JOURNAL OF BACTERIOLOGY, May 1994, p. 2513–2516
0021-9193/94/$04.00+0
Copyright © 1994, American Society for Microbiology
Vol. 176, No. 9

Identification of the Shine-Dalgarno Sequence Required for Expression and Translational Control of the *pyrC* Gene in *Escherichia coli* K-12

JUNKE LIU AND CHARLES L. TURNBOUGH, JR.*

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

Received 28 September 1993/Accepted 20 February 1994

Expression of the *pyrC* gene in *Escherichia coli* K-12 is regulated by a translational control mechanism in which CTP (and perhaps GTP) pool sizes determine the selection of alternative transcriptional start sites at the *pyrC* promoter. High CTP levels cause transcription to start primarily at a site that directs the synthesis of untranslatable *pyrC* transcripts. These transcripts form a hairpin at their 5′ ends that blocks ribosome binding to the Shine-Dalgarno (SD) sequence. The *pyrC* ribosome binding site is unusual in that it contains two potential SD sequences, designated SD1 and SD2, which are located 11 and 4 nucleotides upstream of the translational initiation codon, respectively. In this study, we examined the functions of these two SD sequences in translational initiation. Mutations that inactivate either SD1 or SD2 were constructed and incorporated separately into a *pyrC::lacZ* protein fusion. The effects of the mutations on *pyrC::lacZ* expression, regulation, and transcript levels were determined. The results indicate that SD1 is the only functional *pyrC* SD sequence. The SD2 mutation did cause a small reduction in expression, but this effect appeared to be due to a decrease in transcript stability. In addition, we constructed a mutation that introduces a long spacer region between the hairpin at the 5′ end of the *pyrC* transcript and a new *pyrC* SD sequence. As predicted by the model for translational control, this mutation caused constitutive expression of a *pyrC::lacZ* protein fusion.

In *Escherichia coli* and *Salmonella typhimurium*, the *pyrC* gene encodes the enzyme dihydroorotase, which catalyzes the third step in the pyrimidine nucleotide biosynthetic pathway (6). Expression of the *pyrC* gene is negatively regulated by pyrimidine availability primarily through a translational control mechanism based on the selection of alternative transcriptional start sites at the *pyrC* promoter (5, 12, 13, 15). In *E. coli* K-12, transcriptional initiation at the *pyrC* promoter occurs at four adjacent sites (TCCG in the nontemplate DNA strand) located 6 to 9 bases downstream of the −10 region (Fig. 1) (15, 16). The transcripts initiated at these sites are designated U-6, C-7, C-8, and G-9, respectively, reflecting their initiating nucleotide and position. In cells grown under conditions of pyrimidine excess, which result in a high intracellular level of CTP, C-7 transcripts are synthesized predominantly. However, these transcripts are poorly translated because they form a stable stem-loop structure (hairpin) at their 5′ ends that blocks ribosome binding, resulting in low dihydroorotase synthesis. In contrast, in cells limited for pyrimidines, where the CTP level is low and the GTP level is high (10), G-9 transcripts are synthesized primarily. These shorter transcripts are unable to form the inhibitory hairpin at their 5′ ends, which permits efficient translational initiation and the synthesis of high levels of dihydroorotase (15).

The *pyrC* ribosome binding site in *E. coli* K-12 is unusual in that it contains two potential Shine-Dalgarno (SD) sequences (16). In comparison, the equivalent region in *S. typhimurium* has a different sequence and contains only one possible SD sequence (13). The two potential SD sequences in *E. coli* K-12 are designated SD1 (GGAG) and SD2 (GAG) and are located 11 and 4 nucleotides upstream of the *pyrC* translational initiation codon, respectively (Fig. 1) (16). Although the spacing between either of the two potential SD sequences and the *pyrC* initiation codon is outside the typical 5- to 9-nucleotide spacing for *E. coli* ribosome binding sites (14), each is within the functional range for an SD sequence (7). In addition, ribosome binding to either SD sequence should be inhibited by the hairpin at the 5′ ends of the longer *pyrC* transcripts (2). The sequence of the hairpin overlaps SD1 and is only 5 nucleotides upstream of SD2, which is close enough to interfere with ribosome binding at this sequence (3).

