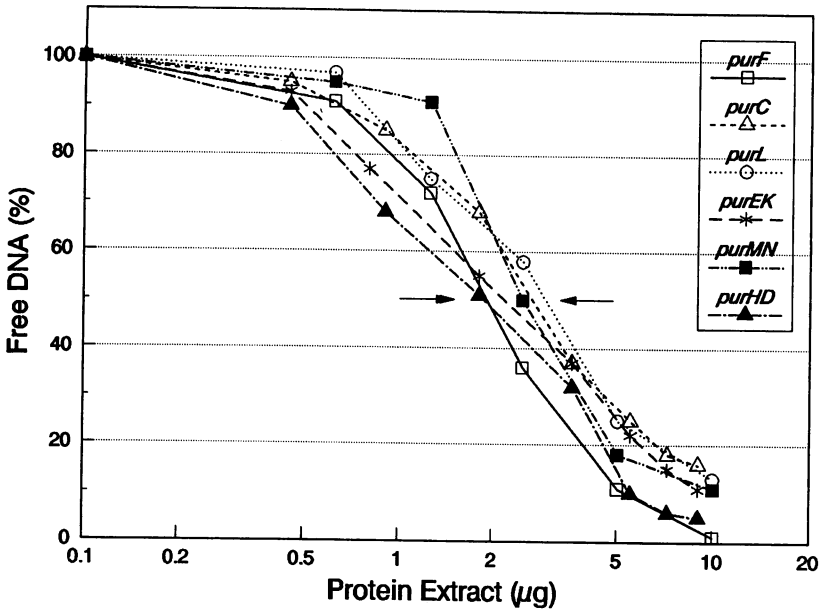

FIG. 4. Binding curves for repressor-*pur* gene control sites. Arrows point to positions corresponding to 50% unbound DNA.

monofunctional enzyme that catalyzes similar reactions in the path to IMP and in the AMP branch. We have constructed *lacZ* fusions to each of these loci in isogenic *purR* and *purR*$^+$ strains in order to assay regulation of gene expression by *purR*. These strains have permitted direct measurements of a well-characterized *purR* mutation on the expression of each of the loci involved in the synthesis of IMP. The measurements of β-galactosidase (Table 2) demonstrate that each locus except *purB* is regulated by *purR* over a 5- to 17-fold range. This coregulation thus defines a *pur* regulon containing *purF*, *purHD*, *purL*, *purMN*, *purE*, and *purC* as well as *purR* (19; R. J. Rolfes and H. Zalkin, submitted for publication). As suggested previously (29, 41), gene *purB* is not coregulated with genes for de novo synthesis of IMP. Recent work by Meng et al. (19) has provided evidence for 2.5-fold coregulation of *purB* by *purR*. Isolation of the *purB* control region is needed to determine whether it contains a PurR binding site. The data in Table 2 thus confirm earlier observations implicating *purR* in the control of the eight enzymes that are specifically involved in the pathway to IMP (reviewed in reference 21). Of the *pur* regulon genes, only *purF* is cotranscribed with nonpurine genes. The *purF* operon also contains genes *cpvA*, which is required for colicin V production (7), and *dedF*, a gene of unknown function (22).

Data summarized in Fig. 4 and 6 demonstrate that purine repressor binds to a conserved 16-bp operator site in the promoter region of each of the operons studied. The sequence of gene *purB* was not available, and this gene is not included in the survey. Presumably, the conserved bases in this operator consensus sequence, NCGCAAAC · GTT TNCNT, are important for binding of repressor (Fig. 2). The operator consensus sequence is a variant of the perfect dyad symmetry, ACGCAAAC · GTTTGCGT. There are 1-bp departures from the consensus in the operators for *purMN*, *purEK*, and *purC*. The *purMN* and *purC* operators bind repressor with somewhat lower affinity than the operators having no departures from the consensus sequence. It is of interest that the deviations from the operator consensus in *purL* and *purC* are in positions 1 and 2, respectively, of the right-hand symmetry. These positions were shown to be important for repressor binding to *purF* (24, 25). We cannot explain why a deviation in the equivalent position in the left-hand symmetry of the *purEK* operator does not adversely affect repressor binding.

