Effects of Transcriptional Start Site Sequence and Position on Nucleotide-Sensitive Selection of Alternative Start Sites at the *pyrC* Promoter in *Escherichia coli*

JUNKE LIU AND CHARLES L. TURNBOUGH, JR.*

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

Received 17 September 1993/Accepted 8 March 1994

In *Escherichia coli*, expression of the *pyrC* gene is regulated primarily by a translational control mechanism based on nucleotide-sensitive selection of transcriptional start sites at the *pyrC* promoter. When intracellular levels of CTP are high, *pyrC* transcripts are initiated predominantly with CTP at a site 7 bases downstream of the Pribnow box. These transcripts form a stable hairpin at their 5' ends that blocks ribosome binding. When the CTP level is low and the GTP level is high, conditions found in pyrimidine-limited cells, transcripts are initiated primarily with GTP at a site 9 bases downstream of the Pribnow box. These shorter transcripts are unable to form a hairpin at their 5' ends and are readily translated. In this study, we examined the effects of nucleotide sequence and position on the selection of transcriptional start sites at the *pyrC* promoter. We characterized promoter mutations that systematically alter the sequence at position 7 or 9 downstream of the Pribnow box or vary the spacing between the Pribnow box and wild-type transcriptional initiation region. The results reveal preferences for particular initiating nucleotides (ATP $\geq$ GTP > UTP >> CTP) and for starting positions downstream of the Pribnow box (7 >> 6 and 8 > 9 > 10). The results indicate that optimal nucleotide-sensitive start site switching at the wild-type *pyrC* promoter is the result of competition between the preferred start site (position 7) that uses the poorest initiating nucleotide (CTP) and a weak start site (position 9) that uses a good initiating nucleotide (GTP). The sequence of the *pyrC* promoter also minimizes the synthesis of untranslatable transcripts and provides for maximum stability of the regulatory transcript hairpin. In addition, the results show that the effects of the mutations on *pyrC* expression and regulation are consistent with the current model for translational control. Possible effects of preferences for initiating nucleotides and start sites on the expression and regulation of other genes are discussed.

In *Escherichia coli* and *Salmonella typhimurium*, six unlinked genes and operons designated *carAB*, *pyrBI*, *pyrC*, *pyrD*, *pyrE*, and *pyrF* encode the six enzymes that catalyze the de novo synthesis of UMP, the precursor of all pyrimidine nucleotides. These genes and operons are subject to negative and noncoordinate regulation mediated by pyrimidine availability (15). Expression of the *pyrC* gene, which encodes the enzyme dihydroorotase, is regulated primarily by a translational control mechanism that is based on nucleotide-sensitive selection of transcriptional start sites (22, 23). Transcriptional initiation at the *pyrC* promoter can occur at four adjacent sites (specifying UCCG) located 6 to 9 bases downstream of the Pribnow box (24). The transcripts initiated at these sites are designated U-6, C-7, C-8, and G-9. According to the model for regulation, when the intracellular level of CTP is high, C-7 transcripts are synthesized predominantly. These transcripts are poorly translated, however, because they form a stable hairpin at their 5' ends that blocks ribosome binding, resulting in low *pyrC* expression. In contrast, when the CTP level is low and the GTP level is high, conditions found in cells limited for pyrimidines, G-9 transcripts are synthesized primarily. These shorter transcripts are unable to form the inhibitory hairpin at their 5' ends and thus are readily translated, resulting in a high level of dihydroorotase synthesis.

In this study, we have examined the effects of nucleotide sequence and position on the selection of transcriptional start sites at the *pyrC* promoter in *E. coli* K-12. Site-directed

mutations were constructed that systematically alter the sequence of the transcriptional initiation region or change the position of the wild-type initiation region relative to the Pribnow box. The effects of these mutations on start site selection were measured by quantitative primer extension mapping of the 5' ends of *pyrC* transcripts isolated from cells grown under conditions of pyrimidine excess and limitation. The half-lives of wild-type and mutant transcripts were determined to correct for different mRNA stabilities, which is essential for an accurate comparison of the rates of *pyrC* transcript synthesis. We also determined the effects of the mutations on *pyrC* expression and pyrimidine-mediated regulation. Our results show that there are preferences for particular nucleoside triphosphates and positions downstream of the Pribnow box that control the selection of transcriptional start sites. Conflicting preferences (i.e., nucleotide versus position) establish competition between start sites at the *pyrC* promoter that are necessary for pyrimidine-mediated regulation. These preferences presumably apply to other promoters and have the potential to affect the level of expression and regulation of many genes.

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 ffbB5301 rbsR*] (19), CLT99 [F⁻ *pyrB482::kan recA56* Δ (*argF-lac*)*U169 rpsL150 thiA1 relA1 deoC1 ptsF25 ffbB5301 rbsR*] (23), and their derivatives were used in this study. Plasmid pDLC126, which contains a *galK::lacZ* protein fusion that has the *galK* ribosome binding site but lacks a promoter (23), was

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

used to generate *pyrC::lacZ* operon fusions. These fusions were constructed by inserting a 152-bp *Bst*NI *pyrC* fragment (overhanging ends filled in), which contained either a wild-type or mutant *pyrC* promoter region, into plasmid pDLC126 at its unique *Sma*I site preceding the *galK::lacZ* protein fusion. The 152-bp *Bst*NI fragments were derived from plasmid pUC19-based *pyrC* subclones or from *pyrC* fragments synthesized by PCR. In each operon fusion, translation begins at the *pyrC* initiation codon and proceeds through the first 5 *pyrC* codons plus 43 codons of vector sequence before terminating upstream of the *galK* ribosome binding site. Plasmid pMLB1034, which contains the *lacZ* gene lacking a promoter, a ribosome binding site, and the first eight codons for β -galactosidase (20), was used to construct *pyrC::lacZ* protein fusions. These fusions were made by inserting a 727-bp *Eco*RI-*Bam*HI *pyrC* fragment, which contained either a wild-type or mutant *pyrC* promoter region, between the unique *Eco*RI and *Bam*HI sites of the plasmid. The 727-bp *Eco*RI-*Bam*HI fragments were derived essentially as previously described (23). In each protein fusion, the first 86 *pyrC* codons are fused in frame to codon 9 of *lacZ*.

