JOURNAL OF BACTERIOLOGY, July 1987, p. 3051–3058
0021-9193/87/073051-08$02.00/0

Vol. 169, No. 7

Nucleotide Sequence and Expression of the *pyrC* Gene of *Escherichia coli* K-12

HELEN R. WILSON, PETER T. CHAN, AND CHARLES L. TURNBOUGH, JR.*

Department of Microbiology, University of Alabama at Birmingham, Birmingham, Alabama 35294

Received 15 December 1986/Accepted 30 March 1987

The *pyrC* gene of *Escherichia coli* K-12, which encodes the pyrimidine biosynthetic enzyme dihydroorotase, was cloned as part of a 1.6-kilobase-pair chromosomal fragment. The nucleotide sequence of this fragment was determined. An open reading frame encoding a 348-amino acid polypeptide (M_r = 38,827) was identified as the *pyrC* structural gene by comparing the amino acid composition predicted from the DNA sequence with that previously determined for the dihydroorotase subunit. The *pyrC* promoter was mapped by primer extension of in vivo transcripts. Transcriptional initiation was shown to occur within a region located 36 to 39 base pairs upstream of the *pyrC* structural gene. Pyrimidine availability appears to affect the use of the minor transcriptional initiation sites. The level of *pyrC* transcription and dihydroorotase synthesis was coordinately derepressed by pyrimidine limitation, indicating that regulation occurs, at least primarily, at the transcriptional level. Inspection of the *pyrC* nucleotide sequence indicates that gene expression is not regulated by an attenuation control mechanism similar to that described for the *pyrBI* operon and the *pyrE* gene. A possible mechanism of transcriptional control involving a common repressor protein is suggested by the identification of a highly conserved, operatorlike sequence in the promoter regions of *pyrC* and the other pyrimidine genes (i.e., *pyrD* and *carAB*) whose expression is negatively regulated by a cytidine nucleotide effector.

In *Escherichia coli* and *Salmonella typhimurium*, the de novo synthesis of UMP, the precursor of all pyrimidine nucleotides, is catalyzed by six enzymes encoded by six unlinked genes and operons. The expression of these genes and operons is noncoordinately regulated by the intracellular levels of pyrimidine nucleotides. The expression of *pyrBI* (designated *pyrB* in *S. typhimurium*), *pyrE*, and *pyrF* appears to be repressed by a uridine nucleotide, whereas *pyrC* and *pyrD* expression appears to be repressed primarily by a cytidine nucleotide (21, 30). The expression of *carAB* (designated *pyrA* in *S. typhimurium*), which is essential for both pyrimidine and arginine biosynthesis, is subject to cumulative repression by pyrimidine (apparently uridine and cytidine) nucleotides and arginine (1, 21). Recent studies indicate that *pyrBI* and *pyrE* expression is regulated by similar attenuation control mechanisms (5, 7, 13, 19, 23, 26, 27, 31). In these mechanisms, transcriptional termination at a Rho-independent terminator (attenuator) immediately preceding the *pyr* structural gene(s) is regulated by the relative rates of UTP-sensitive transcription and coupled translation within a leader region upstream of the attenuator. In additional studies, the promoter-regulatory regions of the *carAB* operon (6, 22) and the *pyrD* (12) and *pyrF* (C. L. Turnbough, Jr., K. H. Kerr, W. R. Funderburg, J. P. Donahue, and F. E. Powell, J. Biol. Chem., in press) genes of *E. coli* K-12 and also the *pyrC* gene of *S. typhimurium* (20) were identified and sequenced. Inspection of the nucleotide sequences indicated that attenuation control similar to that described above is not involved in the expression of these genes. Essentially nothing else is known about the pyrimidine-sensitive control mechanisms that regulate *carAB*, *pyrC*, *pyrD*, and *pyrF* expression.

In this paper we present initial results of a study designed to elucidate the mechanism controlling the expression of the *pyrC* gene of *E. coli* K-12, which encodes the pyrimidine biosynthetic enzyme dihydroorotase. We report the cloning and nucleotide sequence of the *pyrC* gene, the mapping of the *pyrC* transcriptional initiation sites, and the effect of pyrimidine availability on *pyrC* transcription in vivo. In addition, we identify an operatorlike sequence present only in the *pyrC*, *pyrD*, and *carAB* promoter regions which may be involved in cytidine nucleotide-mediated regulation.

MATERIALS AND METHODS

Bacterial strains. *E. coli* K-12 strains CLT9 [F⁻ *araD139* Δ(*argF-lac*)*U169 rpsL150 thiA1 relA1 deoC7 ptsF25 flbB5301 rbsR pyrB476*::Mu d1(Ap^r *lac* cts62)], CLT39 [F⁻ *araD139* Δ(*argF-lac*)*U169 rpsL150 thiA1 relA1 deoC7 ptsF25 flbB5301 rbsR pyrB477 srl-300*::Tn*10 recA56*], and CLT49 (*thi-1 pyrC46 relA1 lacZ43 srl-300*::Tn*10 recA56* λ⁻) were used in this study. Strain CLT9 was constructed as previously described (26). Strains CLT39 and CLT49 were constructed by transducing strains CLT19 (26) and 30SOU6 (MA1008; CGSC 5153) (4), respectively, to Tet^r with bacteriophage P1 grown on strain JC10240 (HfrPO45 *srl-300*::Tn*10 recA56 thr-300 ilv-318 spc-300*). The cotransduced *recA56* allele was identified by screening for UV sensitivity (18).

DNA preparations. Plasmid DNA and DNA restriction fragments were prepared as described previously (26). Chromosomal DNA was prepared from a 250-ml stationary-phase culture of strain CLT9 grown at 30°C in LB medium (18). Cells were harvested by centrifugation, washed with 50 mM Tris hydrochloride (pH 7.8)–5 mM EDTA–50 mM NaCl, and suspended in 4 ml of ice-cold 25% sucrose–50 mM Tris hydrochloride (pH 8.0)–1 mM EDTA. The suspension was incubated on ice for 5 min, and 0.8 ml of a freshly prepared egg white lysozyme solution (5 mg/ml in 0.25 M Tris hydrochloride, pH 8.0) was added. After 5 more min on ice, 1.6 ml

* Corresponding author.

of 0.25 M EDTA (pH 8.0) was added. The sample was incubated for an additional 15 min on ice, and 0.5 ml of 25% sodium dodecyl sulfate was added. The sample was heated at 60°C for 10 min, cooled to room temperature, and mixed vigorously with a vortex mixer for 1 min. The DNA was then banded twice by CsCl density gradient centrifugation and dialyzed against 10 mM Tris hydrochloride (pH 7.8)–1 mM EDTA.