The conserved 16-bp operator sequence is located between positions −46 and −13 relative to the start of transcription in the *purF*, *purL*, *purMN*, *purEK*, and *purHD* operons (Fig. 6). However, the transcription start site for *purC* actually lies within the operator sequence. For genes *purF*, *purL*, *purMN*, *purEK*, *purHD*, and *purC*, the promoter can be defined by the position of the −10 hexamer (overlined in Fig. 6) and the transcription start site. Mutational analysis has confirmed the identification of the *purF* −10 promoter element (24). In addition, we have overlined the expected position of the −35 promoter element (9) in the *pur* loci shown in Fig. 6. In each of the genes, three or more of the six positions for the −35 hexamer consensus sequence TTGACA are conserved. Mutational analysis has supported the assignment of the −35 promoter element in gene *purF* (24). However, in gene *purL* the spacing between the −35 and −10 hexamers is less than the optimal 17 to 18 bp, and in *purHD* and *purC* this spacing is 21 and 20 bp, respectively. These deviations imply that promoters in some of the *pur* loci may not utilize −35 elements for RNA polymerase recognition. It has been well documented that several *E. coli* and phage λ promoters can function without a −35 promoter element (4, 11, 23). In these cases, there is always an extended −10 region in which the sequence 5′-TGN precedes the −10 hexamer. Examination by mutational analysis indicates that the TG dinucleotide is important for function (4, 11, 23). Interestingly, the nucleotides TG can be replaced without loss of function if a synthetic −35 hexamer is correctly positioned, indicating that either an extended −10 element or a −35 hexamer can function in RNA polymerase recognition (4) and that other sequences in the promoter region such as an extended −10 region can compensate for a poor −35 element. The 5′-TGN sequence could participate in an extended −10 region in genes *purF*, *purMN*, *purEK*, and *purC*. A CGN variant is found in *purHD*. The position of

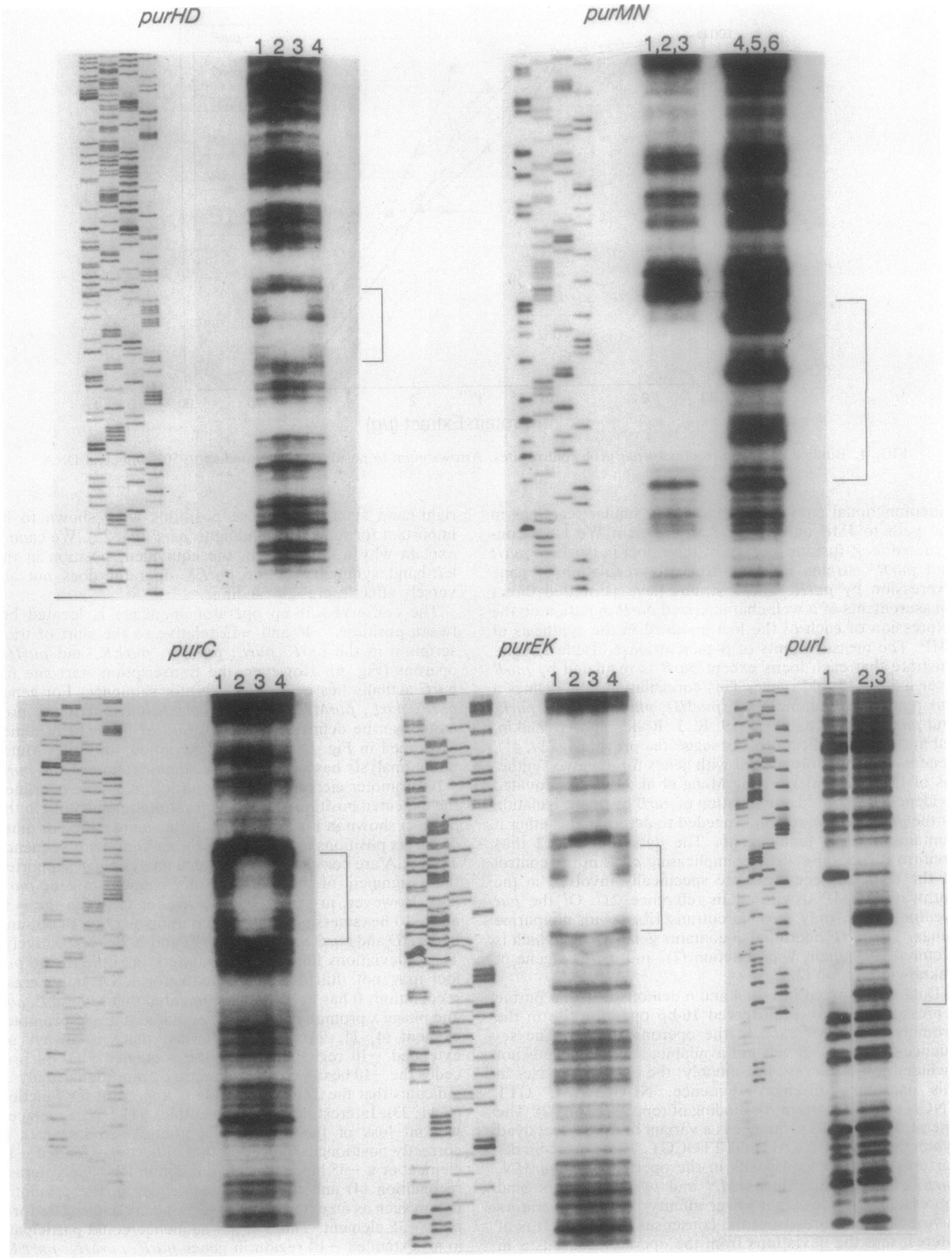
purHD
1 2 3 4
purMN
1,2,3
4,5,6
purC
1 2 3 4
purEK
1 2 3 4
purL
1
2,3