DNA manipulations, transformations, and in vitro oligonucleotide-directed mutagenesis. Plasmid DNA was isolated by CsCl-ethidium bromide density gradient centrifugation (19) or by using Qiagen plasmid kits (Qiagen, Inc., Chatsworth, Calif.). Restriction endonucleases and T4 DNA ligase were purchased from New England BioLabs, Inc. (Beverly, Mass.), and used as recommended by the supplier. DNA restriction fragments were purified from agarose gel slices by the Prep-A-Gene procedure (Bio-Rad Laboratories, Richmond, Calif.). Overhanging 5' ends of restriction fragments were filled in with the Klenow fragment of DNA polymerase I (New England BioLabs). Oligonucleotide-directed mutagenesis of the *pyrC* promoter region carried on recombinant bacteriophage M13mp9-amC5 and confirmation of mutations by sequence analysis were done as previously described (25). Transformations were performed by the method of Chung et al. (1). *E. coli* chromosomal DNA was isolated by using a modification of the procedure described by Gay (6). Cells from a 3-ml overnight $2 \times$ YT (13) culture grown at 37°C with shaking were collected by centrifugation and resuspended in 200 μ l of a solution containing 150 mM NaCl, 100 mM disodium EDTA (pH 8.0), 5 mg of lysozyme (Sigma Chemical Co., St. Louis, Mo.) per ml, and 100 ng of RNase A (Sigma) per ml. The suspension was incubated at 37°C for 30 min, and then 400 μ l of a solution containing 100 mM Tris hydrochloride (pH 9.0), 100 mM NaCl, and 1% sodium dodecyl sulfate was added. The sample was incubated at 37°C for 30 min, mixed with 300 μ l of 7.5 M ammonium acetate, and placed at -20°C for 5 min. The sample was extracted once each with 600 μ l of phenol-chloroform (1:1, saturated with 0.1 M Tris hydrochloride [pH 8.0]) and 600 μ l of chloroform-isoamyl alcohol (24:1). DNA was precipitated from the aqueous phase by adding 700 μ l of isopropanol and incubating the sample at 25°C for 20 min. The DNA was pelleted by centrifugation, washed with 1 ml of 70% ethanol (-20°C), and dissolved in 10 mM Tris hydrochloride (pH 8.0)-0.1 mM disodium EDTA.

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 λ RZ5 (18). The presence of a single prophage at the λ attachment site on the *E. coli* chromosome was confirmed by Southern hybridization (18).

Media and culture methods. Cultures used for enzyme assays or RNA isolations were grown at 30°C to an optical density at 650 nm of 0.5 (exponential phase) in glucose-

FIG. 1. Nucleotide sequence of the *pyrC* transcriptional initiation region. The four wild-type transcriptional start sites T-6, C-7, C-8, and G-9 are indicated by asterisks. Each start site is numbered according to its position downstream of the *pyrC* promoter Pribnow box. (Start sites at mutant promoters also are numbered in this way.) The region of hyphenated dyad symmetry specifying the regulatory hairpin at the 5' ends of *pyrC* transcripts is indicated by horizontal arrows. The *pyrC* Shine-Dalgarno (SD) sequence is labeled. This sequence is the only one of two previously identified potential *pyrC* SD sequences that is used for translational initiation (11, 24). The nine mutations constructed and characterized in this study are shown below the wild-type sequence.

minimal salts medium supplemented with 0.015 mM thiamine hydrochloride, 1 mM arginine, either 1 mM uracil or 0.25 mM UMP as the sole pyrimidine source, and 100 μ g of ampicillin per ml for plasmid-containing strains (23). Culture densities and doubling times were measured as previously described (23).

Enzyme assays. Cell extracts were prepared by sonic oscillation (23). β -Galactosidase (23) and β -lactamase (18) activities were measured and protein determinations were performed (23) as previously described. Enzyme activities presented are averages of results from at least two experiments with variation of less than 5%.

Isolation of cellular RNA. Quantitative isolation of cellular RNA and determination of RNA concentrations were done as previously described (23).

Primer extension mapping. The 5' ends of *pyrC* transcripts were mapped quantitatively by using either 45 μ g of RNA from uracil-grown cells or 27 μ g of RNA from UMP-grown cells, which represents the quantity of RNA isolated from the same volume of equal-density culture. The higher yield of RNA from uracil-grown cells apparently is due to the increased synthesis of stable RNA in faster-growing cells. Each cellular RNA sample was mixed with 5 ng of a 20-nucleotide long, 5' 32 P-end-labeled oligodeoxyribonucleotide primer (23). This primer is complementary to a region of the *pyrC* transcript beginning with the last G of the Shine-Dalgarno sequence (Fig. 1). The RNA mixtures were diluted to 15 μ l by the addition of water and 1.5 μ l of $10 \times$ hybridization buffer (200 mM Tris hydrochloride [pH 8.0], 1 M NaCl, 1 mM disodium EDTA). Samples were heated to 100°C for 2 min and incubated at 50°C for 2 h. After the mixtures were cooled to room temperature, 15 μ l of $2 \times$ reverse transcriptase buffer (0.1 M Tris hydrochloride [pH 8.2 at 37°C], 12 mM $MgCl_2$, 2 mM dithiothreitol, 80 mM KCl, 1 mM each dATP, dGTP, dCTP, and dTTP) and 5 U of avian myeloblastosis virus reverse transcriptase (Boehringer Mannheim Biochemicals, Indianapolis, Ind.) were added, and reaction mixtures were incubated at 37°C for 1 h. The reactions were stopped by adding 15 μ l of sequencing dye (80% [vol/vol] deionized formamide, 10 mM NaOH, 1 mM disodium EDTA, 0.1% xylene cyanol, 0.1% bromophenol blue). An equal volume of each sample was analyzed on a 10% polyacrylamide sequencing gel alongside a dideoxy DNA se-