Restriction digests, ligations, and transformations. Conditions for restriction digests, ligations, and transformations were as previously described (26).

Media and culture methods. Cells used for enzyme assays and for the isolation of RNA were grown in N^-C^- medium (2) supplemented with 10 mM NH_4Cl, 0.4% glucose, 0.015 mM thiamine hydrochloride, 100 μg of ampicillin per ml [with strain CLT39(pBHM105) only], and either 1 mM uracil or 0.24 mM UMP. Cultures (100 ml in a 500-ml flask) were grown at 30°C with shaking. The solid media used were LB (with ampicillin or tetracycline added at 25 μg/ml when required) and VBCG (minimal glucose) (32) containing 1.5% Difco agar. Growth on solid media was at 37°C.

Dihydroorotase assay. Cultures were grown to an A_{650} of 0.5, and 30-ml samples were taken. Cells were collected by centrifugation (4°C), washed with ice-cold 50 mM sodium phosphate (pH 7.0), and stored at −70°C for 1 to 2 days without loss of enzymatic activity. Cells were resuspended in 5 ml of 50 mM sodium phosphate (pH 7.0) and disrupted by sonic oscillation at 0°C. Cell debris was removed by centrifugation at 27,000 × *g* for 30 min at 4°C. Dihydroorotase activity in the cell extracts was measured by the rate of conversion of dihydroorotate to carbamyl aspartate. Reaction mixtures (0.5 ml) contained 2 mM dihydroorotate, 2 mM EDTA, 0.1 M Tris hydrochloride (pH 8.6), and 50 μl of cell extract diluted appropriately in 50 mM sodium phosphate (pH 7.0). Reactions were initiated by adding cell extract, incubated at 30°C, and stopped at various times (15 to 30 min) by adding 1 ml of color mix (24) and 0.5 ml of water. Color mix and water were also added to a series of standards that were identical to the reaction mixes except that they contained known amounts of carbamyl aspartate (0 to 100 nmol) and no cell extract. Color was developed in the samples as described previously (24) and read at 466 nm. The amount of carbamyl aspartate formed in the assays was determined by comparison with the standards. Assay conditions were chosen so that the amount of carbamyl aspartate formed was directly proportional to enzyme concentration and time. Protein in the cell extracts was measured by the method of Lowry et al. (14), using crystalline bovine serum albumin as the standard.

Quantitation of plasmid DNA. Cells from triplicate 10-ml samples of culture were harvested by centrifugation (4°C) and stored at −70°C until analyzed. Plasmid DNA was extracted, and relative levels were measured by densitometric scanning of photographic negatives of ethidium bromide-stained agarose gels as previously described (26).

Isolation of cellular RNA. Cellular RNA was prepared by a procedure similar to that described by Hagen and Young (8). A 25-ml sample was removed from a culture of strain CLT39(pBHM105) (A_{650} = 0.5) and was added directly to 2.5 ml of 0.5 M Tris hydrochloride (pH 6.8)–20 mM EDTA–10% sodium dodecyl sulfate in a flask held in a boiling water bath. After 2 min, the flask was removed from the water bath and allowed to cool to room temperature, and 2.5 ml of 2 M sodium acetate (pH 5.2) was added to the cell lysate. The lysate was extracted twice with an equal volume of water-saturated phenol and once with an equal volume of chloroform. The aqueous phase was dialyzed overnight against 2 liters of diethyl pyrocarbonate-treated (16) sterile water (4°C) to remove interfering phosphate contributed by the growth medium. The RNA was precipitated from the dialysate by adding 0.1 volume of 3 M sodium acetate (pH 5.2) and 2.5 volumes of 95% ethanol and allowing the sample to stand overnight at −20°C. The precipitate was collected by centrifugation (4°C) and dissolved in 0.4 ml of 10 mM Tris hydrochloride (pH 8.0)–1 mM EDTA. The RNA was reprecipitated by adding sodium acetate and 95% ethanol as described above and placing the sample at −70°C for 1 h. The precipitate was collected by centrifugation (4°C), washed with 70% ethanol (−20°C), dried in vacuo, and dissolved in 0.2 ml of 50 mM sodium acetate (pH 6.5)–10 mM $MgCl_2$–2 mM $CaCl_2$. Contaminating DNA was removed from the sample by adding RNase-free DNase I (5 μg/ml; Bethesda Research Laboratories, Inc., Gaithersburg, Md.) and incubating at 37°C for 15 min. The sample was then extracted once with an equal volume of water-saturated phenol and twice with an equal volume of chloroform. The RNA was precipitated, collected, washed, and dried as described in the previous precipitation step. The RNA was dissolved in 0.2 ml of diethyl pyrocarbonate-treated sterile water and stored at −20°C. Each step of the RNA isolation procedure was performed quantitatively to permit a comparison of *pyrC* transcript levels in cells grown under different conditions.