FIG. 5. DNase I footprints for interaction of repressor with *pur* gene control sites. For each experiment, 10 fmol of DNA fragment was 5′ end labeled at either the *Hin*dIII or *Eco*RI polylinker end from pUC118. (*purHD*) Lanes: 1 and 4, DNA; 2 and 3, DNA-protein. Sequencing ladder is *purC* plasmid pCo. (*purMN*) Lanes: 1 to 3, DNA-protein; 4 to 6, DNA. Sequencing ladder is *purL* plasmid pLo. (*purC*) Lanes: 1 and 4, DNA; 2 and 3, DNA-protein. Sequencing ladder is *purC* plasmid pCo. (*purEK*) Lanes: 1 and 4, DNA; 2 and 3, DNA-protein; sequencing ladder is *purEK* plasmid purEo. (*purL*) Lanes: 1, DNA-protein; 2 and 3, DNA. Sequencing ladder is *purL* plasmid pLo.

the C nucleotide in this potential extended −10 region may be restricted because it is a conserved position in the *purC* operator. There is at present no good explanation for why the highly expressed *E. coli pur* genes have relatively poor matches to the −35 promoter consensus sequence. One possibility that had been considered was that the sequence of the −35 region was constrained by the requirement for an overlapping *pur* operator (24). The summary in Fig. 6 shows, however, that the operators in *purHD* and *purC* do not extend to the −35 region, yet these genes have nonoptimally positioned −35 hexamers, with only three of six matches to the −35 hexamer consensus.

From the relative locations of the *pur* operator and promoter (Fig. 6), it is apparent that binding of purine repressor and RNA polymerase should be competitive. Thus, binding of repressor should inhibit gene expression by blocking transcription initiation. In the case of *purC*, there could be competitive binding if the proposed extended −10 region functioned in RNA polymerase recognition. Alternatively, purine repressor could prevent open complex formation in *purC*. In either case, binding of repressor would inhibit transcription initiation. In the simplest model, regulation of gene expression by *purR* would depend on the relative repressor-operator and RNA polymerase-promoter affinities. Quantitative measurements of binding affinities must await experiments with purified repressor. In addition to the limitations imposed by use of crude repressor, the binding curves shown in Fig. 4 were obtained in the absence of added purine or purine nucleotide coeffector. Coeffector-independent binding results from the crude repressor containing bound coeffector and from in vitro conditions that fortuitously bypass a coeffector requirement (Rolfes and Zalkin, unpublished).

