JOURNAL OF BACTERIOLOGY, Jan. 2001, p. 221–228 Vol. 183, No. 1
0021-9193/01/$04.00+0 DOI: 10.1128/JB.183.1.221–228.2001

A Long T · A Tract in the *upp* Initially Transcribed Region Is Required for Regulation of *upp* Expression by UTP-Dependent Reiterative Transcription in *Escherichia coli*

YULIN CHENG, SARA M. DYLLA, AND CHARLES L. TURNBOUGH, JR.*

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

Received 8 May 2000/Accepted 11 October 2000

In *Escherichia coli*, pyrimidine-mediated regulation of *upp* expression occurs by UTP-sensitive selection of alternative transcriptional start sites, which produces transcripts that differ in the ability to be elongated. The *upp* initially transcribed region contains the sequence GATTTTTTTTG (nontemplate strand). Initiation can occur at either the first or the second base in this sequence (designated G6 and A7, with numbering from the promoter −10 region). High intracellular UTP levels favor initiation at position A7; however, the resulting transcripts are subject to reiterative transcription (i.e., repetitive UMP addition) within the 8-bp T · A tract in the initially transcribed region and are aborted. In contrast, low intracellular UTP levels favor initiation at position G6, which results in transcripts that can, in part, avoid reiterative transcription and be elongated normally. In this study, we examined the regulatory requirement for the long T · A tract in the *upp* initially transcribed region. We constructed *upp* promoter mutations that shorten the T · A tract to 7, 6, 5, 4, 3, or 2 bp and examined the effects of these mutations on *upp* expression and regulation. The results indicate that pyrimidine-mediated regulation is gradually reduced as the T · A tract is shortened from 7 to 3 bp; at which point regulation ceases. This reduction in regulation is due to large-percentage increases in *upp* expression in cells grown under conditions of pyrimidine excess. Quantitation of cellular transcripts and in vitro transcription studies indicate that the observed effects of a shortened T · A tract on *upp* expression and regulation are due to increases in the fraction of both G6- and A7-initiated transcripts that avoid reiterative transcription and are elongated normally.

Reiterative transcription (also referred to as pseudotemplated transcription, transcriptional slippage, and RNA polymerase stuttering) is a reaction catalyzed by a number of bacterial, phage, viral, and eukaryotic RNA polymerases (15, 16, 19). In this reaction, nucleotides are repetitively added to the 3′ end of a nascent transcript due to slippage between the transcript and the DNA or RNA template. Typically, slippage occurs between a homopolymeric sequence in the transcript and at least three complementary bases in the template (37). In most cases, the mechanism involves one or more rounds of a one-base upstream shift of the transcript so that the same nucleotide in the template specifies multiple residues in the transcript (10, 13). Recent studies indicate that reiterative transcription plays an important role in the expression and regulation of a number of bacterial and viral genes by a variety of mechanisms (13, 18, 20, 29).

One of these genes is the *upp* gene of *Escherichia coli*. The *upp* gene encodes the pyrimidine salvage enzyme uracil phosphoribosyltransferase, which catalyzes the formation of UMP from uracil and phosphoribosylpyrophosphate (4). The *upp* gene appears to be the first gene of an operon that also contains the *uraA* gene, encoding uracil permease, and perhaps a third, uncharacterized gene designated *b2496* (3, 7). A sequence resembling a strong intrinsic transcriptional terminator is located between the *upp* and *uraA* genes (3). The function of this terminator-like sequence in the expression and regulation of downstream genes is unknown.

Expression of the *upp* gene, and presumably cotranscribed genes, is negatively regulated over a sixfold range by pyrimidine availability (4, 30, 35). This regulation occurs mainly by UTP-sensitive selection of alternative transcriptional start sites, which produces transcripts that differ in the ability to be productively elongated (35). The *upp* initially transcribed region contains the sequence GATTTTTTTTG (nontemplate strand) (Fig. 1). Transcription is initiated primarily at the first two bases in this sequence, designated G6 and A7 (numbering from the promoter −10 region). High intracellular levels of UTP, due to ample pyrimidine availability or synthesis, favor initiation at position A7. However, the resulting transcripts are subject to reiterative transcription (i.e., repetitive UMP addition) within the 8-bp T · A tract in the initially transcribed region. These transcripts, with the general sequence $AUUUU_n$ (where n equals 1 to >50), are not extended to include downstream sequences and are eventually aborted. In contrast, low intracellular levels of UTP, caused by pyrimidine limitation, strongly favor initiation at position G6. This start site switch appears to be caused by inhibition of initiation at position A7, which relies on a high concentration of UTP to form the critical first internucleotide bond of the transcript (23, 26). Transcripts initiated at position G6 can, at least in part, avoid reiterative transcription and be elongated normally. This effect is apparently due to the formation of a relatively stable hybrid between the 5′ end of the G6 transcript and the DNA template.

Several other *E. coli* operons encoding pyrimidine metabolic

* Corresponding author. Mailing address: UAB Department of Microbiology, BBRB 409, 1530 3rd Ave. S., Birmingham AL 35294-2170. Phone: (205) 934-6289. Fax: (205) 975-5479. E-mail: ChuckT@uab.edu.

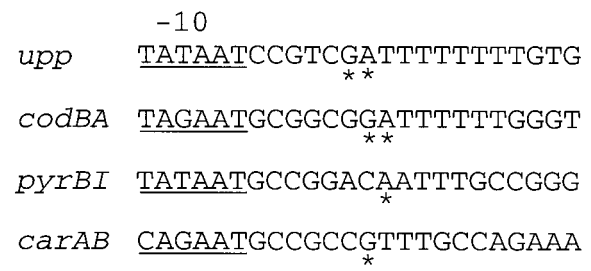

FIG. 1. Sequences of the initially transcribed regions of the *upp*, *codBA*, *pyrBI*, and *carAB* P1 promoters. The nontemplate strand sequence is shown. The −10 region of each promoter is underlined, and the major transcriptional start sites are indicated by asterisks. Start sites are identified in the text according to their positions downstream from the −10 region; e.g., the major *upp* start sites are designated G6 and A7.

enzymes are regulated by mechanisms that employ reiterative transcription. The *codBA* operon, which encodes the pyrimidine salvage enzymes cytosine permease and cytosine deaminase, is regulated over a 30-fold range by a mechanism that is entirely analogous to that described for the *upp* gene (29). Interestingly, the *codBA* initially transcribed region, GATTTTTTG, contains only a 6-bp T · A tract, and the G and A start sites (i.e., G7 and A8) are 1 base further downstream from the promoter −10 region than in the *upp* promoter (Fig. 1). The latter difference contributes significantly to the higher range of *codBA* regulation. (S. M. Dylla and C. L. Turnbough, Jr., unpublished data).