Primer extension mapping of *pyrC* transcripts. The 5′ termini of *pyrC* transcripts were mapped by primer extension essentially as described previously (15). Cellular RNA isolated from strain CLT39(pBHM105) grown on uracil or UMP was used as a source of *pyrC* transcripts. The amount of cellular RNA added to the reaction mixtures was either 15 μg from uracil-grown cells or 8.6 μg from UMP-grown cells. Although the amounts of cellular RNA added are different, these quantities were isolated from the same volume of culture. The higher yield of total RNA from uracil-grown cells, which was observed in three independent experiments, is presumably due to the increased synthesis of stable RNA in faster-growing cells. Cellular RNA was mixed with 16 ng (8×10^5 dpm) of a 5′ ^{32}P-end-labeled oligodeoxyribonucleotide DNA synthesis primer (see text) which was synthesized with an Applied Biosystems model 380A DNA synthesizer and end labeled as described previously (17). This mixture was dried in vacuo and redissolved in 25 μl of hybridization buffer containing 20 mM Tris hydrochloride (pH 8.0), 0.1 M NaCl, and 0.1 mM EDTA. The sample was heated at 100°C for 2 min and then allowed to hybridize at 50°C for 5 h. The sample was cooled to room temperature, and 25 μl of 2× reverse transcriptase reaction mix (0.2 M Tris hydrochloride [pH 8.3], 20 mM $MgCl_2$, 0.1 M KCl, 20 mM dithiothreitol, 1 mM each dATP, dGTP, dCTP, and dTTP) containing 12.5 U of avian myeloblastosis virus reverse transcriptase (Boehringer Mannheim Biochemicals, Indianapolis, Ind.) was added. The reaction mixture was incubated for 1 h at 42°C. The sample was quantitatively extracted once with an equal volume of water-saturated phenol and once with an equal volume of chloroform. The primer extension products were precipitated by adding 0.1 volume of 3 M sodium acetate and 2.5 volumes of 95% ethanol and placing the sample in a dry ice-ethanol bath for 15 min. The precipitate was collected by centrifugation (4°C), washed with 70% ethanol (−20°C), and dried in vacuo. The dried sample was dissolved in DNA-sequencing dye (17), and a portion was run on a 10% polyacrylamide sequencing gel (17) alongside an appropriate DNA sequencing ladder. Under the condi-

tions described, it was shown that there is a linear relationship between the amount of cellular RNA preparation added to the primer extension reaction mixture and the level of extended fragments synthesized by reverse transcriptase.

RESULTS

Cloning the *pyrC* gene. The *pyrC* gene was cloned from chromosomal DNA isolated from *E. coli* K-12 CLT9. *Ava*I-cut chromosomal DNA was ligated to *Ava*I-cut plasmid pBR322, and a *pyrC*-containing recombinant plasmid was isolated by transforming strain CLT49 (*pyrC46*) with selection for pyrimidine prototrophy. This plasmid, designated pBHM52 (Fig. 1), contains three *Ava*I fragments of chromosomal DNA and two copies of pBR322. The *pyrC* gene was subcloned from plasmid pBHM52 as outlined in Fig. 1. All plasmids described were isolated by transforming strain CLT49 and selecting for pyrimidine prototrophy. The chromosomal fragment carrying the *pyrC* gene was reduced to 1.6 kilobase pairs in plasmids pBHM105 and pBHM107.

Two additional plasmids were constructed by inverting the chromosomal insert in plasmids pBHM105 and pBHM107. Both plasmids were shown to complement the *pyrC46* mutation in strain CLT49, indicating that the *pyrC* promoter is included on the 1.6-kilobase-pair chromosomal fragment.

Nucleotide sequence of the *pyrC* gene. The sequence of the entire *pyrC*-containing chromosomal insert of plasmid pBHM107 (Fig. 1) was determined as summarized in Fig. 2. The nucleotide sequence is shown in Fig. 3. A single open reading frame capable of encoding a 38-kilodalton polypeptide, which is the size of the dihydroorotase subunit measured by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (34), was located between base pairs (bp) 467 and 1510 (Fig. 3). This open reading frame begins with an ATG codon preceded by two Shine-Dalgarno-like sequences (SD_1 and SD_2 in Fig. 3) (11) and would encode a 348-amino acid polypeptide with a molecular weight of 38,827. The assignment of this open reading frame as the *pyrC* structural gene was confirmed by comparing the amino acid composition predicted from the DNA sequence and that determined for the dihydroorotase subunit (34). The predicted and determined values are nearly identical (Table 1). The end of the *pyrC* structural gene is followed closely by a region of hyphenated dyad symmetry (bp 1520 to 1584), which may be involved in the termination (33) or nucleolytic processing (25, 28) of the *pyrC* transcript.

A second long open reading frame was identified in the sequence of the chromosomal insert of plasmid pBHM107 (bp 1 to 358). This open reading frame, which extends beyond the sequence shown in Fig. 3, is presumably part of another structural gene in this region. The end of this second open reading frame is followed by a sequence typical of a Rho-independent transcriptional terminator (bp 367 to 392) (33).

Mapping the *pyrC* promoter and transcriptional regulation. The *pyrC* promoter was located by primer extension mapping the 5′ termini of *pyrC* transcripts as described in Materials and Methods. The cellular RNA used as a source of *pyrC* transcripts was isolated from the pyrimidine-auxotrophic strain CLT39(pBHM105) grown with either uracil or UMP as the sole pyrimidine source. Growth on uracil causes repressed *pyr* gene expression, while growth on UMP, which is only slowly utilized by cells under the present conditions, causes derepressed *pyr* gene expression. A synthetic 5′ ^{32}P-end-labeled oligonucleotide complementary to nucleotides 473 to 497 in Fig. 3 was used as the DNA