In addition to controlling *pur* genes, *purR* expression is autoregulated (19) by noncooperative binding of repressor to the two operators shown in Fig. 2 (Rolfes and Zalkin, submitted). Other genes subject to *purR* control include *guaBA* (19), *glyA* (J. G. Steiert, R. J. Rolfes, H. Zalkin, and G. V. Stauffer, submitted for publication), and *codA*, which encodes the pyrimidine salvage enzyme cytosine deaminase (12). The connection of *purR* with pyrimidine metabolism also extends to de novo biosynthesis. Wilson et al. (38) noted sequences similar to the *pur* operator in the 5′-flanking regions of genes *pyrC*, *pyrD*, *carAB*, and *prsA*. Choi and Zalkin (5) have shown that binding of purine repressor to the *pur* operator in the *pyrC* promoter region represses expression twofold. Wilson and Turnbough (39) also reported *purR* regulation of *pyrC* as well as *pyrD* expression. Given the multiple roles for purine nucleotides in metabolism and

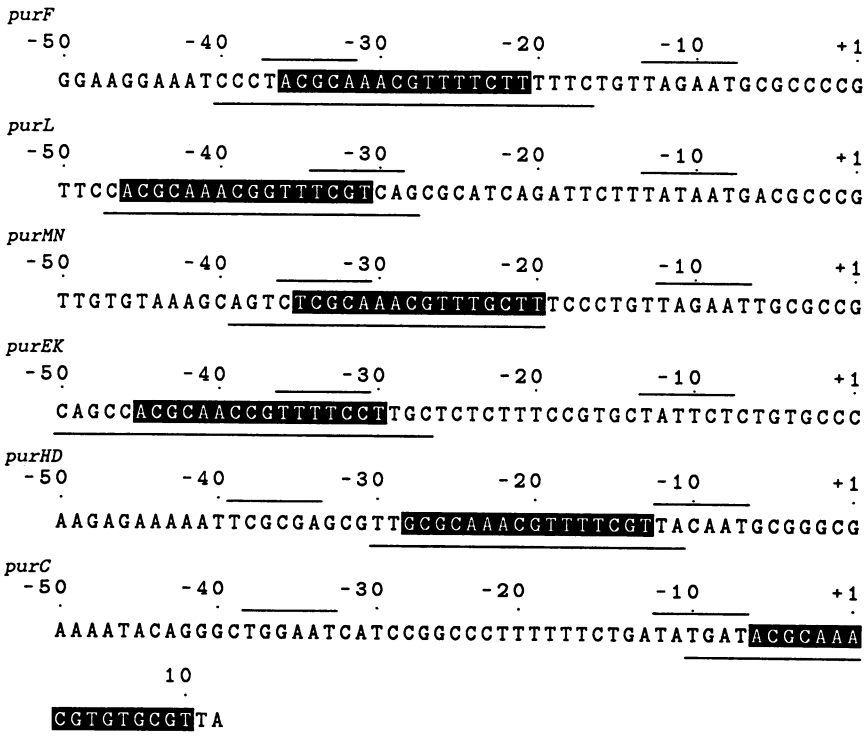

FIG. 6. Summary of DNase I footprinting of repressor interaction with *pur* gene control sites. Sequences are numbered from the transcription start site; regions corresponding to possible −10 and −35 promoter sites are overlined; regions protected by purine repressor are underlined; the *pur* operator is highlighted.

biosynthesis, cross-pathway regulation by *purR* may be one mechanism to coordinate metabolic functions in *E. coli*.

ACKNOWLEDGMENTS

This work was supported by Public Health grants GM24658 (H.Z.) and AI20068 (J.M.S.) from the National Institutes of Health.