quencing ladder of *pyrC* sequence generated with the same primer used for primer extension mapping. Primer extension products were visualized by autoradiography. Transcripts initiated at each start site were quantitated by scanning gels with a Molecular Dynamics (Sunnyvale, Calif.) PhosphorImager. Under the conditions described, there was a linear relationship between the amount of cellular RNA added to the primer extension reaction mixture and the level of extended fragments synthesized by reverse transcriptase.

Transcript half-life and synthesis rate determinations. The rates of decay of *pyrC* transcripts in rifampin-treated cells were determined as previously described (23). Transcript half-lives were measured by curve fitting with the Micro Math (Salt Lake City, Utah) MINSQ program. Half-lives were essentially the same in cells grown on uracil or UMP. Determined half-lives and steady-state *pyrC* transcript levels were used to calculate the relative rates of transcript synthesis, assuming that the rate of synthesis is equal to the rate of degradation at steady state. The rate of degradation was determined by using the first-order rate equation $N_t = N_0 (0.5)^{t/\tau}$, where N_0 and N_t are the amounts of transcript at steady state and at time t after the addition of rifampin, respectively, and τ is the transcript half-life. A time of 1 s, which is much shorter than any of the transcript half-lives, was used in these calculations so that N_t and N_0 were nearly the same.

RESULTS

Characterization of *pyrC* promoter mutations: an overview. The *pyrC* promoter mutations constructed in this study are shown in Fig. 1. These mutations systematically alter the nucleotide sequence at position 7 or 9 downstream of the Pribnow box or vary the spacing between the Pribnow box and the wild-type transcriptional initiation region. These changes were designed to identify preferences for nucleotides and positions during transcriptional initiation and factors controlling start site switching at the *pyrC* promoter. The effects of the mutations were examined by introducing each mutation into a *pyrC::lacZ* operon fusion carried on a multicopy plasmid. Operon fusions were used in this case because the encoded *pyrC::lacZ* transcripts, which are initiated at the same multiple sites as *pyrC* transcripts, have longer and more similar half-lives than their *pyrC* counterparts (23). These characteristics greatly improve the accuracy of measuring transcript steady-state levels and half-lives. Wild-type and mutant operon fusion plasmids were introduced by transformation into the pyrimidine-auxotrophic strain CLT99 (*pyrB482 Δlac recA56*). Transformants were grown in glucose-minimal salts medium containing either uracil or UMP as the pyrimidine source. Growth on uracil provides a condition of pyrimidine excess, whereas growth on UMP, which is only slowly used by cells under the present growth conditions, results in pyrimidine limitation. RNA was isolated quantitatively from these cells, and the levels and start sites of the *pyrC::lacZ* transcripts were determined by primer extension mapping. The half-lives of all *pyrC::lacZ* transcripts were measured and used, along with steady-state transcript levels, to calculate the relative rates of transcript synthesis as described in Materials and Methods. In selected cases (i.e., for position 7 mutants), operon fusion-encoded β-galactosidase activities were measured and compared with transcript levels. Expression of β-galactosidase activities from operon fusions is not subject to pyrimidine-mediated translational control (23).

As judged by plasmid-encoded β-lactamase activities, the copy numbers of operon fusion plasmids were the same for wild-type and mutant plasmids in cells grown on either uracil

FIG. 2. Steady-state levels of *pyrC::lacZ* transcripts initiated at the wild-type promoter and mutant promoters containing base substitutions at position 7 downstream of the Pribnow box. Transcripts were quantitatively isolated from transformants of strain CLT99 carrying the *pyrC::lacZ* operon fusion plasmids pBHM507 (wild type [WT]), pBHM525 (T-7), pBHM526 (G-7), and pBHM527 (A-7) and from strain CLT99 carrying the vector pMLB1034 (equivalent to pDLC126) 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. An autoradiogram of a sequencing gel used to analyze the primer extension products is shown. 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 Pribnow box. The *pyrC* dideoxy sequencing ladder used to identify transcripts was generated with the same primer that was used for primer extension.

or UMP. However, direct measurement of plasmid DNA indicated that cells grown on uracil contained approximately 1.4-fold higher levels of plasmid than did cells grown on UMP (23). This difference may account for the generally higher levels of operon fusion transcripts isolated from uracil-grown cells. No corrections for copy number were made in the experiments using *pyrC::lacZ* operon fusion plasmids.

We also determined the effects of all promoter mutations on pyrimidine-mediated regulation of *pyrC* expression. Each mutation was introduced into a *pyrC::lacZ* protein fusion that is regulated exactly like the wild-type *pyrC* gene (23). These fusions along with a wild-type version were incorporated into bacteriophage λRZ5, and the resulting phages were used to infect the pyrimidine-auxotrophic strain CLT42 (*car-94 Δlac*). Lysogens carrying a single prophage at the chromosomal λ attachment site were grown in glucose-minimal salts medium containing either uracil or UMP, and these cells were used to measure β-galactosidase activities encoded by the protein fusions.