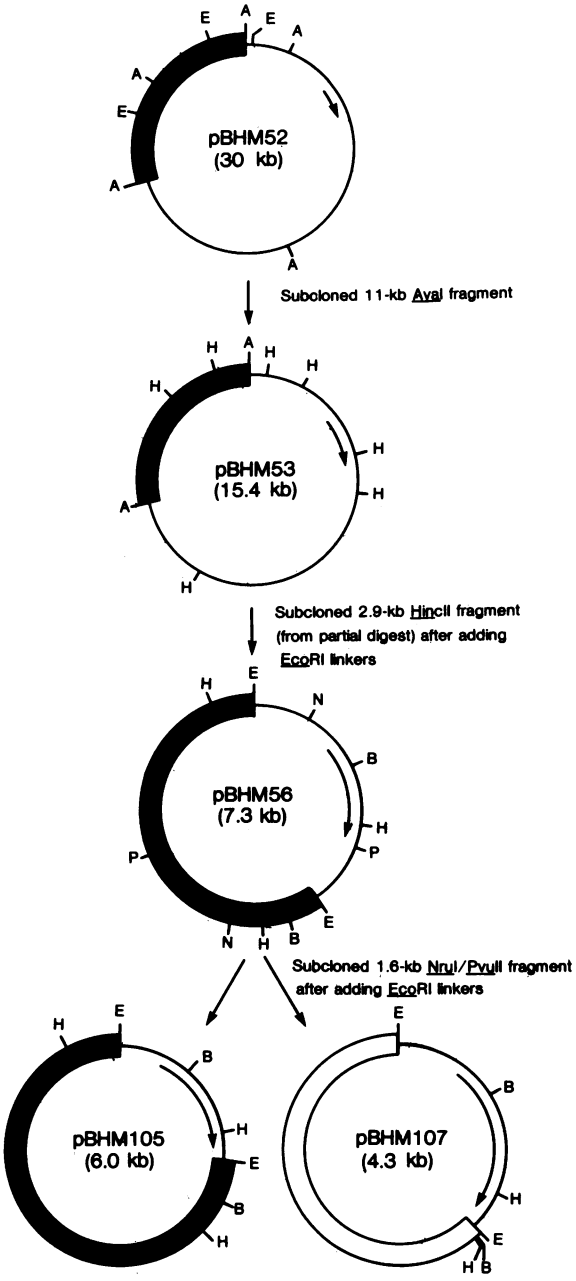

FIG. 1. Cloning the *pyrC* gene of *E. coli* K-12. All plasmids constructed during the cloning and subcloning of the *pyrC* gene are described, and the subcloning protocol is summarized. *E. coli* chromosomal DNA is represented by the thin line, with the arrow indicating the position and direction of transcription of the *pyrC* structural gene. The filled bar represents plasmid pBR322 DNA, and the open bar indicates plasmid pUC18 DNA. Cleavage sites for restriction endonucleases are indicated as follows: A, *Ava*I; B, *Bam*HI; E, *Eco*RI; H, *Hin*cII; N, *Nru*I; and P, *Pvu*II. kb, Kilobase pairs.

synthesis primer. The DNA fragments synthesized by reverse transcriptase were analyzed on a sequencing gel (Fig. 4) containing a dideoxy sequencing ladder that was generated by using the same oligonucleotide primer and a template DNA containing bp 1 to 713 in Fig. 3. Because comigrating fragments in all the lanes of this gel are identical, the sequence of each fragment synthesized in the primer exten-

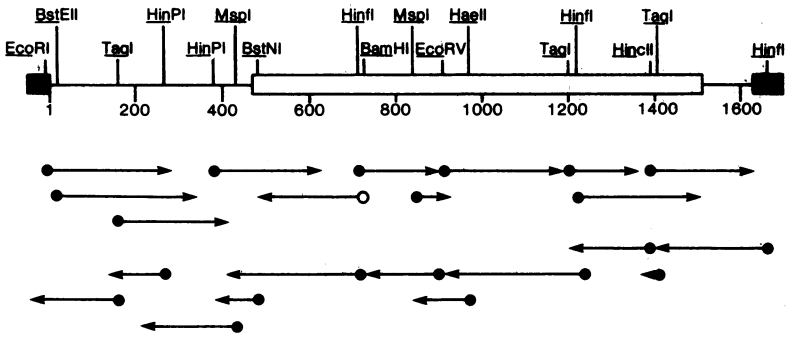

FIG. 2. Strategy and restriction sites used to sequence the *E. coli* K-12 *pyrC* gene and flanking regions cloned in plasmid pBHM107. The filled bars indicate pUC18 and linker DNA, which flank the 1.6-kilobase-pair chromosomal insert of plasmid pBHM107. The open bar represents the *pyrC* coding sequence. Nucleotide sequencing was done by the method of Maxam and Gilbert (17). Arrows indicate the direction and extent of each sequence determination. Arrows beginning with filled or open circles indicate sequences that were determined with 5′ or 3′ ^{32}P-end-labeled restriction fragments, respectively.

sion reactions could be determined by identifying the comigrating fragment in the sequencing ladder. The 3′ termini of the primer extension fragments are complementary to the 5′ termini of the *pyrC* transcripts.

The data indicate that in cells grown on uracil or UMP the primary *pyrC* transcriptional initiation site is at position 431 in Fig. 3. (Note that the sequences shown in Fig. 3 and 4 are of opposite strands.) There appear to be three minor transcriptional initiation sites at positions 428 to 430. The percentage of transcripts initiated at positions 428 and 429 appears to be much greater in uracil-grown cells than in UMP-grown cells. The transcriptional initiation sites are preceded by a −10 sequence and a −35 sequence (Fig. 3) that are typical of an *E. coli* promoter (9). The only other readily detectable extended fragment in Fig. 4 could indicate a minor transcriptional initiation site at position 440. Because this site is not preceded by a promoterlike sequence, however, it is presumed that this fragment is the result of premature termination by reverse transcriptase. The four *pyrC* transcriptional initiation sites corresponding to bp 428 to 431 also were identified by S1 nuclease mapping (35) (data not shown). It was not possible to identify major and minor transcriptional initiation sites by S1 nuclease mapping, however, because the relative amounts of S1 nuclease-resistant DNA-RNA hybrids were dependent on the reaction conditions used.

The effect of pyrimidine availability on the level of *pyrC* transcription was also measured from the data shown in Fig. 4. In the two primer extension reactions analyzed, the amount of cellular RNA added to each reaction mixture was isolated from the same number of cells. Under the assay conditions used, the level of primer extension products should correspond to the amount of *pyrC* transcripts present in the cells. The results indicate that the level of *pyrC* transcripts in cells grown on UMP was approximately fivefold greater than that in cells grown on uracil. The level of dihydroorotase was measured in the same cultures used to isolate the cellular RNA. The total activity of dihydroorotase in UMP-grown cells was 5.7-fold higher than that in uracil-grown cells (Table 2). These results indicate that *pyrC* expression is regulated, at least in large part, at the transcriptional level.

To confirm that the regulation observed in strain CLT39(pBHM105) was the same as that in an *E. coli* strain that carries only the chromosomal *pyrC* gene, dihydroorotase levels in strain CLT39 grown on uracil or UMP were measured (Table 2). After correction for plasmid copy number in strain CLT39(pBHM105), the level of derepression of *pyrC* expression in both strains was approximately 12-fold (Table 2).

Identification of a possible operator sequence. The expression of *pyrC* and *pyrD* is negatively regulated over a similar range (30) by a cytidine nucleotide, which suggests that a common repressor protein may be involved in regulation. Such a repressor protein would presumably bind to a highly conserved sequence (i.e., an operator) in the promoter-regulatory region of the two genes. To identify a possible operator sequence, the nucleotide sequences of the *pyrC* and *pyrD* structural genes and flanking regions were compared by using the COMPARE and BESTFIT computer programs supplied by the University of Wisconsin Genetics Computer Group. The comparison showed that although the two sequences are not strikingly similar overall, there is a highly conserved (16 of 19 bp) sequence present in the promoter regions. This highly conserved sequence is located between the −10 and −35 hexamers of the *pyrC* promoter and is found 15 bp upstream of the putative −35 sequence of the *pyrD* promoter (Fig. 5). The conserved sequence contains a hyphenated dyad symmetry (Fig. 3), which is characteristic of many operator sequences (10).