The *pyrBI* operon, which encodes the two subunits of the pyrimidine biosynthetic enzyme aspartate transcarbamylase, is regulated over a sevenfold range by a UTP-sensitive reiterative transcription mechanism that differs fundamentally from the *upp* and *codBA* mechanisms (20). The *pyrBI* initially transcribed region, AATTTGCC, contains a 3-bp T · A tract and a single transcriptional start site (Fig. 1). The key regulatory event occurs after the synthesis of the first five bases of the transcript, AAUUU, at which point the transcript can reversibly slip on the DNA template and has the potential to engage in reiterative transcription. The extent of reiterative transcription versus strictly templated transcription is controlled by the intracellular level of UTP. High UTP levels favor reiterative transcription, which produces transcripts with the sequence $AAUUUU_n$ (where n equals 1 to >30). These transcripts are always aborted. In contrast, low UTP levels favor the addition of a G residue as the sixth base in the transcript. This addition precludes reiterative transcription, and the AAUUUG transcript may be elongated to a full-length *pyrBI* transcript. The *carAB* operon, which encodes the two subunits of the pyrimidine biosynthetic enzyme carbamoyl phosphate synthetase, is regulated over a threefold range by a mechanism equivalent to that described for the *pyrBI* operon (Fig. 1) (12). The function of the four control mechanisms described above is to adjust pyrimidine salvage and biosynthetic enzyme levels to the cellular need for pyrimidine nucleotides.

A comparison of the reiterative transcription control mechanisms for the *upp*, *codBA*, *pyrBI*, and *carAB* operons and a rudimentary understanding of the requirements for reiterative transcription raise an obvious question. Is the very long T · A tract in the *upp* initially transcribed region required for regulation? In this study, we investigated this question. We constructed *upp* promoter mutations that systematically shorten the T · A tract in the initially transcribed region and examined the effects of these mutations on *upp* expression and regulation. The results indicate, unexpectedly, that a very long T · A tract (i.e., 7 or 8 bp) is, in fact, required for normal regulation. Additional examination of *upp* transcription provided an explanation for this requirement.

MATERIALS AND METHODS

Bacterial strains. *E. coli* K-12 strain CLT42 [F^- *car-94* Δ(*argF-lac*)*U169 rpsL150 thiA1 relA1 deoC1 ptsF25 flbB5301 rbsR*] (32) was used as the parent in the construction of seven lambda lysogens used in this study. These strains were constructed by inserting into the CLT42 chromosomal lambda attachment site a single copy of a recombinant lambda bacteriophage that carries either the wild type or a mutant version of a *upp*::*lacZ* gene fusion. The wild-type fusion contains the wild-type *upp* promoter region, while the six mutant fusions contain 1- to 6-bp deletions in the T · A tract of the *upp* initially transcribed region (i.e., T tract in Fig. 1). The construction of two of the lysogens, which carry either the wild-type or 6-bp deletion promoter region in the gene fusion, was described previously (35). Essentially the same procedure was used here to construct five additional lysogens in which the *upp*::*lacZ* gene fusions contain 1- to 5-bp deletions in the T · A tract of the *upp* initially transcribed region. A brief summary of the steps involved in the construction of all seven lysogens is provided below.

Construction of gene fusions. Construction of *upp*::*lacZ* gene fusions employed the multicopy plasmid pMLB1034, which contains the *lacZ* gene without a promoter, a ribosomal binding site, and the first eight codons for β-galactosidase (34). This plasmid contains an *Eco*RI/*Sma*I/*Bam*HI cloning site immediately preceding the *lacZ* gene. Gene fusions were made by first digesting plasmid pMLB1034 with *Eco*RI and *Bam*HI and then ligating the linear plasmid to an *Eco*RI-*Bam*HI restriction fragment containing the wild-type *upp* promoter region (from −100 to +127, numbering from the first transcriptional initiation site) or an equivalent fragment containing a mutant promoter region. The downstream sequence in these promoter region fragments extends through *upp* codon 30. The resulting plasmids were transformed into strain CLT240 (CLT42 *pcnB80 zad*::Tn*10*). The *pcnB80* mutation of this strain reduces the plasmid copy number, which is essential for maintaining (at least some) mutant fusion plasmids free of secondary mutations that reduce *upp*::*lacZ* expression. All fusion constructions were confirmed by DNA sequence analysis.

Transfer of gene fusions from plasmids to the *E. coli* chromosome. Wild-type and mutant *upp*::*lacZ* gene fusions carried on derivatives of plasmid pMLB1034 (in strain CLT240) were individually transferred to the chromosome of strain CLT42 by using phage lambda RZ5 (31). The presence of a single prophage at the lambda attachment site on the chromosome was determined by PCR analysis (35).

Restriction digests, ligations, transformations, PCR, site-directed mutagenesis, and DNA preparations. Conditions for restriction digests, ligations, and transformations were as previously described (32). PCR amplification of DNA was performed with *Pfu* DNA polymerase (Stratagene), using the reaction mixture recommended by the supplier. PCR conditions were: 95°C for 5 min; then 95°C for 1 min, 60°C for 1 min, and 72°C for 2 min for 30 cycles; and finally 72°C for 5 min. Site-directed mutations were introduced into the *upp* promoter region by using a PCR-based procedure similar to that described by Barettino et al. (6). The resulting mutations were verified by DNA sequence analysis. DNA was prepared essentially as previously described (35).

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% (wt/vol) glucose, 0.015 mM thiamine hydrochloride, 1 mM arginine, and either 1 mM uracil or 0.25 mM UMP. Cultures were incubated at 37°C with shaking and grown to an optical density at 650 nm of 0.5 (mid-log phase). Culture densities were measured with a Gilford model 260 spectrophotometer, and doubling times (which vary slightly with culture density when UMP is the pyrimidine source) were determined between optical densities of 0.1 and 0.2

Enzyme assays. Cell extracts were prepared by sonic oscillation (31). β-Galactosidase activity (24) and protein concentration (22) were determined as previously described.