Effects of base substitutions at position 7 downstream of the *pyrC* Pribnow box. The effects of mutations altering position 7 downstream of the *pyrC* Pribnow box on steady-state levels of *pyrC::lacZ* transcripts encoded by operon fusion plasmids are shown in Fig. 2. These levels and transcript half-lives were measured and used to calculate relative rates of transcript synthesis (Table 1). Comparisons of the relative rates showed interesting similarities and striking differences between transcriptional initiation at the wild-type and mutant promoters. Changing the wild-type C-7 to T still permitted start site switching primarily from position 7 to position G-9 when cells were limited for pyrimidines. However, this switching was not as extensive as observed at the wild-type promoter. In contrast,

TABLE 1. Effects of base substitutions at position 7 downstream of the *pyrC* Pribnow box on transcriptional initiation of *pyrC::lacZ* operon fusions

<i>pyrC</i> promoter	<i>pyrC::lacZ</i> transcript	Half-life (s)	Steady-state level ^a		Relative rate of synthesis	
			Uracil	UMP	Uracil	UMP
C-7 (wild type)	U-6	212	3,160	822	103 (10) ^b	27 (2)
	C-7	225	22,300	6,140	686 (64)	189 (15)
	C-8	131	2,460	3,880	130 (12)	205 (16)
	G-9	139	3,090	16,800	154 (14)	836 (67)
T-7	U-6	132	5,290	2,890	277 (7)	152 (8)
	U-7	114	56,600	14,400	3,430 (87)	873 (48)
	C-8	110	2,570	3,640	160 (4)	229 (13)
	G-9	126	1,100	10,400	61 (2)	570 (31)
G-7	U-6	110	9,610	8,590	604 (13)	470 (22)
	G-7	111	65,200	27,000	4,060 (87)	1,680 (78)
	C-8		ND ^c	ND		
	G-9		ND	ND		
A-7	U-6	92	3,530	2,570	264 (6)	193 (8)
	A-7	98	56,500	32,000	3,980 (94)	2,250 (92)
	C-8		ND	ND		
	G-9		ND	ND		

^a Steady-state levels of transcripts are in arbitrary units based on PhosphorImager counts and represent averages of results from at least two experiments. Transcript levels in independent experiments were normalized by using wild-type values as a reference. Chromosomal background was subtracted.

^b Numbers in parentheses indicate rate of synthesis of the indicated transcript as a percentage of the rate of total transcript synthesis at a particular promoter.

^c ND, not detectable.

changing position 7 to either G or A completely eliminated switching. Each of these mutations caused initiation to occur almost entirely at position 7. All three of the mutations had a large effect on the rate of transcript synthesis. Each resulted in an increase of five- to sixfold in the rate of synthesis of transcripts initiated at position 7 in uracil-grown cells, the predominant start site under these conditions.

The increases in transcript synthesis rates caused by the mutations also could be detected by comparing β -galactosidase activities encoded by the wild-type and mutant *pyrC::lacZ*

TABLE 2. Effects of base substitutions at position 7 downstream of the *pyrC* Pribnow box on expression of *pyrC::lacZ* fusions^a

Plasmid or strain	Type of fusion	Mutation	β -Galactosidase activity (nmol/min/mg of protein)		Fold regulation
			Uracil	UMP	
Plasmids					
pBHM507	Operon	Wild type	2,960	5,080	1.7
pBHM525	Operon	T-7	5,300	5,700	1.1
pBHM526	Operon	G-7	7,200	7,350	1.0
pBHM527	Operon	A-7	5,000	5,940	1.2
Strains					
CLT5087	Protein	Wild type	358	4,220	11.8
CLT5096	Protein	T-7	439	3,680	8.4
CLT5089	Protein	G-7	166	344	2.1
CLT5100	Protein	A-7	137	460	3.4

^a Doubling times for strain CLT99 transformants carrying the indicated pBHM plasmids were 75 min on uracil and 125 min on UMP. Doubling times for strains CLT5087, CLT5096, CLT5089, and CLT5100 were 62 min on uracil and 110 min on UMP.

FIG. 3. Steady-state levels of *pyrC::lacZ* transcripts initiated at the wild-type promoter and mutant promoters containing base substitutions at position 9 downstream of the Pribnow box. Transcripts were isolated from transformants of strain CLT99 carrying the *pyrC::lacZ* operon fusion plasmids pBHM507 (wild type [WT]), pBHM528 (C-9), and pBHM529 (A-9) and the vector pMLB1034 (Chr.). Other details are as described in the legend to Fig. 2.

operon fusions in uracil-grown cells (Table 2). The mutations resulted in 1.7- to 2.4-fold-higher levels of enzyme activity. The increases were not as high as the five- to sixfold increases seen with transcript synthesis rates because β -galactosidase activities reflect steady-state levels of total transcripts. These levels for mutant transcripts were only 1.9- to 2.4-fold higher than for wild-type transcripts because the mutant transcripts have shorter half-lives (Table 1). A large difference between β -galactosidase activities encoded by the wild-type and mutant *pyrC::lacZ* operon fusions was not observed in UMP-grown cells (Table 2). This result was due to a nearly twofold increase in β -galactosidase activity encoded by the wild-type operon fusion, which presumably is related to transcriptional start site switching from position C-7 to G-9.