To determine whether other pyrimidine genes contained the highly conserved, operatorlike sequence or a closely related sequence, all the *E. coli* pyrimidine gene and flanking region sequences were screened. No sequences strongly resembling the highly conserved sequence were found in the *pyrBI*, *pyrE*, or *pyrF* regions. A closely related sequence was found in the opposite orientation near the *carAB* promoter region. This sequence is located 85 bp upstream of the pyrimidine-regulated *carAB* promoter (Fig. 5). The highly conserved sequence was also found in the *S. typhimurium* *pyrC* promoter (Fig. 5). This result was expected and is not particularly useful in identifying a possible operator sequence, because the sequence of the entire *pyrC* region of *E. coli* and *S. typhimurium* are highly conserved.

DISCUSSION

The results presented in this study provide an initial step toward the understanding of the mechanism controlling *pyrC* expression in *E. coli* K-12. The nucleotide sequence determined in this paper shows that the *pyrC* leader region and the beginning of the *pyrC* structural gene do not contain a Rho-independent terminator, indicating that regulation does not involve attenuation control similar to that described for

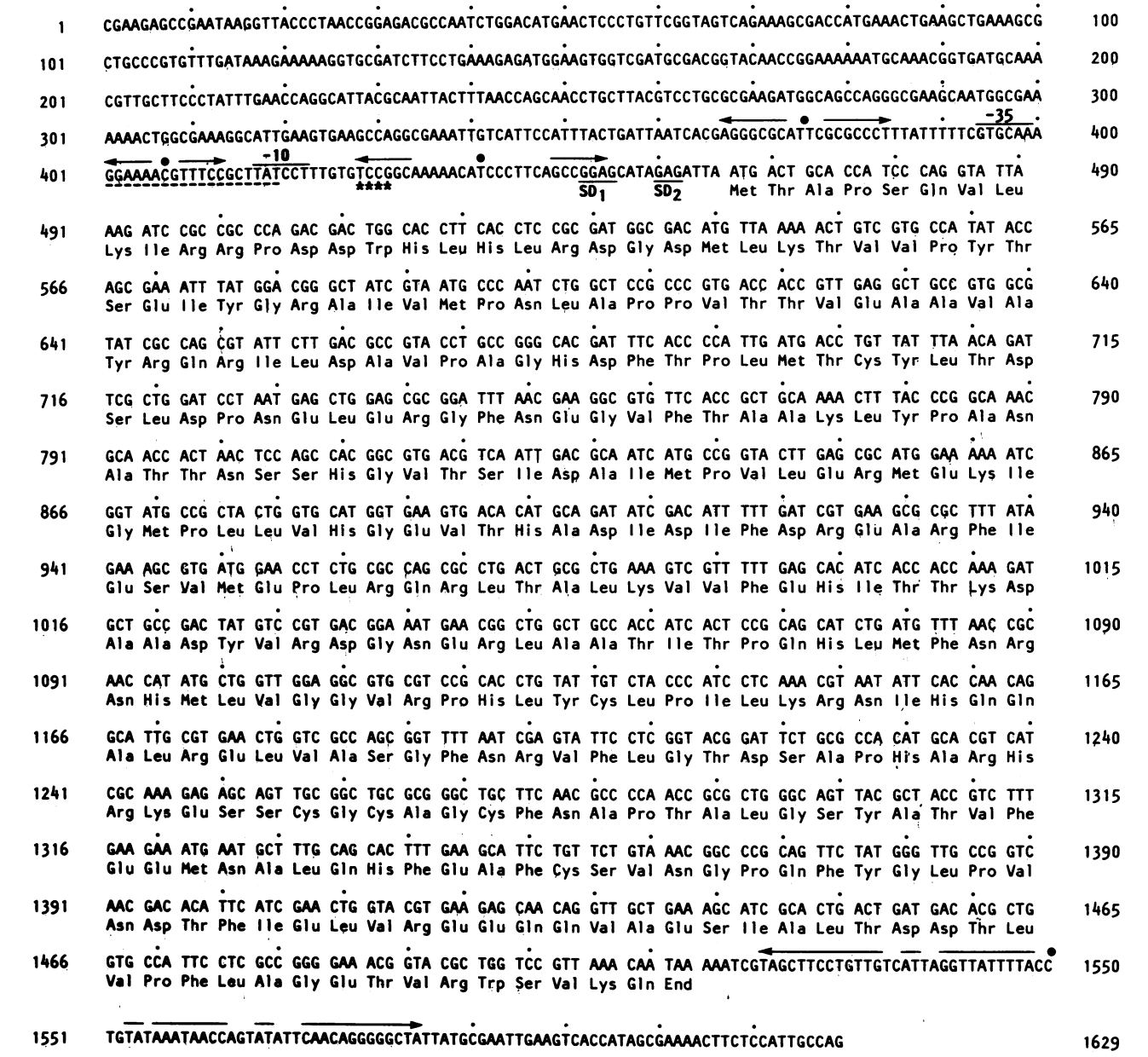

FIG. 3. Nucleotide sequence of the 1,629-bp chromosomal insert of plasmid pBHM107 containing the *pyrC* gene of *E. coli* K-12. Only the sequence of the antisense strand is shown; numbering is from the 5′ end. The deduced amino acid sequence of dihydroorotase is shown. Two possible *pyrC* Shine-Dalgarno (SD) sequences are underlined and labeled. The −10 and −35 hexamers of the *pyrC* promoter are overlined and labeled. The asterisks below nucleotides 428 to 431 indicate *pyrC* transcriptional initiation sites. The sequence underlined with a dashed line is a possible operator sequence, which is discussed in detail in the text. Dyad symmetries are indicated by the arrows with the centers of symmetry shown by the large dots.

the *pyrBI* operon and the *pyrE* gene. A Rho-independent termination sequence is present immediately upstream of the *pyrC* promoter; however, this sequence appears to be involved in the termination of transcription of an adjacent structural gene. There is no indication that the expression of this adjacent gene, which apparently is not intact in plasmid pBHM105, is required for *pyrC* expression and regulation (Table 2).

