AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology®

RESEARCH ARTICLE

Novel Translation Initiation Regulation Mechanism in *Escherichia coli ptrB* Mediated by a 5′-Terminal AUG

Heather J. Beck,* Gary R. Janssen†

Department of Microbiology, Miami University, Oxford, Ohio, USA

ABSTRACT Alternative translation initiation mechanisms, distinct from the Shine-Dalgarno (SD) sequence-dependent mechanism, are more prevalent in bacteria than once anticipated. Translation of *Escherichia coli ptrB* instead requires an AUG triplet at the 5′ terminus of its mRNA. The 5′-terminal AUG (5′-uAUG) acts as a ribosomal recognition signal to attract ribosomes to the *ptrB* mRNA rather than functioning as an initiation codon to support translation of an upstream open reading frame. *ptrB* expression exhibits a stronger dependence on the 5′-uAUG than the predicted SD sequence; however, strengthening the predicted *ptrB* SD sequence relieves the necessity for the 5′-uAUG. Additional sequences within the *ptrB* 5′ untranslated region (5′-UTR) work cumulatively with the 5′-uAUG to control expression of the downstream *ptrB* coding sequence (CDS), thereby compensating for the weak SD sequence. Replacement of 5′-UTRs from other mRNAs with the *ptrB* 5′-UTR sequence showed a similar dependence on the 5′-uAUG for CDS expression, suggesting that the regulatory features contained within the *ptrB* 5′-UTR are sufficient to control the expression of other *E. coli* CDSs. Demonstration that the 5′-uAUG present on the *ptrB* leader mRNA is involved in ribosome binding and expression of the downstream *ptrB* CDS revealed a novel form of translational regulation. Due to the abundance of AUG triplets at the 5′ termini of *E. coli* mRNAs and the ability of *ptrB* 5′-UTR regulation to function independently of gene context, the regulatory effects of 5′-uAUGs on downstream CDSs may be widespread throughout the *E. coli* genome.

IMPORTANCE As the field of synthetic biology continues to grow, a complete understanding of basic biological principles will be necessary. The increasing complexity of the synthetic systems highlights the gaps in our current knowledge of RNA regulation. This study demonstrates that there are novel ways to regulate canonical Shine-Dalgarno-led mRNAs in *Escherichia coli*, illustrating that our understanding of the fundamental processes of translation and RNA regulation is still incomplete. Even for *E. coli*, one of the most-studied model organisms, genes with translation initiation mechanisms that do not fit the canonical Shine-Dalgarno sequence paradigm are being revealed. Uncovering diverse mechanisms that control translational expression will allow synthetic biologists to finely tune protein production of desired gene products.

KEYWORDS 5′ upstream AUG, Shine-Dalgarno, noncanonical initiation, translation initiation, translational regulation, upstream open reading frame

Received 16 February 2017 **Accepted** 1 May 2017

Accepted manuscript posted online 8 May 2017

Citation Beck HJ, Janssen GR. 2017. Novel translation initiation regulation mechanism in *Escherichia coli ptrB* mediated by a 5′-terminal AUG. J Bacteriol 199:e00091-17. https://doi.org/10.1128/JB.00091-17.

Editor Tina M. Henkin, Ohio State University

Address correspondence to Heather J. Beck, heather.beck@univie.ac.at.

* Present address: Heather J. Beck, Max F. Perutz Laboratories, Vienna, Austria.

† Deceased.

This article is dedicated to my mentor and beloved friend, Gary Janssen.

For a commentary on this article, see https://doi.org/10.1128/JB.00295-17.

Protein synthesis, in which the ribosome translates an mRNA sequence to form polypeptides, is a four-step process that includes initiation, elongation, termination, and ribosome recycling. Translation initiation is energy dependent and the most highly regulated phase of translation (1). The translational machinery interacts with the mRNA at its ribosome binding site (RBS) along with initiator tRNA to form a ternary complex that is equipped for polypeptide production. Both the mRNA sequence and its sec-

ondary structure affect how it interacts with the translational machinery and therefore influence its translation efficiency. The RBS of a canonical mRNA contains an initiation codon, a purine-rich Shine-Dalgarno (SD) sequence (2), and an appropriately sized spacer region between the two aforementioned elements (3, 4). These primary structural elements, as well as potential upstream and downstream enhancer sequences, contribute to translation efficiency.

An mRNA secondary structure can also impact the rate of translation. Ribosomes require a single-stranded region for efficient binding (5–7). A secondary structure that occludes the RBS will cause the region to be unrecognizable to the ribosome and therefore negatively impact translation (8–10). Thermodynamic changes in mRNA secondary structure as a result of physiological conditions can also regulate translation by either opening or occluding the RBS, thereby resulting in changes in initiation efficiency (11–13).