The effects of the base substitutions at position 7 on pyrimidine-mediated regulation of *pyrC::lacZ* protein fusions also are shown in Table 2. Changing the wild-type C to T had relatively minor effects on the levels of β -galactosidase activity in cells grown on uracil or UMP, resulting in only a slight reduction in regulation. Regulation occurred over an 8.4-fold range, similar to the 11.8-fold range observed for the wild-type fusion. This result provided additional evidence that start site switching occurred with UTP as the initiating nucleotide at position 7. This result also suggested that a U-7·G-31 base pair could substitute, at least to some extent, for the wild-type C-7·G-31 base pair in the regulatory transcript hairpin. In contrast, the change from C to G or A at position 7 reduced the range of regulation to a low level, most of which appears to be independent of translational control (23). This effect on regulation is a consequence of similar low levels of β -galactosidase activity in the G-7 and A-7 mutant strains grown on uracil or UMP. These low levels are comparable to the β -galactosidase activity in uracil-grown wild-type cells. It appears that G-7 and A-7 mutant transcripts, which are the predominant *pyrC* transcripts in the corresponding mutant strains grown on either pyrimidine source, are not efficiently translated because they form a 4-bp hairpin (including positions 8 to 11 and 27 to 30 in Fig. 1) that inhibits ribosome binding. It should be noted, however, that this inhibition is not necessarily as efficient as that with the 5-bp regulatory hairpin in wild-type C-7 transcripts, because the G-7 and A-7 mutations also eliminate the synthesis of readily translatable G-9 transcripts (Table 1). It is

TABLE 3. Effects of base substitutions at position 9 downstream of the *pyrC* Pribnow box on transcriptional initiation of *pyrC::lacZ* operon fusions

<i>pyrC</i> promoter	<i>pyrC::lacZ</i> transcript	Half-life (s)	Steady-state level ^a		Relative rate of synthesis	
			Uracil	UMP	Uracil	UMP
G-9 (wild type)	U-6	212	3,160	822	103 (10) ^b	27 (2)
	C-7	225	22,300	6,140	686 (64)	189 (15)
	C-8	131	2,460	3,880	130 (12)	205 (16)
	G-9	139	3,090	16,800	154 (14)	836 (67)
C-9	U-6	186	3,610	1,850	134 (16)	69 (14)
	C-7	151	14,400	8,590	659 (81)	393 (77)
	C-8	160	595	764	26 (3)	33 (6)
	C-9	160	ND ^c	320		14 (3)
A-9	U-6	96	1,450	407	104 (8)	29 (4)
	C-7	136	12,800	2,510	650 (49)	128 (15)
	C-8	123	4,480	2,360	252 (19)	133 (16)
	A-9	102	4,760	8,110	322 (24)	549 (65)

^a Except for wild-type values, which are averages of three determinations, steady-state levels of transcripts are PhosphorImager counts from a typical experiment. Transcript levels were normalized to those in Table 1, using wild-type values as a reference. Other details are as described for Table 1.

^b See Table 1, footnote b.

^c ND, not detectable.

translation of low levels of G-9 transcripts, and not translation of the predominant C-7 transcripts, that apparently accounts for most β -galactosidase activity in the wild-type strain grown on uracil.

Effects of base substitutions at position 9 downstream of the *pyrC* Pribnow box. The effects of mutations altering position 9 on the synthesis of *pyrC::lacZ* transcripts encoded by operon fusion plasmids are shown in Fig. 3 and Table 3. Changing the wild-type G-9 to C caused transcriptional initiation to occur predominantly at position C-7 in cells grown on uracil or UMP. Start site switching was virtually eliminated. In contrast, changing position 9 to A resulted in a pattern of transcript synthesis and start site switching similar to that seen with the wild-type promoter. The only apparent difference was a slight preference for transcriptional initiation at A-9 compared with wild-type G-9 in cells grown on uracil.

The effects of the position 9 mutations on expression of *pyrC::lacZ* protein fusions are shown in Table 4. β -Galactosidase activities were much (6- and 12-fold) higher than the wild-type level in uracil-grown cells and only slightly increased in UMP-grown cells, resulting in minimal regulation. These results were expected because each mutation disrupts a central base pair in the regulatory transcript hairpin, which should permit efficient translation of the mutant transcripts initiated at all start sites.

TABLE 4. Effects of base substitutions at position 9 downstream of the *pyrC* Pribnow box on expression of *pyrC::lacZ* protein fusions^a

Strain	Mutation	β -Galactosidase activity (nmol/min/mg of protein)		Fold regulation
		Uracil	UMP	
CLT5087	None (wild type)	358	4,220	11.8
CLT5097	C-9	2,130	5,310	2.5
CLT5161	A-9	4,360	5,100	1.2

^a Doubling times for all strains were 62 min on uracil and 110 min on UMP.

FIG. 4. Steady-state levels of *pyrC::lacZ* transcripts initiated at the wild-type promoter and mutant promoters containing deletion and insertion mutations between the *pyrC* Pribnow box and transcriptional initiation region. Transcripts were isolated from transformants of strain CLT99 carrying the *pyrC::lacZ* operon fusion plasmids pBHM507 (wild type [WT]), pBHM508 (ΔT), pBHM509 (ΔTT), pBHM510 (ΔTTG), and pBHM451 (+T). Other details are as described in the legend to Fig. 2, except that the bands are labeled only by the initiating nucleotide of the transcript. Asterisks indicate transcripts initiated at wild-type or corresponding start sites.