Isolation of cellular RNA and primer extension mapping. Cellular RNA was isolated quantitatively as described by Wilson et al. (36). Primer extension mapping of the 5′ ends of *upp*::*lacZ* transcripts was performed as described by Liu and Turnbough (21), except that 50 μg of RNA from uracil-grown cells and 42

μg of RNA from UMP-grown cells were used for analysis. The different amounts of RNA, which were isolated from the same mass of cells, reflect the different levels of stable RNA in cells growing at different rates. The primer used in these experiments was 5′TTTTCCCAGTCACGACGTTG, which was labeled with ^{32}P at the 5′ end (5). This primer hybridizes to the *lacZ* sequence just downstream from the fusion junction in the *upp*::*lacZ* transcript. In these experiments, the primer was in large molar excess. In addition to the standard procedure of identifying transcript-specific primer extension products by alignment with bands in a sequencing ladder, identities were confirmed by spiking sequencing reactions with primer extension products prior to analysis by gel electrophoresis (data not shown).

In vitro transcription. Purified RNA polymerase holoenzyme containing σ^{70} was prepared as previously described (8, 9, 11). DNA templates for in vitro transcription were gel-purified *Pvu*II restriction fragments (derived from plasmid DNA) containing either the wild-type or mutant *upp* promoter regions. These fragments were identical to the *Eco*RI-*Bam*HI restriction fragments used for construction of *upp*::*lacZ* gene fusions, except for several base pairs at each end. Transcription reaction mixtures (10 μl) contained 10 nM DNA template; 100 nM RNA polymerase; 20 mM Tris-acetate (pH 7.9); 10 mM magnesium acetate; 100 mM potassium glutamate; 0.2 mM Na_2EDTA; 0.1 mM dithiothreitol; 200 μM each ATP, CTP, and GTP; and either 50 or 1,000 μM UTP. The reaction mixture included either [γ-^{32}P]ATP or [γ-^{32}P]GTP (purchased from NEN) at a specific activity of 0.625 Ci/mmol. Reactions were initiated by addition of RNA polymerase, and the reaction mixtures were incubated at 37°C for 15 min. Heparin (1 μl of a 1-mg/ml solution) was then added to the mixture, and incubation was continued for an additional 10 min to permit the completion of elongating transcripts. Reactions were terminated by adding 10 μl of stop solution (7 M urea, 2 mM Na_2EDTA, 0.025% [wt/vol] each bromophenol blue and xylene cyanol) and placing the samples on ice. The samples were heated at 100°C for 3 min, and an equal volume of each sample was removed and run on a 25% polyacrylamide (29:1 acrylamide-bisacrylamide ratio)–50 mM Tris-borate (pH 8.3)–1 mM Na_2EDTA sequencing gel containing 7 M urea (28). Transcripts were visualized by autoradiography and quantitated by scanning gels with a Molecular Dynamics PhosphorImager. Transcripts were identified as previously described (35).

RESULTS

Effects of shortening the T · A tract in the *upp* initially transcribed region on *upp*::*lacZ* expression and regulation. To examine the regulatory role of the long T · A tract in the *upp* initially transcribed region, we first constructed seven isogenic *E. coli* strains each containing either a wild-type or a mutant *upp*::*lacZ* gene fusion. Each gene fusion was carried on a lambda bacteriophage and inserted in single copy into the chromosomal lambda attachment site of strain CLT42 (*car-94* Δ*lacZYA*). The wild-type fusion contains the wild-type *upp* promoter region, and the six mutant fusions contain 1- to 6-bp deletions in the 8-bp T · A tract of the *upp* initially transcribed region. The promoters for the *upp*::*lacZ* gene fusions (and the fusions themselves) are hereafter referred to as T_n promoters (and fusions), where *n* equals the number of T · A base pairs present in the initially transcribed region.

The seven strains were then used to measure the effects of the deletions on *upp*::*lacZ* expression and regulation. These strains, which are pyrimidine auxotrophs because the *car-94* mutation inactivates the first enzyme of the pyrimidine nucleotide biosynthetic pathway, 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 and high intracellular levels of UTP, while growth on UMP, which is only slowly used by cells growing in this medium, results in pyrimidine limitation and low intracellular levels of UTP (C. L. Turnbough, Jr., unpublished data). The level of *upp*::*lacZ* expression in each culture was determined by measuring the fusion-encoded β-galactosidase activity.

TABLE 1. Effects of $\Delta(T \cdot A)_n$ mutations on *upp*::*lacZ* expression and regulation[a]

Strain (T · A tract)	β-Galactosidase activity (nmol/min/mg)[b]		Fold regulation
	Uracil	UMP	
CLT5178 (T_8 [wild type])	1,580	9,020	5.71
CLT5249 (T_7)	1,570	8,610	5.48
CLT5248 (T_6)	3,200	12,700	3.97
CLT5247 (T_5)	4,540	15,300	3.37
CLT5246 (T_4)	5,690	13,900	2.44
CLT5245 (T_3)	6,990	10,300	1.47
CLT5218 (T_2)	5,940	9,030	1.52

[a] Doubling times were 47 ± 1 min for cells grown on uracil and 66 ± 2 min for cells grown on UMP.
[b] Means of six experiments with standard deviations of ≤7%.

The results show that wild-type (T_8) fusion expression was regulated over a nearly sixfold range by pyrimidine availability (Table 1). Expression and regulation of the T_7 fusion were essentially the same as those of the wild-type fusion. In contrast, regulation of *upp*::*lacZ* expression was significantly and steadily decreased as the length of the T · A tract was gradually reduced to 3 bp, at which point only a basal level of regulation (i.e., 1.5-fold) was detectable. This basal level of regulation, detected with both the T_3 and T_2 fusions, is most likely unrelated to control by reiterative transcription, because this reaction does not occur at the T_2 promoter (see below). The steady decrease in regulation observed with the T_6, T_5, T_4, and T_3 fusions was due primarily to large percentage increases of *upp*::*lacZ* expression in cells grown under conditions of pyrimidine excess (i.e., on uracil). Expression of these fusions also changed in cells grown under conditions of pyrimidine limitation (i.e., on UMP), but not as uniformly or dramatically (on a percentage basis). Expression levels increased gradually as the T · A tract was shortened to 5 bp and then decreased gradually as the T · A tract was shortened further. Possible reasons for this pattern are discussed below. Overall, the results indicate that shortening of the T · A tract relieves negative regulation imposed by the reiterative transcription control mechanism.