In this study, we examined the roles of the two potential *pyrC* SD sequences in translational initiation in *E. coli* K-12. Mutations that individually inactivate SD1 and SD2 were constructed and incorporated separately into a *pyrC::lacZ* protein fusion, which is expressed and regulated in essentially the same way as the *pyrC* gene (15). The effects of the mutations on *pyrC::lacZ* expression, pyrimidine-mediated regulation, and transcript synthesis were measured. The results indicate that SD1 is the only functional *pyrC* SD sequence. In addition, we constructed another mutation that replaces SD2 (GAG) with a 23-bp insertion containing a new *pyrC* SD sequence at the downstream end. This insertion alters *pyrC* transcripts by introducing a 19-nucleotide spacer region between the regulatory hairpin and the new SD sequence. This mutation was incorporated into a *pyrC::lacZ* protein fusion, and its effects on expression and regulation were measured. The results show that the mutation causes constitutive *pyrC::lacZ* expression as predicted by the current model for *pyrC* translational control.

MATERIALS AND METHODS

Bacterial strains and plasmids. *E. coli* K-12 strains CLT42 [F⁻ *car-94* Δ(*argF-lac*)*U169 rpsL150 thiA1 relA1 deoC1 ptsF25 flbB5301 rbsR*] (9) and CLT99 [F⁻ *pyrB482::kan recA56* Δ(*argF-lac*)*U169 rpsL150 thiA1 relA1 deoC1 ptsF25 flbB5301 rbsR*] (15) were used as the parent strains in the construction of λ lysogens and transformants carrying derivatives of plasmid

* Corresponding author. Phone: (205) 934-6289. Fax: (205) 975-5479. Electronic mail address: Chuck_Turnbough@micro.microbio.uab.edu.

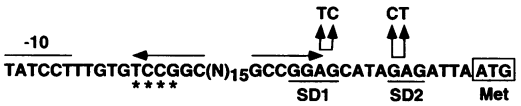

FIG. 1. Nucleotide sequence of the promoter/translational initiation region of the *pyrC* gene. The promoter −10 region is overlined and labeled, and transcriptional start sites are indicated by asterisks. The region of hyphenated dyad symmetry specifying the regulatory hairpin at the 5′ ends of *pyrC* transcripts is indicated by horizontal arrows. The ATG translational initiation codon is enclosed in a box. Potential *pyrC* SD sequences are underlined and labeled SD1 and SD2. Two of the mutations constructed and characterized in this study are shown above the wild-type sequence. Each of these mutations changes a 2-bp region in either SD1 or SD2. Neither mutation causes a major change in the predicted secondary structure at the 5′ end of the *pyrC* transcript (through codon 10).

pMLB1034 (11), respectively. Plasmid pMLB1034, which contains the *lacZ* gene lacking a promoter, a ribosome binding site, and the first eight codons for β-galactosidase, was used to construct *pyrC*::*lacZ* protein fusions. A 727-bp (or 747-bp in the case of the spacer mutation) *Eco*RI-*Bam*HI *pyrC* fragment containing either a wild-type or mutant *pyrC* SD region was inserted between the unique *Eco*RI and *Bam*HI sites of the plasmid. The *Eco*RI-*Bam*HI fragments were derived essentially as previously described (15). In each protein fusion, the first 86 *pyrC* codons are fused in frame to codon 9 of *lacZ*.

DNA preparations and transformations. DNA preparations, restriction digestions, ligations, and transformations were performed as previously described (5).

In vitro oligonucleotide-directed mutagenesis. Oligonucleotide-directed mutagenesis of the *pyrC* SD region carried on recombinant bacteriophage M13mp9amC5 and confirmation of mutations by sequence analysis were done as previously described (17).

Transfer of *pyrC*::*lacZ* protein fusions from plasmids to the *E. coli* chromosome. Wild-type and mutant *pyrC*::*lacZ* protein fusions carried on derivatives of plasmid pMLB1034 were transferred individually to the chromosome of strain CLT42 by using bacteriophage λRZ5 (8). The presence of a single prophage at the λ attachment site on the *E. coli* chromosome was confirmed by Southern hybridization (8).