Effects of mutations that alter the spacing between the *pyrC* Pribnow box and wild-type transcriptional initiation region. The effects of altering the spacing between the *pyrC* Pribnow box and wild-type transcriptional initiation region on the synthesis of *pyrC::lacZ* transcripts encoded by operon fusion plasmids are shown in Fig. 4 and Table 5. Changes in the selection of start sites are summarized in Fig. 5. A decrease in

TABLE 5. Effects of the deletion and insertion mutations downstream of the *pyrC* Pribnow box on transcriptional initiation of *pyrC::lacZ* operon fusions^a

<i>pyrC</i> promoter	<i>pyrC::lacZ</i> transcript ^b	Half-life (s)	Steady-state level		Relative rate of synthesis	
			Uracil	UMP	Uracil	UMP
Wild type	U-6	212	3,160	822	103 (10)	27 (2)
	C-7	225	22,300	6,140	686 (64)	189 (15)
	C-8	131	2,460	3,880	130 (12)	205 (16)
	G-9	139	3,090	16,800	154 (14)	836 (67)
ΔT	C-7	131	9,370	10,400	494 (32)	549 (20)
	G-8	139	19,300	41,900	960 (63)	2,080 (75)
	G-9	134	1,620	2,670	84 (5)	138 (5)
ΔTT	C-6	131	4,220	5,240	223 (8)	277 (13)
	G-7	139	42,700	31,300	2,120 (73)	1,560 (75)
	G-8	134	10,700	4,610	552 (19)	238 (12)
ΔTTG	G-6	139	4,440	5,210	221 (18)	259 (22)
	G-7	134	18,900	17,600	975 (82)	908 (78)
+T	G-6	224	2,470	2,130	76 (38)	66 (37)
	U-7	212	2,910	1,220	95 (47)	40 (23)
	C-8	225	985	1,050	30 (15)	32 (18)
	C-9	131	ND ^c	ND		
	G-10	139	ND	780		39 (22)

^a Details are as described in Table 3, footnote a, and Table 1, footnote b.

^b Note that the same transcript (e.g., wild-type G-9 and G-8 with the ΔT promoter) is labeled with a different number for each *pyrC* promoter because the distance between the same start site and the Pribnow box varies.

^c ND, not detectable.

	Pribnow box
WT	TATCCTTTGTGTCCGGC *****
ΔT	TATCCTTGTGTCCGGC **
ΔTT	TATCCTGTGTCCGGC ***
ΔTTG	TATCCTTGTCCGGC **
+T	TATCCTTTTGTGTCCGGC *****

FIG. 5. Summary of the effects of the deletion and insertion mutations on the selection of transcriptional start sites shown in Fig. 4. Start sites are indicated by asterisks.

spacing of 1 bp, caused by the ΔT mutation, resulted in a pattern of transcriptional initiation drastically different from the wild-type pattern. The large majority of transcripts were initiated at position G-8 (equivalent to wild-type G-9), with the remainder initiated primarily at C-7 (equivalent to wild-type C-8) in cells grown on uracil or UMP. There was a slightly higher percentage of transcripts initiated at G-8 in UMP-grown cells, which apparently permitted a higher level of transcript synthesis. A decrease in spacing of 2 bp, caused by the ΔTT mutation, had a similar effect on start site selection. Three-quarters of the transcripts were initiated at position G-7 (equivalent to wild-type G-9), with the remainder initiated at positions G-8 and C-6 (equivalent to wild-type G-10 and C-8, respectively) in cells grown on uracil or UMP. The effect was somewhat different when the spacing was reduced by 3 bp. With the ΔTTG mutation, 80% of the transcripts were initiated at position G-7 (equivalent to wild-type G-10), with the remainder initiated at position G-6 (equivalent to wild-type G-9) in cells grown on uracil or UMP. In sharp contrast, an increase in spacing of 1 bp caused by the +T mutation resulted in a unique, pyrimidine-sensitive pattern. In cells grown on uracil, nearly half of the transcripts were initiated at position T-7 (equivalent to wild-type T-6). A slightly lower level was initiated at position G-6 (equivalent to wild-type G-5), with the remainder initiated at position C-8 (equivalent to wild-type C-7). In cells grown on UMP, a similar (37%) level of the transcripts was initiated at position G-6, with the remainder initiated at roughly the same (20%) level at positions T-7, C-8, and G-10 (the latter equivalent to wild-type G-9). The +T mutation also caused a significant decrease in the level of total

transcripts compared with the wild type. The reason for this decrease is not clear.

The effects of the three deletion mutations on pyrimidine-mediated regulation of *pyrC::lacZ* protein fusions were similar (Table 6). Regulation was reduced to a nearly basal level as a result of high levels of β-galactosidase activity in cells grown on uracil (7- to 15-fold higher than the wild-type level) and only approximately 2-fold higher levels in cells grown on UMP. These results were as expected because each mutation caused transcriptional initiation to occur predominantly at sites that specify transcripts unable to form the regulatory hairpin in cells grown on either pyrimidine source. On the other hand, the +T mutation caused large reductions in protein fusion-encoded β-galactosidase activities in cells grown on uracil (166-fold) and UMP (37-fold), which resulted in a nearly 5-fold increase in the range of regulation compared with the wild type. The large reductions in enzyme activities apparently are due to generally reduced rates of transcriptional initiation, which occurs primarily at sites directing the synthesis of transcripts capable of forming very stable hairpins at their 5' ends. These hairpins should block translation. The increase in regulation is caused by the virtual elimination of translatable transcripts in cells grown on uracil coupled with a low but presumably 52-fold-higher level of translatable (i.e., G-10) transcripts in cells grown on UMP.

DISCUSSION

The key element in translational control of *pyrC* expression is a promoter at which the position of transcriptional initiation can be switched among multiple start sites (primarily between C-7 and G-9) in response to CTP and GTP levels. To identify features of the *pyrC* promoter that permit this start site switching and that possibly influence the level of *pyrC* transcript synthesis, we compared transcriptional initiation at wild-type and mutant promoters in cells grown under conditions of pyrimidine excess or limitation. The results show preferences for particular initiating nucleotides and start site positions which play major roles in *pyrC* expression and regulation.