Analysis of *upp*::*lacZ* transcripts initiated at wild-type and mutant promoters. To further elucidate the effects of the deletions in the T · A tract, we used quantitative primer extension mapping to measure the levels and determine the start sites of *upp*::*lacZ* transcripts synthesized in wild-type and mutant fusion strains grown on uracil or UMP. Cellular RNA was isolated from cultures that were essentially identical to those described in Table 1. The primer used in these experiments hybridizes to the *lacZ* sequence contained in the fusions and therefore detects only *upp*::*lacZ* transcripts in the Δ*lacZYA* strains. It should also be noted that nonproductive transcripts initiated at positions G6 and A7 (e.g., GAU_n and AU_n) are not detected in this assay. The results are shown in Fig. 2. In the case of the wild-type (T_8) and T_7 fusions, only transcripts initiated at the G6 start site were detected in uracil- and UMP-grown cells (Fig. 2A). The relative levels of these transcripts closely paralleled the cellular β-galactosidase activities described above (Fig. 2B and Table 1), as expected for transcriptionally controlled fusions. In the case of the fusions containing shorter T · A tracts, transcripts initiated at both *upp* start sites

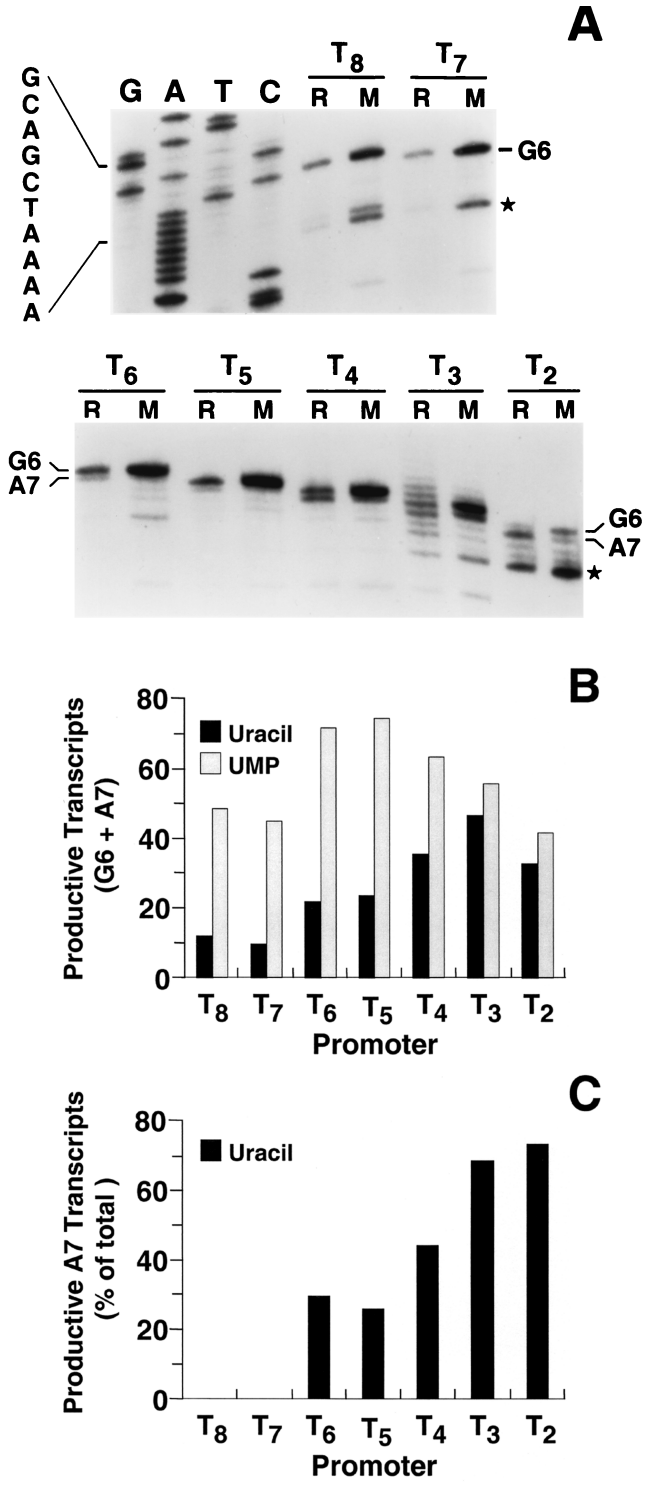

FIG. 2. Levels of *upp*::*lacZ* transcripts initiated at the wild-type (T_8) and mutant (T_n) promoters. Cellular RNA was quantitatively isolated from cells grown on either uracil (R) or UMP (M). Transcript levels were measured by primer extension mapping as described in Materials and Methods. (A) Autoradiograph of the 10% polyacrylamide sequencing gel that was used to separate the primer extension products. The autoradiograph was cut between the T_7 and T_6 lanes, and one part was placed beneath the other to facilitate labeling. The dideoxy sequencing ladder of the wild-type *upp* promoter region, which was used to identify transcripts, was generated with the same primer that was used for primer extension. (Note that the DNA sequence is for the template strand and is complementary to the transcript sequences.) The bands corresponding to G6 and (where detected) A7 transcripts are indicated. The positions of these bands in the gel are different for each promoter, reflecting the number of U residues specified by the T · A tract. Bands marked with an asterisk represent major transcripts produced by in vivo degradation of G6 and A7 transcripts. (B) Total productive (i.e., *upp*::*lacZ*) transcript levels in uracil- and UMP-grown cells were quantitated using a PhosphorImager and plotted in arbitrary units. Total transcript levels included G6 and A7 transcripts, degradation products, and escaped stuttering products in the case of the T_3 fusion. (C) The levels of productive A7-initiated transcripts in uracil-grown cells were plotted as percentages of total productive transcript levels, calculated using the formula (A7/G6 + A7) × 100. In the case of the T_3 fusion, escaped stuttering products were counted as A7 transcripts.

(i.e., G6 and A7) were detected, indicating avoidance of reiterative transcription by A7 transcripts. At least with the T_6, T_5, and T_4 fusions, G6 transcript levels were also increased (Fig. 2A), indicating that these transcripts more efficiently avoided reiterative transcription. As with the wild-type and T_7 fusions, the relative levels of total *upp*::*lacZ* transcripts specified by the T_6, T_5, T_4, T_3, and T_2 fusions closely reflected cellular β-galactosidase activities.

The length of the T · A tract in the *upp* initially transcribed region appeared to have a major effect on the ability of fusion transcripts to avoid reiterative transcription. This effect was most clearly seen with A7 transcripts in uracil-grown cells. In this case, a clear inverse relationship was observed between the length of the T · A tract and the percentage of *upp*::*lacZ* transcripts that were initiated at position A7 (Fig. 2C).