Media and culture methods. Cells used for enzyme assays and RNA isolations were grown in N^-C^- medium (1) supplemented with 10 mM NH_4Cl, 0.4% glucose, 5 μg of thiamine hydrochloride per ml, 1 mM arginine, either 1 mM uracil or 0.25 mM UMP as the sole pyrimidine source, and 100 μg of ampicillin per ml for plasmid-containing strains. Cultures were grown at 30°C with shaking to an optical density at 650 nm of 0.5 (exponential phase). Culture densities and doubling times were measured as previously described (15).

Enzyme assays. Cell extracts were prepared by sonic oscillation (15). β-Galactosidase (15) and β-lactamase (8) activities were measured and protein determinations were performed (15) as previously described. Enzyme activities presented in the tables are averages of results from at least two experiments with variation of less than ±5% for β-galactosidase and ±10% for β-lactamase.

Isolation of cellular RNA and primer extension mapping. Quantitative isolation of cellular RNA (15) and primer extension mapping of the 5′ ends of *pyrC* transcripts (5) were done as previously described. A 25-nucleotide-long oligodeoxyribonucleotide primer with the sequence 5′-GGATCTTTAATAC CTGGGATGGTGC was used for primer extension mapping. This primer is complementary to a region in the *pyrC* transcript between codons 3 and 11 (16).

TABLE 1. Effects of *pyrC* SD mutations on expression and regulation of *pyrC*::*lacZ* protein fusions carried on multicopy plasmids[a]

Plasmid	*pyrC* SD region	β-Galactosidase activity (nmol/min/mg of protein)[b]		Fold regulation[c]
		Uracil	UMP	
pBHM338	Wild type	12,200	80,500	8.8
pBHM344	SD1⁻	33	276	8.9
pBHM345	SD2⁻	7,390	55,100	10.9

[a] Doubling times for transformants of strain CLT99 carrying the indicated plasmid were 82 ± 2 min on uracil and 119 ± 2 min on UMP.

[b] Not adjusted for plasmid copy number. On the basis of plasmid-encoded β-lactamase activities, relative plasmid copy numbers in cells grown on uracil and UMP were 1.00 and 0.75 for pBHM338, 1.30 and 1.22 for pBHM344, and 1.10 and 0.75 for pBHM345, respectively.

[c] Corrected for plasmid copy number.

RESULTS AND DISCUSSION

Effects of eliminating SD1 or SD2 on the expression and regulation of a *pyrC*::*lacZ* protein fusion. The mutations designed to prevent ribosome binding at SD1 and SD2 are shown in Fig. 1. Each mutation was incorporated separately into a *pyrC*::*lacZ* protein fusion carried on a multicopy plasmid as described in Materials and Methods. Wild-type and mutant protein fusion plasmids were introduced by transformation into the pyrimidine-auxotrophic strain CLT99 (*pyrB482* Δ*lac recA56*). Transformants were grown in minimal medium containing either uracil or UMP as the sole pyrimidine source. Growth on uracil provides a condition of pyrimidine excess, whereas growth on UMP, which is only slowly metabolized by the cells, results in pyrimidine limitation. The effects of each SD mutation on *pyrC*::*lacZ* expression and regulation, as determined by β-galactosidase activities, are shown in Table 1. The SD1 mutation virtually eliminated *pyrC*::*lacZ* expression, reducing β-galactosidase activities to approximately 0.3% of wild-type levels in cells grown on uracil or UMP. In sharp contrast, the SD2 mutation reduced β-galactosidase activities by only about one-third in cells grown on either pyrimidine source. Pyrimidine-mediated regulation of β-galactosidase activities was not significantly affected by either SD mutation. None of the effects of the SD mutations appeared to be associated with differences in plasmid copy numbers (Table 1, footnote *b*). These results indicate that SD1 is the only significant SD sequence for the *pyrC* gene and that the SD2 mutation has an indirect negative effect on translation of *pyrC* transcripts.