There are also examples of *cis*-acting elements, such as upstream open reading frames (uORFs), which can influence the expression of downstream open reading frames (ORFs) on the same transcript. In prokaryotes, uORFs exert their effects primarily through translational coupling (14–16). These effects occur after the ribosome completes translation of the uORF and, rather than dissociating, remains bound to the mRNA and repositions to the downstream initiation region via a scanning-like movement (17, 18). uORFs can also encode nascent peptides that cause ribosomal stalling by binding to regions of the peptidyltransferase center, blocking the ribosome exit tunnel, or through attenuation (19, 20). This stalling can then impact downstream translation by blocking the RBS, causing conformational changes in the secondary structure, or increasing degradation of the mRNA (21, 22).

We recently reported another example of the influences of uORFs on downstream expression (23). In that study, we identified novel examples of uORFs positioned at the 5′ terminus of SD sequence-led mRNAs in *Escherichia coli*, which were therefore classified as leaderless mRNAs. Leaderless mRNAs lack a 5′ untranslated region (5′-UTR) and are thought to initiate translation via a mechanism distinct from leadered mRNAs, in which an intact 70S ribosome binds to the 5′ terminus of the mRNA (24–26). In our previous study, the uORF of the *ptrB* mRNA was found to greatly influence expression of the *ptrB* coding sequence (CDS). Downstream expression of *ptrB* appeared to be dependent upon the 5′-terminal AUG (5′-uAUG), even though the 5′-terminal uORF (5′-uORF) was not efficiently translated (23).

The necessity of the *ptrB* 5′-uAUG and its role in the regulation of the *ptrB* CDS is the focus of this study. The annotated *ptrB* CDS produces a 686-amino-acid protease II (27), but little is currently known about its regulation. This study reinforces the dependency of *ptrB* translation on the 5′-uAUG, based on its ability to act as a ribosome binding signal rather than through translation of the uORF. Ribosome binding and expression assays revealed a regulatory process distinct from the canonical SD sequence-mediated regulation, suggesting a novel mechanism for translation regulation directed by a 5′-uAUG. These findings are significant because the *ptrB* 5′-UTR regulatory sequences can act autonomously regardless of gene context, suggesting that this form of regulation may be widespread in *E. coli*.

RESULTS

***ptrB* 5′-uAUG is a ribosome binding signal.** Our recent study showed that the *ptrB* 5′-uAUG is necessary for efficient downstream *ptrB* CDS expression, although the 5′-uORF is expressed at a low level (23). Due to the use of the *lacZ* reporter gene, any disruption of the ribosomal progression through the *ptrB* uORF sequence would result in loss of *lacZ* expression and may not accurately reflect the initiation frequency at the 5′-uAUG. To examine if low 5′-uORF expression is due to ribosomal stalling, drop-off, or frameshifting, *lacZ* fusions 5, 10, and 15 codons beyond the uORF were constructed (Table 1, pACptrB.lacfuse5, pACptrB.lacfuse10, and pACptrB.lacfuse15). If an impediment in ribosomal movement resulting in ribosomal stalling, drop-off, or a change in frame is the primary cause of the low *lacZ* reporter expression, moving the position of

TABLE 1 DNA sequences of *ptrB* gene fragments, with and without mutations, used in this study