Primer extension mapping revealed a number of anomalous transcripts. Transcripts shorter than the G6 and A7 transcripts were detected (some marked with asterisks in Fig. 2A). These transcripts were previously characterized as products of in vivo *upp* transcript degradation (35). It was also shown that G6 and A7 transcripts are degraded similarly. Another set of anomalous transcripts was observed with the T_3 fusion (Fig. 2A). A ladder of transcripts longer than the G6 transcript was detected. The transcripts in this ladder are apparently transcripts that are initiated at position A7, undergo a limited number of UMP additions as a result of reiterative transcription within the 3-bp T · A tract, and then switch to a normal mode of transcriptional elongation. Thus, these transcripts are full-length *upp*::*lacZ* transcripts containing a 5′ A residue, followed by a run of more than the three U residues specified by the T · A tract. Similar switching from reiterative transcription to normal elongation has been observed with a mutant *pyrBI* promoter, with an ATTTG initially transcribed region (F. Qi and C. L. Turnbough, Jr., unpublished data), and other mutant bacterial and phage promoters (14, 37).

Analysis of short transcripts initiated at wild-type and mutant *upp* promoters in vitro. To directly analyze the effects of the deletions in the T · A tract of the *upp* initially transcribed region on reiterative transcription, we examined the short transcripts produced by transcription of wild-type and mutant *upp* promoter regions in vitro. These short transcripts include both the aborted products of reiterative transcription and the products of simple abortive initiation involving strictly templated transcription. Transcripts up to approximately 12 bases in length can be produced by simple abortive initiation at the *upp*

promoter. We analyzed G6 and A7 transcripts separately by using either [γ-^{32}P]GTP or [γ-^{32}P]ATP, respectively, to label the 5′ ends of transcripts. Transcription reaction mixtures contained either 50 or 1,000 μM UTP (200 μM each other nucleoside triphosphate) for the synthesis of G6 and A7 transcripts, respectively, to maximize initiation at the start site of interest. These UTP concentrations (i.e., 50 and 1,000 μM) roughly mimic those found in cells grown under conditions of pyrimidine limitation or excess, respectively (2, 25). Transcripts produced in vitro were separated on a 25% polyacrylamide sequencing gel and visualized by autoradiography.

In the case of the G6 transcripts, the shortest transcripts that can be unambiguously attributed to reiterative transcription contain the sequence GAU_{n+1}, where *n* equals the number of base pairs in the T · A tract (e.g., GAU_9 for the wild-type promoter, GAU_8 for the T_7 promoter, etc.). If extensive reiterative transcription occurs at a particular promoter, then G6 transcripts with longer terminal U tracts can also be detected (e.g., GAU_{10} and GAU_{11} for the wild-type promoter, GAU_9 and GAU_{10} for the T_7 promoter, etc.). Transcripts (either G6 or A7) with these longer terminal U tracts are hereafter referred to as longer ladder transcripts. The shortest G6 transcripts that can be unambiguously attributed to strictly templated transcription (i.e., simple abortive initiation) contain the sequence GAU_nG (e.g., GAU_8G for the wild-type promoter, GAU_7G for the T_7 promoter, etc.). A GAU_nG transcript will migrate slightly slower in the gel than an equal-length GAU_{n+1} transcript because of sequence effects on transcript mobility, thus allowing resolution of the two transcripts (35) (e.g., in Fig. 3, the 11-mer transcripts GAU_8G and GAU_9 initiated at the wild-type promoter, the 10-mers GAU_7G and GAU_8 initiated at the T_7 promoter, etc.). Similarly, longer abortive initiation products such as GAU_nGU and GAU_nGUG can be resolved from equal-length G6 transcripts produced by reiterative transcription (i.e., with a terminal U tract) because of different nucleotide content (e.g., in Fig. 3, the 12-mers GAU_8GU and GAU_{10} initiated at the wild-type promoter). Under the conditions employed here, nucleotide addition to a transcript retards its gel mobility in the following order: G > A ≈ U > C (29).

Transcription from the wild-type (T_8) and T_7 promoters produced levels of GAU_{n+1} and longer ladder transcripts that were comparable in most cases to those of equal-length transcripts produced by simple abortive initiation (Fig. 3). For example, compare the levels of 11-mers (i.e., GAU_8G and GAU_9) in the T_8 lane and the levels of 10-mers (i.e., GAU_7G and GAU_8) in the T_7 lane in Fig. 3. This result indicated that a significant fraction of G6 transcripts initiated at these promoters were subject to reiterative transcription. Transcription from the promoters containing shorter T · A tracts produced a substantially lower level of GAU_{n+1} and longer ladder transcripts compared to transcripts produced by simple abortive initiation. For example, compare the levels of 9-mers (i.e., GAU_6G and GAU_7) in the T_6 lane and the levels of 8-mers (i.e., GAU_5G and GAU_6) in the T_5 lane in Fig. 3. In addition, the level of GAU_{n+1} and longer ladder transcripts compared to simple aborted transcripts decreased in proportion to the size of the deletion in the T · A tract for all promoters. Synthesis of GAU_{n+1} and longer ladder transcripts was effectively eliminated at the T_3 and T_2 promoters. These results clearly illustrate the need for a long T · A tract to permit high levels of reiterative transcription with G6 transcripts. We also examined transcription of the wild-type and mutant promoters in reaction mixtures containing either 200 or 1,000 μM UTP (data not shown). The effects of the mutations on reiterative transcription were similar to those observed with 50 μM UTP; however, increasing the UTP concentration did stimulate slightly the level of reiterative transcription at all but the wild-type and T_2 promoters.