Effects of the SD1 and SD2 mutations on steady-state levels of *pyrC*::*lacZ* protein fusion transcripts. The CLT99 transformants carrying the wild-type and mutant protein fusion plasmids also were used to examine the effects of the SD mutations on steady-state levels of *pyrC*::*lacZ* transcripts. Cellular RNA was isolated quantitatively from strains grown on uracil and UMP, and the levels and start sites of *pyrC*::*lacZ* transcripts were determined by primer extension mapping (Fig. 2 and Table 2). The results show a pattern of transcripts for the wild-type fusion that is essentially the same as that observed with wild-type *pyrC* transcripts (15). In each case, only low levels of U-6 and C-7 transcripts are detected in uracil-grown cells, presumably because these untranslated transcripts have similar, short half-lives. In the case of *pyrC* transcripts, previous measurements indicate that the half-lives of U-6 and C-7 transcripts are about 1/10 as long as those of the efficiently translated C-8 and G-9 transcripts (5, 15).

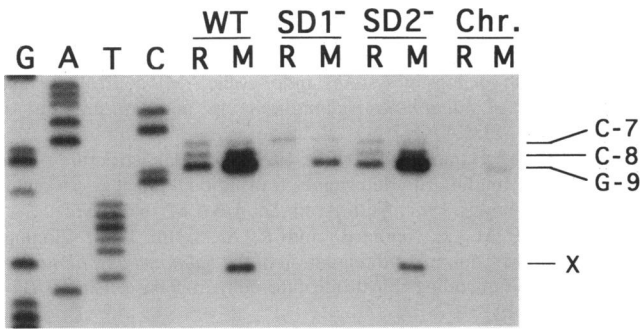

FIG. 2. Steady-state levels of *pyrC*::*lacZ* transcripts containing wild-type and mutant *pyrC* SD regions. Transcripts were quantitatively isolated from transformants of strain CLT99 carrying the *pyrC*::*lacZ* protein fusion plasmids pBHM338 (wild-type SD), pBHM344 (SD1⁻), and pBHM345 (SD2⁻) and from strain CLT99 carrying the vector pMLB1034 to correct for chromosomally encoded *pyrC* transcripts (Chr.). Cells were grown on either uracil (R) or UMP (M). Transcript levels were measured by primer extension mapping as described in Materials and Methods. Shown is an autoradiogram of a sequencing gel used to analyze the primer extension products. The positions of the bands corresponding to *pyrC*::*lacZ* transcripts (or *pyrC* transcripts in the control experiment) are indicated at the right. The bands are labeled according to the initiating nucleotide of the transcript and the position of the transcript start site downstream of the *pyrC* promoter −10 region. Bands corresponding to U-6 transcripts initiated at position T-6 are not visible in this exposure, and their location above C-7 is not marked. The band marked with an X corresponds to a *pyrC* transcript degradation product (15). The *pyrC* dideoxy sequencing ladder used to identify transcripts was generated with the same primer that was used for primer extension.

Compared with the pattern of wild-type *pyrC*::*lacZ* transcripts, the pattern of fusion transcripts with the SD1 mutation was much different (Fig. 2). The major change was that the levels of C-8 and G-9 transcripts were reduced to approximately 10% of wild-type levels in cells grown on uracil or UMP (Table 2). Although sizable, these reductions are much too small to account for the several-hundredfold reduction in *pyrC*::*lacZ* expression caused by the SD1 mutation (Table 1).

TABLE 2. Effects of *pyrC* SD mutations on steady-state levels of transcripts encoded by *pyrC*::*lacZ* protein fusions carried on multicopy plasmids

Plasmid (SD region)	*pyrC*::*lacZ* transcript	Steady-state level[a]	
		Uracil	UMP
pBHM338 (wild type)	U-6	304	273
	C-7	1,470	1,690
	C-8	2,010	7,000
	G-9	5,000	27,700
pBHM344 (SD1⁻)	U-6	201	103
	C-7	1,070	491
	C-8	240	552
	G-9	536	2,650
pBHM345 (SD2⁻)	U-6	170	140
	C-7	937	901
	C-8	1,120	3,790
	G-9	3,180	17,800

[a] Expressed in arbitrary units based on PhosphorImager counts. Background counts contributed by chromosomally encoded wild-type *pyrC* transcripts were subtracted.