The quantitation of *pyrC* transcripts in cells grown under conditions of pyrimidine excess or limitation indicates that the regulation of *pyrC* expression occurs, at least primarily, at the transcriptional level. A possible mechanism for transcriptional control involving a common repressor protein is suggested by the identification of a highly conserved, operatorlike sequence in or near the promoter regions of the pyrimidine genes whose expression is regulated by a cytidine nucleotide effector (i.e., *pyrC*, *pyrD*, and *carAB*). For *pyrC*, the binding of a repressor protein to the possible operator sequence, which is included in the promoter, should prevent the binding of RNA polymerase and the initiation of transcription. If a repressor protein is involved in the control mechanism, it must be present at a level high enough to permit pyrimidine-mediated regulation of *pyrC* expression directed by the multicopy plasmid pBHM105. Experiments

TABLE 1. Amino acid composition of *E. coli* dihydroorotase

Amino acid	Subunit composition from:	
	Amino acid analysis[a]	DNA sequence
Ala	33	33
Arg	23	23
Asx	33	(33)
Asp		19
Asn		14
Cys	6	6
Glx	34	(34)
Glu		23
Gln		11
Gly	19	19
His	13	14
Ile	16	16
Leu	33	33
Lys	9	9
Met	9	11
Phe	16	16
Pro	20	21
Ser	16	15
Thr	24	25
Trp	2	2
Tyr	9	9
Val	26	29
Total	341	348
M_r	38,050	38,827

[a] Values from Washabaugh and Collins (34).

are in progress to determine the role of the possible operator sequences and *trans*-acting factors in the regulation of *pyrC*, *pyrD*, and *carAB* expression.

In addition to affecting the level of *pyrC* transcripts, pyrimidine availability appears to affect the site of *pyrC* transcriptional initiation (refer to Fig. 4). The primer extension mapping data show that in uracil-grown cells two minor transcriptional initiation sites (corresponding to bp 428 and 429 in Fig. 3) are used much more frequently than in UMP-grown cells. The major transcriptional initiation site (corresponding to bp 431), however, is the same in cells grown with either pyrimidine source. A possible (presumably secondary) regulatory role for the use of the additional minor transcriptional initiation sites is suggested by the fact that nucleotides 428 to 433 (Fig. 3) are complementary to nucleotides 449 to 454. This latter sequence overlaps one of the possible *pyrC* Shine-Dalgarno sequences (SD_1 in Fig. 3). Transcripts initiated at bp 428 and 429, but not at bp 430 and 431, may be able to form a stable secondary structure that could inhibit translational initiation and further reduce the synthesis of dihydroorotase in cells grown under conditions of pyrimidine excess.

On the basis of their recent studies of *S. typhimurium*, Neuhard et al. (20) suggested that *pyrC* expression is regulated by a novel attenuation control mechanism in which the formation of the secondary structure in the *pyrC* transcript described above would control transcriptional termination (presumably Rho dependent) within the structural gene. Such a mechanism appears to be excluded by the results described in the present study. The quantitative primer extension mapping of *pyrC* transcripts shown in Fig. 4 indicates that regulation occurs at a step before the transcription of the first 11 codons of the *pyrC* structural gene. (Note that the oligonucleotide primer used in this experiment is complementary to codons 3 to 11.) Rho-dependent termination could not occur within this early stage of *pyrC*

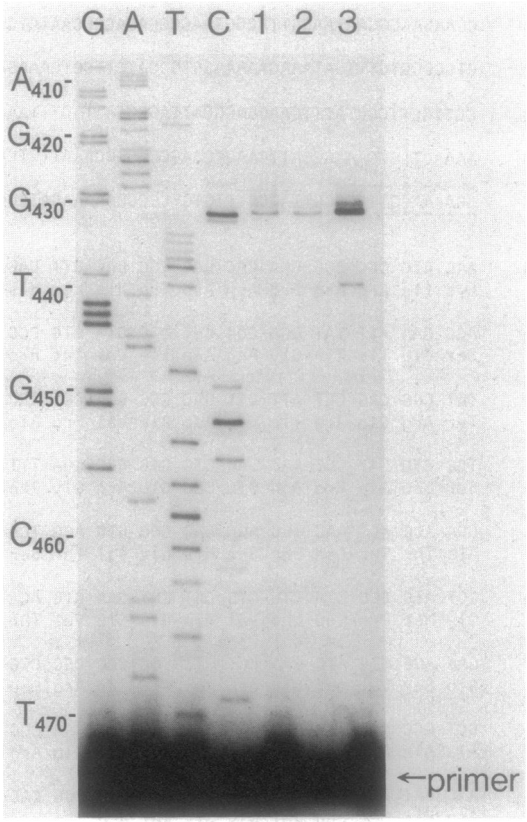

FIG. 4. Primer extension mapping of the 5′ termini of *pyrC* transcripts. An autoradiogram of the 10% polyacrylamide sequencing gel used to analyze the DNA primers extended by reverse transcriptase is shown. The two primer extension reaction mixtures analyzed contained cellular RNA isolated from strain CLT39(pBHM105) grown on uracil (lane 1) or UMP (lanes 2 and 3). The same amount of primer extension reaction mixture (1/10th of the total) was loaded in lanes 1 and 3. The sample loaded in lane 2 was one-fifth of that loaded in lane 3. The dideoxy sequencing ladder (29) was generated by using the same primer that was used in the primer extension reactions. The template used for sequencing was a restriction fragment containing bp 1 to 713 in Fig. 3. Nucleotide positions are numbered in accordance with the numbering of the complementary sequence shown in Fig. 3. There is a C at position 425 that is not detected by dideoxy sequencing.

TABLE 2. Effect of pyrimidine availability on the synthesis of dihydroorotase in strains CLT39(pBHM105) and CLT39

Strain	Pyrimidine source	Total activity (nmol/min per ml)		Sp act (nmol/min per mg)	
CLT39(pBHM105)[a]	Uracil	1,560	(10.9)[b]	2,970	(12.0)[b]
	UMP	8,960		18,800	
CLT39[c]	Uracil	81.9	(11.5)[d]	168	(12.9)[d]
	UMP	938		2,170	

[a] Doubling times were 86 and 130 min on uracil and UMP, respectively.
[b] Numbers in parentheses indicate fold derepression after correction for plasmid copy number. Plasmid copy number was 1.9-fold higher in the uracil-grown cells.
[c] Doubling times were 86 and 150 min on uracil and UMP, respectively.
[d] Numbers in parentheses indicate fold derepression.