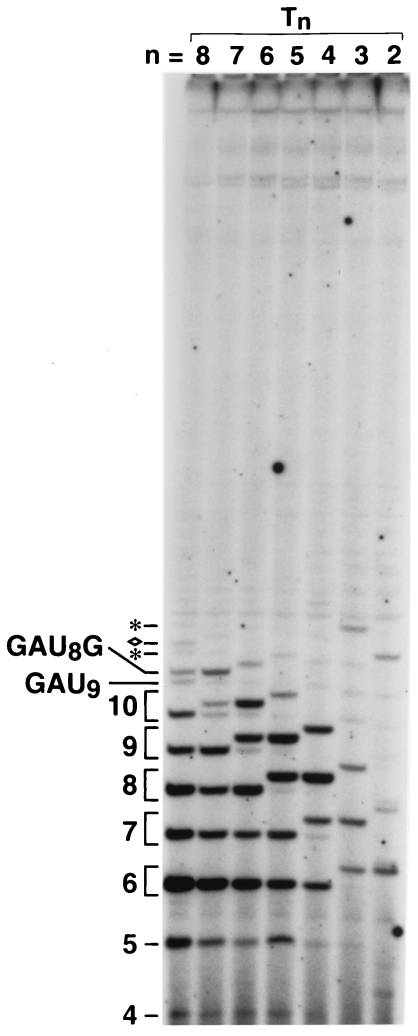

FIG. 3. Analysis of short G6 transcripts initiated at wild-type and mutant *upp* promoters. DNA templates containing either the wild-type (T_8) or a T_n mutant promoter were transcribed in vitro in reaction mixtures containing 200 μM each ATP, CTP, and [γ-^{32}P]GTP and 50 μM UTP. The autoradiograph is of the 25% polyacrylamide gel used to separate the transcripts. The numbers on the left indicate transcript length (in nucleotides); brackets are used to indicate equal-length transcripts with different sequences, which in most cases cause these transcripts to migrate differently in the gel. In the case of the wild-type promoter (i.e., lane T_8 only), several other transcripts are marked. Transcripts with the sequences GAU_9 and GAU_8G are labeled. Two longer GAU_n transcripts (i.e., GAU_{10} and GAU_{11}) are marked with asterisks. The longest unambiguously assigned product of simple abortive initiation, GAU_8GU, is marked with a diamond.

Examination of A7 transcript synthesis revealed a similar pattern; namely, a longer T · A tract favored reiterative tran-

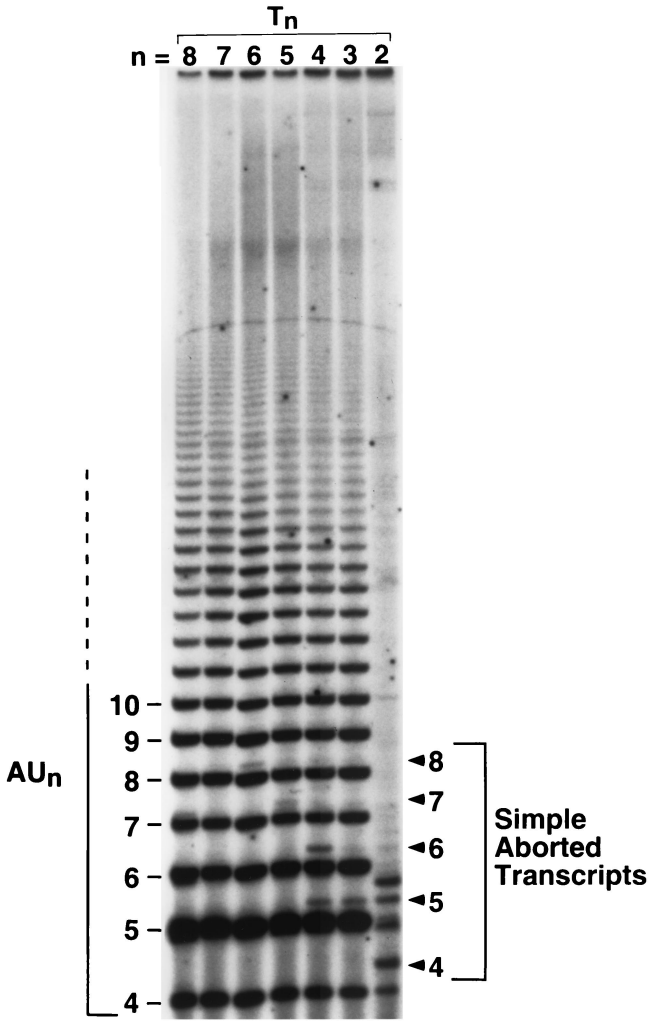

FIG. 4. Analysis of short A7 transcripts initiated at wild-type and mutant *upp* promoters. DNA templates containing either the wild-type (T_8) or a T_n mutant promoter were transcribed in vitro in reaction mixtures containing 200 μM each [γ-^{32}P]ATP, CTP, and GTP and 1,000 μM UTP. The autoradiograph is of the 25% polyacrylamide gel used to separate the transcripts. The lengths (in nucleotides) of transcripts with the sequence AU_n (where n is ≥3) are shown on the left. The lengths of transcripts produced by simple abortive initiation are shown on the right. Note that the identity of the minor 5-mer transcript initiated at the T_4 promoter is unknown. Also note that the AU_2GU transcript, which is produced by simple abortive initiation at the T_2 promoter, comigrates with the AU_3G transcript initiated at the T_3 promoter. Unlabeled transcripts initiated at the T_2 promoter are discussed in the text.

scription. The shortest A7 transcripts that can be unambiguously attributed to either reiterative transcription or simple abortive initiation contain the sequence AU_{n+1} or AU_nG, respectively. Again, because of sequence effects, the AU_nG transcript will migrate slightly slower in the gel than an AU_{n+1} transcript, allowing resolution of the two transcripts (e.g., in Fig. 4, see the 8-mer transcripts AU_6G and AU_7 initiated at the T_6 promoter). Transcription from the wild-type and T_7 promoters in vitro always appeared to enter the reiterative mode, as indicated by high levels of AU_{n+1} and longer ladder transcripts and the total absence of AU_nG (and longer) transcripts produced by abortive initiation (Fig. 4). Transcription from the T_6, T_5, T_4, and T_3 promoters also produced high levels of AU_{n+1} and longer ladder transcripts; however, a low level of AU_nG (and longer) transcripts produced by simple abortive initiation could also be detected. This result indicated that at least a low level of strictly templated transcription (compared to reiterative transcription) can occur at these promoters. In the case of the T_2 promoter, reiterative transcription was dramatically reduced, if not totally eliminated, as indicated by the absence of a longer ladder of A7 transcripts. Low levels of several unidentified short transcripts were detected, two of which migrate like AU_3 and AU_4 transcripts (Fig. 4). Synthesis of the latter transcripts was nearly eliminated when the UTP concentration in the reaction mixture was reduced to 200 μM (data not shown), suggesting that their synthesis was the result of misincorporation enhanced by unbalanced (and nonphysiological) nucleotide concentrations.