The experiments in which position 7 downstream of the *pyrC* Pribnow box was changed from C to either T, G, or A were designed primarily to determine if there were advantages associated with using CTP to initiate the predominant C-7 transcript at the wild-type promoter in cells grown on uracil. This transcript contains the regulatory hairpin at its 5' end which apparently prevents translation. The results indicate that there are in fact several important advantages. First, initiation with CTP minimizes the synthesis of untranslatable *pyrC* transcripts because it is the least efficient initiating nucleotide. As measured by relative rates of transcript synthesis, initiating with CTP is approximately one-sixth as efficient as initiating with UTP, GTP, or ATP at position 7 in cells grown on uracil (Table 1). It should be noted that nucleoside triphosphate pools are essentially the same in pyrimidine auxotrophs grown on uracil, the condition of pyrimidine excess used in this study, and in pyrimidine prototrophs grown on minimal medium (15, 17). Reduced transcriptional initiation with CTP also can be detected by comparing the levels of β-galactosidase activities encoded by wild-type and mutant *pyrC::lacZ* operon fusions in cells grown on uracil (Table 2). Second, initiation with CTP permits the maximum level of start site switching when cells are limited for pyrimidines (Fig. 2 and Table 1). Start site switching does occur at the mutant T-7 promoter, but not as extensively because of a relatively high level of initiation with UTP at position 7 in cells grown on UMP. No switching is

TABLE 6. Effects of deletion and insertion mutations between the *pyrC* Pribnow box and transcriptional initiation region on expression of *pyrC::lacZ* protein fusions^a

Strain	Mutation	β-Galactosidase activity (nmol/min/mg of protein)		Fold regulation
		Uracil	UMP	
CLT5087	None (wild type)	358	4,220	11.8
CLT5088	ΔT	2,640	6,680	2.5
CLT5098	ΔTT	4,370	8,370	1.9
CLT5162	ΔTTG	5,200	8,590	1.7
CLT5099	+T	2.16	113	52.3

^a Doubling times for all strains were 62 min on uracil and 110 min on UMP, except for strain CLT5098, which grew with a doubling time of 122 min on UMP.

observed with the G-7 and A-7 promoters. For these promoters, initiation occurs predominantly at position 7 in cells grown on either pyrimidine source, with a slightly higher percentage at position A-7 than at position G-7. Third, initiation with CTP, as opposed to UTP, would permit the formation of a more stable regulatory hairpin as a result of the contribution of a strong C·G base pair. Presumably, a stronger hairpin blocks translation more effectively, thereby permitting a wider range of regulation as observed in Table 2.

The experiments changing position 9 downstream of the Pribnow box from G to either C or A provide additional information about the requirements for switching and efficient transcriptional initiation. The data show that at the C-9 promoter, transcriptional initiation with CTP at position 9 is essentially eliminated in cells grown on either pyrimidine source (Fig. 3 and Table 3). Nearly all transcriptional initiation at this mutant promoter occurs with CTP at position 7, indicating a strong preference for initiation at this position over position 9 (and position 8). These results show that switching requires the use of a strong initiating nucleotide at a weak start site, position 9 in this case. In contrast, transcription and switching at the A-9 promoter is very similar to that observed with the wild-type promoter, with only a slight preference for initiation at position A-9 compared with G-9 in cells grown on uracil. These results suggest that the only important difference between a *pyrC* promoter with either G or A at position 9 is that the G-9 (wild-type) promoter permits the synthesis of C-7 transcripts capable of forming a more stable regulatory hairpin as a result of the contribution of a strong G-9·C-29 base pair.

The experiments in which we varied the spacing between the Pribnow box and wild-type transcriptional initiation region were designed primarily to identify preferred start site positions. Using wild-type and mutant promoters, we compared transcript synthesis initiated with CTP at adjacent sites corresponding to wild-type C-7 and C-8 and with GTP at adjacent sites corresponding to wild-type G-9 and G-10 (Fig. 4 and Table 5). The data with the wild-type and +T promoters show that when transcription initiates with CTP, position 7 is strongly preferred over position 8 and that position 8 is preferred over position 9. The data with the ΔT , ΔTT , and ΔTTG promoters show that when transcription initiates with GTP, position 7 is strongly preferred over positions 6 and 8 and that position 8 is strongly preferred over position 9. The pattern of transcription at the ΔT promoter also provides additional evidence that GTP is a much better initiating nucleotide than CTP. In this case, most transcripts are initiated with GTP at position 8, with a much smaller fraction initiated with CTP at the highly preferred position 7. Another major effect of the deletion mutations is that essentially all switching is eliminated as a result of the existence of a single predominant start site using GTP. In contrast, switching still occurs at the +T promoter, primarily between positions U-7 and G-10. Also of interest is that transcriptional initiation at the +T promoter, which uses UTP at position 7, is widely distributed compared with initiation at other mutant promoters, which use ATP or GTP as the initiating nucleotide at position 7.

The characterization of transcriptional initiation at wild-type and mutant *pyrC* promoters described above permits the following assignments of preferences for initiating nucleotides and start site positions: ATP $\geq$ GTP > UTP >> CTP and position 7 >> 6 and 8 > 9 > 10. These preferences provide the basis for nucleotide-sensitive start site switching at the wild-type *pyrC* promoter. The poorest initiating nucleotide (CTP) is used to start transcripts at the most preferred start site (position 7), and a very good initiating nucleotide (GTP) is

used to start transcripts at a weak start site (position 9). These factors establish competition between initiation at positions C-7 and G-9 that can be influenced over a wide range by changes in intracellular levels of CTP and GTP that occur in cells grown under conditions of pyrimidine excess and limitation. These same preferences restrict transcriptional initiation at positions T-6 and C-8, which requires a combination of weaker start sites and poorer initiating nucleotides. Recent studies measuring steady-state levels of *pyrC::lacZ* protein fusion transcripts initiated at wild-type and several mutant *pyrC* promoters in *S. typhimurium* indicate the same preferences for initiating nucleotides as listed above and the importance of start site position in nucleotide-sensitive switching (21, 22).