LITERATURE CITED

1. **Aiba, A., and K. Mizobuchi.** 1989. Nucleotide sequence analysis of genes *purH* and *purD* involved in the *de novo* purine nucleotide biosynthesis of *Escherichia coli*. J. Biol. Chem. **264**:21239–21246.
2. **Bachmann, B. J.** 1972. Pedigrees of some mutant strains of *Escherichia coli* K-12. Bacteriol. Rev. **36**:525–557.
3. **Benson, C. E., and J. S. Gots.** 1976. Occurrence of a regulatory deficiency in purine biosynthesis among *purA* mutants of *Salmonella typhimurium*. Mol. Gen. Genet. **145**:31–36.
4. **Chan, B., and S. Busby.** 1989. Recognition of nucleotide sequences at the *Escherichia coli* galactose operon P1 promoter by RNA polymerase. Gene **84**:227–236.
5. **Choi, K. Y., and H. Zalkin.** 1990. Regulation of *Escherichia coli pyrC* by the purine regulon repressor protein. J. Bacteriol. **172**:3201–3207.
6. **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.
7. **Fath, M. J., H. K. Mahanty, and R. Kolter.** 1989. Characterization of a *purF* operon mutation which affects colicin V production. J. Bacteriol. **171**:3158–3161.

7a. **Flannigan, K. A., S. H. Hennigan, H. H. Vogelbacker, J. S. Gots, and J. M. Smith.** 1990. Purine biosynthesis in *Escherichia coli* K12: structure and DNA sequence studies of the *purHD* locus. Mol. Microbiol. **4**:381–392.