DISCUSSION

Although extensive reiterative transcription can occur at promoters that contain a T · A tract in the initially transcribed region as short as 3 bp (12, 17, 20), a much longer T · A tract is clearly required for normal pyrimidine-mediated regulation of *upp* expression in *E. coli*. The 8-bp T · A tract in the wild-type *upp* promoter can be shortened by a single base pair with minimal effect, but longer deletions significantly and progressively reduce the range of regulation. Regulation involving reiterative transcription appears to be completely eliminated when the T · A tract contains three or fewer base pairs.

The progressive loss of regulation caused by the deletion mutations is due primarily to regular increases in the level of *upp* expression in cells grown under conditions of pyrimidine excess (Table 1). These increases are due, in large part, to the avoidance of reiterative transcription by a significant fraction of A7 transcripts (Fig. 2A), which are the predominant products of transcriptional initiation at the *upp* promoter in uracil-grown cells (35). The fraction of A7 transcripts that avoid reiterative transcription increases as the length of the T · A tract decreases, as clearly indicated by the quantitative primer extension mapping of cellular transcripts (Fig. 2C).

A similar but less dramatic pattern of avoidance of reiterative transcription by A7 transcripts is observed in vitro (Fig. 4). Transcription of wild-type and mutant *upp* promoters indicates that avoidance of reiterative transcription by A7 transcripts, resulting in strictly templated transcription, can occur only when the T · A tract is shortened to six or fewer base pairs. The level of strictly templated transcription, as indicated by the products of simple abortive initiation, increases as the length of the T · A tract is reduced. However, a high level of reiterative transcription still occurs in vitro at all promoters except the T_2 promoter, at which reiterative transcription is abolished. The fact that reiterative transcription is so robust at promoters like T_3 and T_4, while regulation is severely restricted at these promoters, is somewhat surprising. One possible explanation is that only a very low percentage of initiation events need to be redirected from the reiterative to the strictly templated transcription pathway to produce the maximum level of full-length transcripts. Another explanation, at least for the T_3 promoter, is that some transcripts can enter the reiterative transcription

pathway and then switch to the strictly templated mode to produce full-length transcripts. Clear evidence for such a switch at the T_3 promoter is provided by the results of quantitative primer extension mapping of cellular transcripts (Fig. 2A). Why such a switch is restricted to the T_3 promoter remains to be determined.

Avoidance of reiterative transcription by A7 transcripts initiated at mutant *upp* promoters may explain the loss of most, if not all, regulation of *upp* expression. However, avoidance of reiterative transcription at the mutant promoters is not restricted to A7 transcripts. The data from quantitative primer extension mapping of cellular transcripts indicate that approximately one-third of the G6 transcripts initiated at the wild-type *upp* promoter are subject to reiterative transcription (Fig. 2B). This reiterative transcription appears to be affected by shortening of the T · A tract similarly to that observed with the A7 transcripts. For example, reducing the length of the T · A tract to six or fewer base pairs causes an increase in the level of cellular G6 transcripts, which in this case can be detected in cells grown under conditions of either pyrimidine excess or limitation (Fig. 2A). However, the pattern of these increases is different than that observed with A7 transcripts. There is not a steady increase in the level of G6 transcripts with progressive shortening of the T · A tract. Instead, the highest levels of G6 transcripts are observed with the T_6 and T_5 promoters, with progressively lower levels of G6 transcripts with promoters T_4, T_3, and T_2. This pattern may indicate that shortening of the T · A tract to 6 bp is sufficient to achieve maximum avoidance of reiterative transcription by G6 transcripts. Consistent with this idea is the observation that in vitro, reiterative transcription involving G6 transcripts was sharply reduced by shortening of the T · A tract to six or fewer base pairs (Fig. 3). The relative decreases observed in vivo in G6 transcript levels with promoters T_4, T_3, and T_2 may reflect secondary effects on promoter strength or transcript stability. Consistent with these proposals, in vitro production of all short G6 transcripts at promoters T_4, T_3, and T_2 was relatively low (Fig. 3), perhaps indicating reduced promoter activity, and in vivo degradation of G6 and A7 *upp*::*lacZ* transcripts was greatly enhanced in the case of the T_2 promoter (Fig. 2A).

Overall, the results of this study indicate that a long T · A tract in the initially transcribed region of the wild-type *upp* promoter is necessary to achieve a particular balance between reiterative and strictly templated transcription involving G6 and A7 transcripts. This balance establishes a maximum or physiologically appropriate level of *upp* regulation. A long run of seven or eight T · A base pairs is required to ensure that essentially all A7 transcripts will engage in nonproductive reiterative transcription. A run of eight T · A base pairs is still short enough to allow the majority of G6 transcripts to avoid reiterative transcription and produce translatable mRNA. Presumably, the avoidance of reiterative transcription by G6 transcripts can occur by virtue of an rG · dC base pair formed by the 5′ G residue of the transcript and its complementary C residue in the DNA template. The formation of this strong base pair inhibits upstream slippage of the nascent transcript that is a prerequisite for reiterative transcription.

The ability of G6 transcripts to avoid reiterative transcription is likely to be related to another feature of the transcriptional initiation complex, and that is the length of the RNA-DNA hybrid formed between the nascent transcript and the DNA template. The results presented in this paper are consistent with the formation of an 8-bp RNA-DNA hybrid during initiation of G6 transcripts. Such a hybrid would explain the elimination of essentially all reiterative transcription involving G6 transcripts at mutant promoters containing six or fewer base pairs in the T · A tract. The nascent transcripts synthesized at these promoters would be part of a relatively stable hybrid, anchored by a 5′ rG · dC base pair, during the addition of all of the U residues specified by the T · A tract. The fact that about one-third of the G6 transcripts initiated at the T_7 and wild-type (T_8) promoters engage in reiterative transcription indicates that the RNA-DNA hybrid, at least a permanent hybrid, does not extend beyond 8 bp. The factors that allow most of the latter transcripts to avoid reiterative transcription (assuming that the 8-bp hybrid is correct) remain to be determined. Thus, the suggested length of the RNA-DNA hybrid at the *upp* promoter is the same as, or similar to, the 8- to 9-bp hybrid detected during transcriptional elongation (27, 33). It will be of interest to determine in future studies if the RNA-DNA hybrid length is the same at all promoters. If not, then the capacity to engage in reiterative transcription during initiation could be affected by this difference.

ACKNOWLEDGMENTS

Y.C. and S.M.D. contributed equally to this work, and either person could have been listed as the first author.