In this study, we also examined the effects of the promoter mutations on expression and regulation of *pyrC::lacZ* protein fusions (Tables 2, 4, and 6). The data show that when mutations cause the synthesis of transcripts capable of forming a regulatory hairpin containing a stem of at least 4 bp (including C-8, G-9, G-10, and C-11), expression is low. When mutations cause the synthesis of shorter or disrupted stems, expression is high. When mutations prevent switching between these two classes of transcripts, regulation is reduced to a basal level. These results provide strong support for the current model of translational control of *pyrC* expression in which start site switching controls the synthesis of multiple transcripts with different potentials for translation.

We expect that the preferences for selection of initiating nucleotides and start site positions exhibited at the *pyrC* promoter will apply at other promoters. Examination of approximately 300 *E. coli* promoters shows frequencies for selecting initiating nucleotides (A [47%], G [28%], T [15%], and C [10%]) (10) and start site positions (7 [40%], 8 [24%], 6 [11%], 9 [10%], other sites [$\leq 5\%$]) (7) that reflect the preferences identified with the *pyrC* promoter. These results suggest that most transcripts start with efficient initiating nucleotides to maximize transcript synthesis. They also raise the possibility that particular initiating nucleotides are evolutionarily selected to control transcript levels, which may be particularly important for low-abundance transcripts. Perhaps more importantly, preferences for initiating nucleotides and start sites may be used for nucleotide-sensitive regulation of a large number of genes. In *E. coli* (12) and *S. typhimurium* (5), the *pyrD* promoter includes C-6, C-7, C-8, and G-9 start sites. Transcriptional initiation appears to occur primarily at positions C-6 and C-7 under conditions of pyrimidine excess and at position G-9 under conditions of pyrimidine limitation, resulting in translational control of *pyrD* expression by a mechanism equivalent to that described for *pyrC* (5, 21). Mutations in the *pyrD* transcriptional initiation region also alter start site switching in a manner similar to that observed with *pyrC* promoter mutations (21). Inspection of published bacterial promoter sequences reveals many other transcriptional initiation regions at which nucleotide-sensitive start site switching is predicted. Such switching can produce transcripts with minor differences in sequence at their 5' ends, which produce major differences in the ability of the transcripts to express the specified gene products. Differences at the 5' ends of transcripts can affect transcript stability (4, 16) or control the formation of secondary structures that inhibit translational initiation as seen with *pyrC* (2, 3). Switching also can result in the selection of start sites with different potentials for abortive initiation (14) and transcriptional slippage during initiation (8). Both of these processes can greatly influence the frequency of productive initiation at a particular promoter (9, 26). We anticipate that the expression of many genes can be regulated by intracellular

nucleotide levels through an assortment of mechanisms based on conflicting preferences for initiating nucleotides and transcriptional start sites.

ACKNOWLEDGMENTS

We thank Tom Elliott for helpful suggestions regarding the manuscript.

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

REFERENCES

1. Chung, C. T., S. L. Niemela, and R. H. Miller. 1989. One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc. Natl. Acad. Sci. USA* **86**:2172–2175.
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. de Smit, M. H., and J. van Duin. 1990. Control of prokaryotic translational initiation by mRNA secondary structure. *Prog. Nucleic Acid Res. Mol. Biol.* **38**:1–35.
4. Emory, S. A., P. Bouvet, and J. G. Belasco. 1992. A 5'-terminal stem-loop structure can stabilize mRNA in *Escherichia coli*. *Genes Dev.* **6**:135–148.
5. 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.
6. Gay, N. J. 1984. Construction and characterization of an *Escherichia coli* strain with a *uncI* mutation. *J. Bacteriol.* **158**:820–825.
7. Harley, C. B., and R. P. Reynolds. 1987. Analysis of *E. coli* promoter sequences. *Nucleic Acids Res.* **15**:2343–2361.
8. Jacques, J.-P., and M. M. Susskind. 1990. Pseudo-templated transcription by *Escherichia coli* RNA polymerase at a mutant promoter. *Genes Dev.* **4**:1801–1810.
9. Kammerer, W., U. Deuschle, R. Gentz, and H. Bujard. 1986. Functional dissection of *Escherichia coli* promoters: information in the transcribed region is involved in late steps of the overall process. *EMBO J.* **5**:2995–3000.
10. Lisser, S., and H. Margalit. 1993. Compilation of *E. coli* mRNA promoter sequences. *Nucleic Acids Res.* **21**:1507–1516.
11. Liu, J., and C. L. Turnbough, Jr. 1994. Identification of the Shine-Dalgarno sequence required for expression and translational control of the *pyrC* gene in *Escherichia coli* K-12. *J. Bacteriol.* **176**:2513–2516.
12. Liu, J., and C. L. Turnbough, Jr. Unpublished data.
13. Miller, J. H. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
14. Munson, L. M., and W. S. Reznikoff. 1981. Abortive initiation and long ribonucleic acid synthesis. *Biochemistry* **20**:2081–2085.
15. 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.
16. Petersen, C. 1992. Control of functional mRNA stability in bacteria: multiple mechanisms of nucleolytic and non-nucleolytic inactivation. *Mol. Microbiol.* **6**:277–282.
17. Poulsen, P., and K. F. Jensen. 1987. Effect of UTP and GTP pools on attenuation at the *pyrE* gene of *Escherichia coli*. *Mol. Gen. Genet.* **208**:152–158.
18. 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.
19. 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.
20. Silhavy, T. J., M. L. Berman, and L. W. Enquist. 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. 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.
22. 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.
23. 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.
24. 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.
25. 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.
26. Xiao, F. X., and W. S. Reznikoff. 1993. Transcriptional slippage during the transcription initiation process at a mutant *lac* promoter in vivo. *J. Mol. Biol.* **231**:569–580.