8. **Gots, J. S., C. E. Benson, B. Jochimsen, and K. R. Koduri.** 1977. Microbial models and regulatory elements in the control of purine metabolism. CIBA Found. Symp. **48**:23–41.
9. **Hawley, D., and W. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11**:2237–2256.
10. **Houlberg, U., and K. F. Jensen.** 1983. Role of hypoxanthine and guanine in regulation of *Salmonella typhimurium* and guanine in regulation of *Salmonella typhimurium pur* gene expression. J. Bacteriol. **153**:837–845.
11. **Keilty, S., and M. Rosenberg.** 1987. Constitutive function of a positively regulated promoter reveals new sequences essential for activity. J. Biol. Chem. **262**:6389–6395.
12. **Kilstrup, M., L. M. Meng, J. Neuhard, and P. Nygaard.** 1989. Genetic evidence for a repressor of synthesis of cytosine deaminase and purine biosynthesis enzymes in *Escherichia coli*. J. Bacteriol. **171**:2124–2127.
13. **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**:494–508.
14. **Layne, E.** 1957. Spectrophotometric and turbidometric methods for measuring proteins. Methods Enzymol. **3**:448–451.
15. **Levine, R. A., and M. W. Taylor.** 1981. Selection for purine regulatory mutants in an *E. coli* hypoxanthine phosphoribosyltransferase-guanine phosphoribosyltransferase double mutant. Mol. Gen. Genet. **181**:313–318.
16. **Levine, R. A., and M. W. Taylor.** 1981. Mechanisms of adenine toxicity in *Escherichia coli*. J. Bacteriol. **149**:923–930.
17. **Levine, R. A., and M. W. Taylor.** 1982. Regulation of *purE* transcription in a *purE*::*lac* fusion strain of *Escherichia coli*. J. Bacteriol. **149**:1041–1049.
18. **Makaroff, C. A., and H. Zalkin.** 1985. Regulation of *Escherichia coli purF*. Analysis of the control region of a *pur* regulon gene. J. Biol. Chem. **260**:10378–10387.
19. **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.
20. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **Neuhard, C. A., and P. Nygaard.** 1987. Purines and pyrimidines, p. 445–473. *In* F. C. Neidhardt, J. L. Ingraham, K. P. 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.
22. **Nonet, M. L., C. C. Marvel, and D. R. Tolan.** 1987. The *hisT-purF* region of the *E. coli* K-12 chromosome. Identification of additional genes of the *hisT* and *purF* operons. J. Biol. Chem. **262**:12209–12217.
23. **Ponnambalam, S., B. Chan, and S. Busby.** 1988. Functional analysis of different sequence elements in the *Escherichia coli* galactose operon P2 promoter. Mol. Microbiol. **2**:165–172.
24. **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.
25. **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.
26. **Schendel, F. J., E. Mueller, J. Stubbe, A. Shiau, and J. M. Smith.** 1989. Formyl-glycinamide ribonucleotide synthetase from *Escherichia coli*: cloning, sequencing, overproduction, isolation and characterization. Biochemistry **28**:2459–2471.
27. **Shiau, A., and J. M. Smith.** 1988. Improved cat gene cassette for promoter analysis and genetic constructions. Gene **67**:295–299.
28. **Smith, J. M., and H. A. Daum.** 1986. Nucleotide sequence of the *purM* gene encoding 5′-phosphoribosyl-5-aminoimidazole synthetase of *Escherichia coli* K12. J. Biol. Chem. **261**:10632–10636.
29. **Thomulka, K. W., and J. S. Gots.** 1982. Isolation and characterization of purine regulatory mutants of *Salmonella typhimurium* with an episomal *purE-lac* fusion. J. Bacteriol. **151**:153–161.
30. **Tiedeman, A. A., J. Keyhani, J. Kamholz, H. A. Daum, J. Gots, and J. M. Smith.** 1989. Nucleotide sequence analysis of the *purEK* operon encoding 5′-phosphoribosyl-5-aminoimidazole carboxylase of *Escherichia coli* K-12. J. Bacteriol. **171**:205–212.
31. **Tiedeman, A. A., and J. M. Smith.** 1988. *lacZY* gene fusion cassettes with KanR resistance. Nucleic Acids Res. **16**:3587.
32. **Tso, J. Y., H. Zalkin, M. van Cleemput, C. Yanofsky, and J. M. Smith.** 1981. Nucleotide sequence of *Escherichia coli purF* and deduced amino acid sequence of glutamine phosphoribosylpyrophosphate amidotransferase. J. Biol. Chem. **257**:3525–3531.
33. **Vidal-Ingigliardi, D., and O. Raibaud.** 1985. A convenient technique to compare the efficiency of promoters in *Escherichia coli*. Nucleic Acids Res. **13**:5919–5926.
34. **Vieira, J., and J. Messing.** 1987. Production of single-stranded plasmid DNA. Methods Enzymol. **153**:3–11.
35. **Vogel, H. J., and D. M. Bonner.** 1956. Acetylornithinase of *Escherichia coli*: partial purification and some properties. J. Biol. Chem. **218**:97–106.
36. **Watanabe, W., G.-I. Sampei, A. Aiba, and K. Mizobuchi.** 1989. Identification and sequence analysis of *Escherichia coli purE* and *purK* genes encoding 5′-phosphoribosyl-5-amino-4-imidazole carboxylase for de novo purine biosynthesis. J. Bacteriol. **171**:198–204.
37. **Way, J. C., M. A. Davis, D. Morisato, D. E. Roberts, and N. Kleckner.** 1984. New Tn10 derivatives for transposon mutagenesis and for construction of *lacZ* operon fusions by transposition. Gene **32**:369–379.
38. **Wilson, H. R., P. T. Chan, and C. L. Turnbough, Jr.** 1987. Nucleotide sequence and expression of the *pyrC* gene of *Escherichia coli* K-12. J. Bacteriol. **169**:3051–3058.
39. **Wilson, H. R., and C. L. Turnbough, Jr.** 1990. Role of the purine repressor in regulation of pyrimidine gene expression in *Escherichia coli*. J. Bacteriol. **172**:3208–3213.
40. **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.
41. **Wolfe, S. A., and J. M. Smith.** 1985. Separate regulation of *purA* and *purB* loci of *Escherichia coli* K-12. J. Bacteriol. **162**: 822–925.