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. **Andersen, J. T., K. F. Jensen, and P. Poulsen.** 1991. Role of transcription pausing in the control of the *pyrE* attenuator in *Escherichia coli*. Mol. Microbiol. **5:**327–333.
3. **Andersen, P. S., D. Frees, R. Fast, and B. Mygind.** 1995. Uracil uptake in *Escherichia coli* K-12: isolation of *uraA* mutants and cloning of the gene. J. Bacteriol. **177:**2008–2013.
4. **Andersen, P. S., J. M. Smith, and B. Mygind.** 1992. Characterization of the *upp* gene encoding uracil phosphoribosyltransferase of *Escherichia coli* K12. Eur. J. Biochem. **204:**51–56.
5. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. S. Struhl (ed.).** 1989. Current protocols in molecular biology. John Wiley & Sons, Inc., New York, N.Y.
6. **Barettino, D., M. Feigenbutz, R. Valcárcel, and H. G. Stunnenberg.** 1993. Improved method for PCR-mediated site-directed mutagenesis. Nucleic Acids Res. **22:**541–542.
7. **Blattner, F. R., G. Plunkett III, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao.** 1997. The complete genome sequence of *Escherichia coli* K-12. Science **277:**1453–1462.
8. **Burgess, R. R., and J. J. Jendrisak.** 1975. A procedure for the rapid, large-scale purification of *Escherichia coli* DNA-dependent RNA polymerase involving polymin P precipitation and DNA-cellulose chromatography. Biochemistry **14:**4634–4638.
9. **Gonzalez, N., J. Wiggs, and M. J. Chamberlin.** 1977. A simple procedure for resolution of *Escherichia coli* RNA polymerase holoenzyme from core polymerase. Arch. Biochem. Biophys. **182:**404–408.
10. **Guo, H.-C., and J. W. Roberts.** 1990. Heterogeneous initiation due to slippage at the bacteriophage 82 late gene promoter in vitro. Biochemistry **29:**10702–10709.
11. **Hager, D. A., D. J. Jin, and R. R. Burgess.** 1990. Use of mono Q high-resolution ion-exchange chromatography to obtain highly pure and active *Escherichia coli* RNA polymerase. Biochemistry **29:**7890–7894.
12. **Han, X., and C. L. Turnbough, Jr.** 1998. Regulation of *carAB* expression in *Escherichia coli* occurs in part through UTP-sensitive reiterative transcription. J. Bacteriol. **180:**705–713.
13. **Hausmann, S., D. Garcin, C. Delenda, and D. Kolakofsky.** 1999. The versa-

tility of paramyxovirus RNA polymerase stuttering. J. Virol. **73:**5568–5576.
14. **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.
15. **Jacques, J.-P., and D. Kolakofsky.** 1991. Pseudo-templated transcription in prokaryotic and eukaryotic organisms. Genes Dev. **5:**707–713.
16. **Jeong, S. W., W. H. Lang, and R. H. Reeder.** 1996. The yeast transcription terminator for RNA polymerase I is designed to prevent polymerase slippage. J. Biol. Chem. **271:**16104–16110.
17. **Jin, D. J.** 1994. Slippage synthesis at the *galP2* promoter of *Escherichia coli* and its regulation by UTP concentration and cAMP-cAMP receptor protein. J. Biol. Chem. **269:**17221–17227.
18. **Larsen, B., N. M. Wills, C. Nelson, J. F. Atkins, and R. F. Gesteland.** 2000. Nonlinearity in genetic decoding: homologous DNA replicase genes use alternatives of transcriptional slippage or translational frameshifting. Proc. Natl. Acad. Sci. USA **97:**1683–1688.
19. **Linton, M. F., M. Raabe, V. Pierotti, and S. G. Young.** 1997. Reading-frame restoration by transcriptional slippage at long stretches of adenine residues in mammalian cells. J. Biol. Chem. **272:**14127–14132.
20. **Liu, C., L. S. Heath, and C. L. Turnbough, Jr.** 1994. Regulation of *pyrBI* operon expression in *Escherichia coli* by UTP-sensitive reiterative RNA synthesis during transcriptional initiation. Genes Dev. **8:**2904–2912.
21. **Liu, J., and C. L. Turnbough, Jr.** 1994. Effects of transcriptional start site sequence and position on nucleotide-sensitive selection of alternative start sites at the *pyrC* promoter in *Escherichia coli*. J. Bacteriol. **176:**2938–2945.
22. **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.
23. **McClure, W. R., C. L. Cech, and D. E. Johnston.** 1978. A steady state assay for the RNA polymerase initiation reaction. J. Biol. Chem. **253:**8941–8948.
24. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **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.
26. **Nierman, W. C., and M. J. Chamberlin.** 1979. Studies of RNA chain initiation by *Escherichia coli* RNA polymerase bound to T7 DNA. Direct analysis of the kinetics and extent of RNA chain initiation at T7 promoter A_1. J. Biol. Chem. **254:**7921–7926.
27. **Nudler, E., A. Mustaev, E. Lukhtanov, and A. Goldfarb.** 1997. The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. Cell **89:**33–41.
28. **Qi, F., C. Liu, L. S. Heath, and C. L. Turnbough, Jr.** 1996. *In vitro* assay for reiterative transcription during transcriptional initiation by *Escherichia coli* RNA polymerase. Methods Enzymol. **273:**71–85.
29. **Qi, F., and C. L. Turnbough, Jr.** 1995. Regulation of *codBA* operon expression in *Escherichia coli* by UTP-dependent reiterative transcription and UTP-sensitive transcriptional start site switching. J. Mol. Biol. **254:**552–565.
30. **Rasmussen, U. B., B. Mygind, and P. Nygaard.** 1986. Purification and some properties of uracil phosphoribosyltransferase from *Escherichia coli* K12. Biochim. Biophys. Acta **881:**268–275.
31. **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.
32. **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.
33. **Sidorenkov, I., N. Komissarova, and M. Kashlev.** 1998. Crucial role of the RNA:DNA hybrid in the processivity of transcription. Mol. Cell **2:**55–64.
34. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
35. **Tu, A.-H. T., and C. L. Turnbough, Jr.** 1997. Regulation of *upp* expression in *Escherichia coli* by UTP-sensitive selection of transcriptional start sites coupled with UTP-dependent reiterative transcription. J. Bacteriol. **179:**6665–6673.
36. **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*. J. Bacteriol. **174:**514–524.
37. **Xiong, X. F., 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.