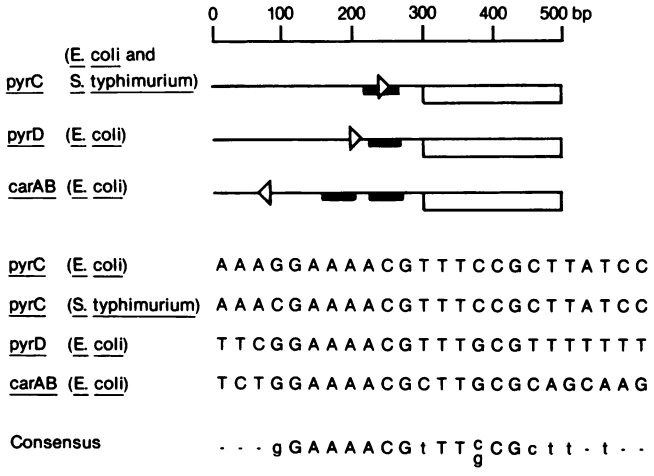

FIG. 5. Location and nucleotide sequence of a highly conserved, operatorlike sequence in the *pyrC*, *pyrD*, and *carAB* promoter-regulatory regions. The open boxes represent the beginning of the coding region for the indicated gene or operon. The filled boxes indicate the positions of the promoters. The *carAB* operon has two promoters; pyrimidine availability regulates transcription from the upstream promoter only (6, 22). The arrowheads indicate the location and orientation of the highly conserved sequences. All sequences are written 5′ to 3′.

transcription because the 5′ end of the transcript is too short to permit Rho binding (33). In addition, the majority of *pyrC* transcripts in *E. coli* (i.e., those initiated at bp 431 in Fig. 3) apparently would not be capable of forming a stable secondary structure between the 5′ end of the transcript and Shine-Dalgarno sequence SD_1.

The sequence of the *E. coli pyrC* ribosome-binding site is unusual in that there are two possible Shine-Dalgarno sequences. It is not clear which of the two sequences would be more likely to function in translational initiation. The upstream sequence, SD_1, is one nucleotide longer (GGAG as compared with GAG for SD_2); however, the spacing between each Shine-Dalgarno sequence and the initiation codon is slightly outside the 5- to 9-nucleotide range typically found in a ribosome-binding site (11). Interestingly, the sequence of the *S. typhimurium pyrC* ribosome-binding site (20) does not contain the sequence AGAG (bp 459 to 462 in Fig. 3), which includes Shine-Dalgarno sequence SD_2. This deletion is the most striking difference between the *pyrC* sequences of the two bacteria. The effect of the different ribosome-binding sites on the expression and regulation of the *pyrC* gene remains to be determined.

During the preparation of this manuscript, Bäckström et al. (3) published a paper that contains the same nucleotide sequence as that determined in this study. There are, however, a number of significant differences in the two papers. In the paper by Bäckström et al., S_1 mapping experiments are described (without presenting any primary data) which apparently show that *pyrC* transcription is initiated primarily at the two base pairs (bp 426 and 427 in Fig. 3) immediately upstream of the initiation sites demonstrated in the present study. This discrepancy is of particular concern because of the potential of the transcripts initiated at the more upstream sites to form a secondary structure which could affect translational initiation and perhaps regulation as described above. We think that the initiation sites reported in the present study are correct because they were detected by two different mapping techniques and because the spacing between the *pyrC* −10 sequence and these sites (but not those reported by Bäckström et al.) is typical of that found in *E. coli* (9). There are also differences in the two papers with respect to the assignment of sequences required for transcription and translation. Bäckström et al. report that in the *pyrC* promoter there is no sequence similar to the consensus −35 region (i.e., TTGACA); however, we think that the sequence GTGCAA located 17 bp upstream of the *pyrC* −10 sequence is a reasonable candidate for the −35 sequence (9). Bäckström et al. indicate only one *pyrC* Shine-Dalgarno sequence (SD_2), but as indicated above we have suggested two candidates. Finally, the paper by Bäckström et al. does not include any experiments demonstrating transcriptional control of *pyrC* expression and does not discuss the highly conserved, operatorlike sequence identified in the present study. However, their paper does propose an assortment of possible *trans*-acting factors and *cis*-acting sequences that could function in the regulation of *pyrC* expression.

ACKNOWLEDGMENTS

We thank John Donahue for his assistance with promoter mapping, Jeffrey Engler for oligonucleotide synthesis, and Nita Redmond for typing the manuscript.

This work was supported by Public Health Service grant GM29466 to C.L.T. from the National Institutes of Health. H.R.W. was a predoctoral fellow supported by training grant GM08111 from the National Institutes of Health, and P.T.C. was a postdoctoral fellow supported by training grant AI07041 from the National Institutes of Health.