TABLE 3. Effects of *pyrC* SD mutations on expression and regulation of chromosomal *pyrC*::*lacZ* protein fusions[a]

Strain	*pyrC* SD region	β-Galactosidase activity (nmol/min/mg of protein)		Fold regulation
		Uracil	UMP	
CLT5087	Wild type	358	4,220	11.8
CLT5093	SD1⁻	1.92	35.6	18.5
CLT5094	SD2⁻	224	2,620	11.7

[a] Doubling times for strains were 63 ± 2 min on uracil and 111 ± 1 min on UMP.

These results indicate that the SD1 mutation prevents translation of C-8 and G-9 transcripts, which leads to a 10-fold-higher rate of degradation of these transcripts.

Unlike the SD1 mutation, the SD2 mutation did not cause significant changes in the pattern of fusion transcripts compared with that of the wild type (Fig. 2). The only changes were quantitative. The levels of mutant U-6, C-7, C-8, and G-9 transcripts were all reduced to approximately 60% of the wild-type level (Table 2). This result indicates that the SD2 mutation causes a small reduction in the half-lives of *pyrC*::*lacZ* transcripts, since transcript synthesis is unlikely to be affected by the mutation. Presumably, the SD2 mutation reduces transcript stability directly and not through decreased translational initiation using SD1, because translatable and untranslatable transcripts are affected equally. The destabilization of C-8 and G-9 transcripts apparently accounts for the one-third reduction in β-galactosidase activities caused by the SD2 mutation described in Table 1.

Confirmation of the effects of the SD1 and SD2 mutations using a chromosomal *pyrC*::*lacZ* protein fusion. To eliminate possible artifacts caused by the use of multicopy plasmids, we also determined the effects of the two SD mutations on the expression of a chromosomal *pyrC*::*lacZ* protein fusion. The wild-type and mutant protein fusions were incorporated into bacteriophage λRZ5, and the resulting phages were used to infect the pyrimidine-auxotrophic strain CLT42 (*car-94* Δ*lac*) as described in Materials and Methods. Lysogens carrying a single prophage at the chromosomal λ attachment site were grown in minimal medium containing either uracil or UMP as the pyrimidine source, and these cells were used to measure β-galactosidase activities. The results were very similar to those described with multicopy plasmids in Table 1. The SD1 mutation reduced β-galactosidase activities to 0.5 and 0.8% of wild-type levels in cells grown on uracil and UMP, respectively (Table 3). In this case, the range of pyrimidine-mediated regulation of residual enzyme synthesis was slightly increased compared with the wild type. The SD2 mutation reduced β-galactosidase activities by 38% in cells grown on either pyrimidine source, thus having no effect on regulation (Table 3). The reduction in β-galactosidase activities caused by the SD2 mutation is essentially identical to the effect that the mutation had on steady-state levels of *pyrC*::*lacZ* protein fusion transcripts (Table 2). These results provide additional evidence that SD1 is the only functional SD sequence for the *pyrC* gene and that the SD2 mutation has only a minor negative effect on gene expression that corresponds to a small decrease in transcript half-lives.

Effects of a 19-nucleotide spacer region between the regulatory hairpin and SD sequence in the *pyrC* transcript on gene expression and regulation. Translational control of *pyrC* expression apparently requires the formation of a hairpin at the

TABLE 4. Effects of the spacer mutation in the *pyrC* SD region on *pyrC*::*lacZ* protein fusion expression and regulation[a]

Strain	*pyrC* SD region	β-Galactosidase activity (nmol/min/mg of protein)		Fold regulation
		Uracil	UMP	
CLT5087	Wild type	358	4,220	11.8
CLT5095	+Spacer	1,200	2,670	2.2

[a] Doubling times for strains were 62 ± 1 min on uracil and 111 ± 2 min on UMP.