LITERATURE CITED

1. **Abdelal, A., and J. L. Ingraham.** 1969. Control of carbamyl phosphate synthesis in *Salmonella typhimurium*. J. Biol. Chem. **244**:4033–4038.
2. **Alper, M. D., and B. N. Ames.** 1978. Transport of antibiotics and metabolite analogs by systems under cyclic AMP control: positive selection of *Salmonella typhimurium cya* and *crp* mutants. J. Bacteriol. **133**:149–157.
3. **Bäckström, D., R.-M. Sjöberg, and L. G. Lundberg.** 1986. Nucleotide sequence of the structural gene for dihydroorotase of *Escherichia coli* K12. Eur. J. Biochem. **160**:77–82.
4. **Beckwith, J. R., A. B. Pardee, R. Austrian, and F. Jacob.** 1962. Coordination of the synthesis of the enzymes in the pyrimidine pathway of *E. coli*. J. Mol. Biol. **5**:618–634.
5. **Bonekamp, F., K. Clemmesen, O. Karlström, and K. F. Jensen.** 1984. Mechanism of UTP-modulated attenuation at the *pyrE* gene of *Escherichia coli*: an example of operon polarity control through the coupling of translation to transcription. EMBO J. **3**:2857–2861.
6. **Bouvier, J., J.-C. Patte, and P. Stragier.** 1984. Multiple regulatory signals in the control region of the *Escherichia coli carAB* operon. Proc. Natl. Acad. Sci. USA **81**:4139–4143.
7. **Clemmesen, K., F. Bonekamp, O. Karlström, and K. F. Jensen.** 1985. Role of translation in the UTP-modulated attenuation at the *pyrBI* operon of *Escherichia coli*. Mol. Gen. Genet. **201**:247–251.
8. **Hagen, F. S., and E. T. Young.** 1978. Effect of RNase III on efficiency of translation of bacteriophage T7 lysozyme mRNA. J. Virol. **26**:793–804.
9. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11**:2237–2255.
10. **Himeno, T., T. Imanaka, and S. Aiba.** 1986. Nucleotide sequence of the penicillinase repressor gene *penI* of *Bacillus licheniformis* and regulation of *penP* and *penI* by the repressor. J. Bacteriol. **168**:1128–1132.
11. **Kozak, M.** 1983. Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. Microbiol. Rev. **47**:1–45.
12. **Larsen, J. N., and K. F. Jensen.** 1985. Nucleotide sequence of

the *pyrD* gene of *Escherichia coli* and characterization of the flavoprotein dihydroorotate dehydrogenase. Eur. J. Biochem. **151**:59–65.

13. **Levin, H. L., and H. K. Schachman.** 1985. Regulation of aspartate transcarbamoylase synthesis in *Escherichia coli*: analysis of deletion mutations in the promoter region of the *pyrBI* operon. Proc. Natl. Acad. Sci. USA **82**:4643–4647.
14. **Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall.** 1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem. **193**:265–275.
15. **Luse, D. S., J. R. Haynes, D. VanLeeuwen, E. A. Schon, M. L. Cleary, S. G. Shapiro, J. B. Lingrel, and R. G. Roeder.** 1981. Transcription of the β-like globin genes and pseudogenes of the goat in a cell-free system. Nucleic Acids Res. **9**:4339–4354.
16. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
17. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
18. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. **Navre, M., and H. K. Schachman.** 1983. Synthesis of aspartate transcarbamoylase in *Escherichia coli*: transcriptional regulation of the *pyrB-pyrI* operon. Proc. Natl. Acad. Sci. USA **80**:1207–1211.
20. **Neuhard, J., R. A. Kelln, and E. Stauning.** 1986. Cloning and structural characterization of the *Salmonella typhimurium pyrC* gene encoding dihydroorotase. Eur. J. Biochem. **157**:335–342.
21. **Piérard, A., N. Glansdorff, D. Gigot, M. Crabeel, P. Halleux, and L. Thiry.** 1976. Repression of *Escherichia coli* carbamoylphosphate synthase: relationships with enzyme synthesis in the arginine and pyrimidine pathways. J. Bacteriol. **127**:291–301.
22. **Piette, J., H. Nyunoya, C. J. Lusty, R. Cunin, G. Weyens, M. Crabeel, D. Charlier, N. Glansdorff, and A. Piérard.** 1984. DNA sequence of the *carA* gene and the control region of *carAB*: tandem promoters, respectively controlled by arginine and the pyrimidines, regulate the synthesis of carbamoyl-phosphate synthetase in *Escherichia coli* K-12. Proc. Natl. Acad. Sci. USA **81**:4134–4138.
23. **Poulsen, P., F. Bonekamp, and K. F. Jensen.** 1984. Structure of the *Escherichia coli pyrE* operon and control of *pyrE* expression by a UTP modulated intercistronic attenuation. EMBO J. **3**:1783–1790.
24. **Prescott, L. M., and M. E. Jones.** 1969. Modified methods for the determination of carbamyl aspartate. Anal. Biochem. **32**:408–419.
25. **Robertson, H. D., E. Dickson, and J. J. Dunn.** 1977. A nucleotide sequence from a ribonuclease III processing site in bacteriophage T7 RNA. Proc. Natl. Acad. Sci. USA **74**:822–826.
26. **Roland, K. L., F. E. Powell, and C. L. Turnbough, Jr.** 1985. Role of translation and attenuation in the control of *pyrBI* operon expression in *Escherichia coli* K-12. J. Bacteriol. **163**:991–999.
27. **Roof, W. D., K. F. Foltermann, and J. R. Wild.** 1982. The organization and regulation of the *pyrBI* operon in *E. coli* includes a rho-independent attenuator sequence. Mol. Gen. Genet. **187**:391–400.
28. **Rosenberg, M., and R. A. Kramer.** 1977. Nucleotide sequence surrounding a ribonuclease III processing site in bacteriophage T7 RNA. Proc. Natl. Acad. Sci. USA **74**:984–988.
29. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
30. **Schwartz, M., and J. Neuhard.** 1975. Control of expression of the *pyr* genes in *Salmonella typhimurium*: effects of variations in uridine and cytidine nucleotide pools. J. Bacteriol. **121**:814–822.
31. **Turnbough, C. L., Jr., K. L. Hicks, and J. P. Donahue.** 1983. Attenuation control of *pyrBI* operon expression in *Escherichia coli* K-12. Proc. Natl. Acad. Sci. USA **80**:368–372.
32. **Turnbough, C. L., Jr., R. J. Neill, R. Landsberg, and B. N. Ames.** 1979. Pseudouridylation of tRNAs and its role in regulation in *Salmonella typhimurium*. J. Biol. Chem. **254**:5111–5119.
33. **von Hippel, P. H., D. G. Bear, W. D. Morgan, and J. A. McSwiggen.** 1984. Protein-nucleic acid interactions in transcription: a molecular analysis. Annu. Rev. Biochem. **53**:389–446.
34. **Washabaugh, M. W., and K. D. Collins.** 1986. Dihydroorotase from *Escherichia coli*. Sulfhydryl group-metal ion interactions. J. Biol. Chem. **261**:5920–5929.
35. **Weaver, R. F., and C. Weissmann.** 1979. Mapping of RNA by a modification of the Berk-Sharp procedure: the 5′ termini of 15 S β-globin mRNA precursor and mature 10 S β-globin mRNA have identical map coordinates. Nucleic Acids Res. **7**:1175–1193.