5′ end of the *pyrC* transcript that physically prevents a ribosome from binding to the SD sequence. To test this feature of the model, we constructed a mutation that replaces SD2 (GAG) with a 23-bp region with the sequence 5′-ACAACAA CAACAACAAGGAGCAT. Insertion of this sequence creates a new *pyrC* SD sequence (AAGGAG) located 7 nucleotides upstream of the *pyrC* translational initiation codon and introduces a 19-nucleotide spacer region between this new SD sequence and the regulatory hairpin at the 5′ end of the transcript. Given the size of the ribosome binding site (4), the spacer region should move the new SD sequence far enough downstream of the regulatory hairpin to permit unobstructed interaction between this sequence and an initiating ribosome. A *pyrC*::*lacZ* protein fusion containing the spacer mutation was constructed and incorporated into the chromosome of strain CLT42 as described above. Strains carrying the wild-type and mutant fusions were grown in minimal medium containing either uracil or UMP and then assayed for β-galactosidase activities (Table 4). The results show that the spacer mutation causes a large increase in β-galactosidase activity in cells grown on uracil, indicating that the regulatory hairpin is not effectively blocking translation. The β-galactosidase activity in mutant cells grown on UMP is only twofold higher than that measured in uracil-grown cells. This residual regulation is most likely due to a mechanism unrelated to translational control (15). The β-galactosidase activity in the mutant strain grown on UMP is about one-third less than that measured in the wild-type strain. This reduction may indicate that the spacer mutation has a destabilizing effect on fusion transcripts similar to that observed with the SD2 mutation (Table 2). Taken together, the results with the spacer mutation clearly indicate that the regulatory hairpin needs to be close to the *pyrC* SD region to prevent initiation of translation and also to permit pyrimidine-mediated regulation, as predicted by the model for translational control.

ACKNOWLEDGMENTS

We thank Fengxia Qi for helpful suggestions regarding the manuscript.

This work was supported by Public Health Service grant GM29466 from the National Institutes of Health.

REFERENCES

1. **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.
2. **de Smit, M. H., and J. van Duin.** 1990. Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. Proc. Natl. Acad. Sci. USA **87:**7668–7672.
3. **Frick, M. M., J. Neuhard, and R. A. Kelln.** 1990. Cloning, nucleotide sequence and regulation of the *Salmonella typhimurium pyrD* gene encoding dihydroorotate dehydrogenase. Eur. J. Biochem. **194:**573–578.
4. **Gold, L.** 1988. Posttranscriptional regulatory mechanisms in *Escherichia coli*. Annu. Rev. Biochem. **57:**199–233.
5. **Liu, J., and C. L. Turnbough, Jr.** J. Bacteriol., in press.
6. **Neuhard, J., and P. Nygaard.** 1987. Purines and pyrimidines, p. 445–473. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
7. **Ringquist, S., S. Shinedling, D. Barrick, L. Green, J. Binkley, G. D. Stormo, and L. Gold.** 1992. Translational initiation in *Escherichia coli*: sequences within the ribosome-binding site. Mol. Microbiol. **6:**1219–1229.
8. **Roland, K. L., C. Liu, and C. L. Turnbough, Jr.** 1988. Role of the ribosome in suppressing transcriptional termination at the *pyrBI* attenuator of *Escherichia coli* K-12. Proc. Natl. Acad. Sci. USA **85:**7149–7153.
9. **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.
10. **Sadler, W. C., and R. L. Switzer.** 1977. Regulation of *Salmonella* phosphoribosylpyrophosphate synthetase activity in vivo: deductions from pool measurements. J. Biol. Chem. **252:**8504–8511.
11. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
12. **Sørensen, K. I., K. E. Baker, R. A. Kelln, and J. Neuhard.** 1993. Nucleotide pool-sensitive selection of the transcriptional start site in vivo at the *Salmonella typhimurium pyrC* and *pyrD* promoters. J. Bacteriol. **175:**4137–4144.
13. **Sørensen, K. I., and J. Neuhard.** 1991. Dual transcriptional initiation sites from the *pyrC* promoter control expression of the gene in *Salmonella typhimurium*. Mol. Gen. Genet. **225:**249–256.
14. **Stormo, G. D., T. D. Schneider, and L. M. Gold.** 1982. Characterization of translational initiation sites in *E. coli*. Nucleic Acids Res. **10:**2971–2996.
15. **Wilson, H. R., C. D. Archer, J. Liu, and C. L. Turnbough, Jr.** 1992. Translational control of *pyrC* expression mediated by nucleotide-sensitive selection of transcriptional start sites in *Escherichia coli* K-12. J. Bacteriol. **174:**514–524.
16. **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.
17. **Wilson, H. R., and C. L. Turnbough, Jr.** 1990. Role of the purine repressor in the regulation of pyrimidine gene expression in *Escherichia coli* K-12. J. Bacteriol. **172:**3208–3213.