	Sequence[a]	
Plasmid	**Variable 5′-UTR**	**Variable CDS**
pAptrB.WT	ATGTTTCAACCAGAAAGAACAATAAC	ATGCTACCAAAAGCCGCC
pAptrB.5′KO	AT**C**TTTCAACCAGAAAGAACAATAAC	ATGCTACCAAAAGCCGCC
pACptrB.lacfuse5	ATGTTTCAACCAGAA	
pACptrB.lacfuse10	ATGTTTCAACCAGAAAGAACAATAAC	ATGC
pACptrB.lacfuse15	ATGTTTCAACCAGAAAGAACAATAAC	ATGCTACCAAAAGCCGCCC
pAptrB.prestop1	ATGTTTCAACCAGAA***T***GAACAATAAC	ATGCTACCAAAAGCCGCC
pAptrB.prestop2	ATGTTTCAACCAGAAAGAACAATAAC	ATGCTAC***T****A*AAAGCCGCC
pAptrB.SDmut	ATGTTTCAACCA**CTTTC**AACAATAAC	ATGCTACCAAAAGCCGCC
pAptrB.SDstr	ATGTTTCAACCAG**G**A**G**GAACAATAAC	ATGCTACCAAAAGCCGCC
pAptrB.5′KO.SDstr	AT**C**TTTCAACCAG**G**A**G**GAACAATAAC	ATGCTACCAAAAGCCGCC
pAptrB.5′KO.SDmut	AT**C**TTTCAACCA**CTTTC**AACAATAAC	ATGCTACCAAAAGCCGCC
pAptrB.scan1	ATG**AAAGT**ACCAGAAAGAACAATAAC	ATGCTACCAAAAGCCGCC
pAptrB.scan2	ATGTT**AGTTG**CAGAAAGAACAATAAC	ATGCTACCAAAAGCCGCC
pAptrB.scan3	ATGTTTCA**TGGT**CAAAGAACAATAAC	ATGCTACCAAAAGCCGCC
pAptrB.scan4	ATGTTTCAAC**GT**CCCAGAACAATAAC	ATGCTACCAAAAGCCGCC
pAptrB.scan5	ATGTTTCAACCA**CCCCC**AACAATAAC	ATGCTACCAAAAGCCGCC
pAptrB.scan6	ATGTTTCAACCAGAAAG**TTGTT**TAAC	ATGCTACCAAAAGCCGCC
pAptrB.scan7	ATGTTTCAACCAGAAAGAAC**TTATT**C	ATGCTACCAAAAGCCGCC
pABptrB-pnp	ATGTTTCAACCAGAAAGAACAATAAC	TTGCTTAATCCGATCGTT
pABptrB-pncB	ATGTTTCAACCAGAAAGAACAATAAC	ATGACACAATTCGCTTCT
pABptrB5′KO-pnp	AT**C**TTTCAACCAGAAAGAACAATAAC	TTGCTTAATCCGATCGTT
pABptrB5′KO-pncB	AT**C**TTTCAACCAGAAAGAACAATAAC	ATGACACAATTCGCTTCT
pABptrB-tna.IN	ATGTTTCAACCAGAAAGAACAATAAC	ATGGTGGCGTTCTCTAAC
pABptrB-aroL.IN	ATGTTTCAACCAGAAAGAACAATAAC	ATGACGGTCGCGGAGATC
pABptrB5′KO-tna.IN	AT**C**TTTCAACCAGAAAGAACAATAAC	ATGGTGGCGTTCTCTAAC
pABptrB5′KO-aroL.IN	AT**C**TTTCAACCAGAAAGAACAATAAC	ATGACGGTCGCGGAGATC
pABptrB-pncBdblmut	ATGTTTCAACCAGAAAGAACAATAAC	ATGAC**G**CA**G**TTCGCTTCT
pABptrB5′KO-pncBdblmut	AT**C**TTTCAACCAGAAAGAACAATAAC	ATGAC**G**CA**G**TTCGCTTCT
pAptrBCDSdblmut	ATGTTTCAACCAGAAAGAACAATAAC	ATGCT**G**CC**G**AAAGCCGCC
pAptrB5′KO.CDSdblmut	AT**C**TTTCAACCAGAAAGAACAATAAC	ATGCT**G**CC**G**AAAGCCGCC
pAptrB.9ntadd	ATG**TTTCAACCA**TTTCAACCAGAAAGAACAATAAC	ATGCTACCAAAAGCCGCC
pAptrB.21ntadd	ATG**TTTCAACCACTTTCAACAATAACA**TTTCAACCAGAAAGAACAATAAC	ATGCTACCAAAAGCCGCC
pAptrB.33ntadd	ATG**TTTCAACCACTTTTTCAACCACTTTCAACAATAACA**TTTCAACCAGAAAGAACAATAAC	ATGCTACCAAAAGCCGCC

[a]The predicted SD sequence is underlined, the mutated nucleotides are shown in bold, and the stop codons are italicized. The pA series plasmid constructs contain the *ptrB* coding sequence in frame with *lacZ*, the pAB series constructs contain the CDSs of different genes in frame with *lacZ*, and the pAC series constructs contain the 5′-uORF in frame with *lacZ*.

the fusion upstream would result in an increase in reporter expression. However, these fusions also resulted in similarly low levels of expression (see Fig. S1 in the supplemental material), suggesting that the ribosome does not efficiently translate any region of the uORF.

The 5′-uAUG-defined uORF overlaps the downstream *ptrB* CDS reading frame, such that the uORF stop codon is at the +35 position (with the A of the *ptrB* CDS start codon at +1) (Fig. 1). Due to the overlap of the open reading frames, we wanted to rule out the possibility of translational coupling in *ptrB*. The stop codon of the uORF was mutated, allowing for ribosomal read-through and therefore disruption of potential coupling. Mutation of the uORF stop codon resulted in no discernible change in *ptrB* CDS expression (Fig. S2), suggesting that this regulation is not the result of translational coupling. Additional stop codons in frame with the 5′-uAUG were also introduced upstream of the natural uORF stop codon (Table 1, pAptrB.prestop1 and pAptrB.prestop2) to examine the influences of ribosomal progression on the transcript. The additional stop codons also had minimal effects on *ptrB* CDS expression (Fig. S2),

+1 +35

5′ **atg**tttcaaccagaaagaacaataacATGCTACCAAAAGCCGCCCGCATTCCCCACGCCA**TGA**CGCTTCATGGC GTCGAC…3′

FIG 1 The *ptrB* gene sequence. The 5′-ATG is shown in bold at the 5′ terminus. The 5′-UTR is indicated by lowercase letters. The predicted SD sequence is underlined. The *ptrB* CDS is indicated by uppercase letters. The 5′-uAUG-defined ORF's stop codon is within the *ptrB* coding sequence and is also shown in bold (TGA). The SalI site used to construct the in-frame *lacZ* fusion is indicated by the dotted underline.

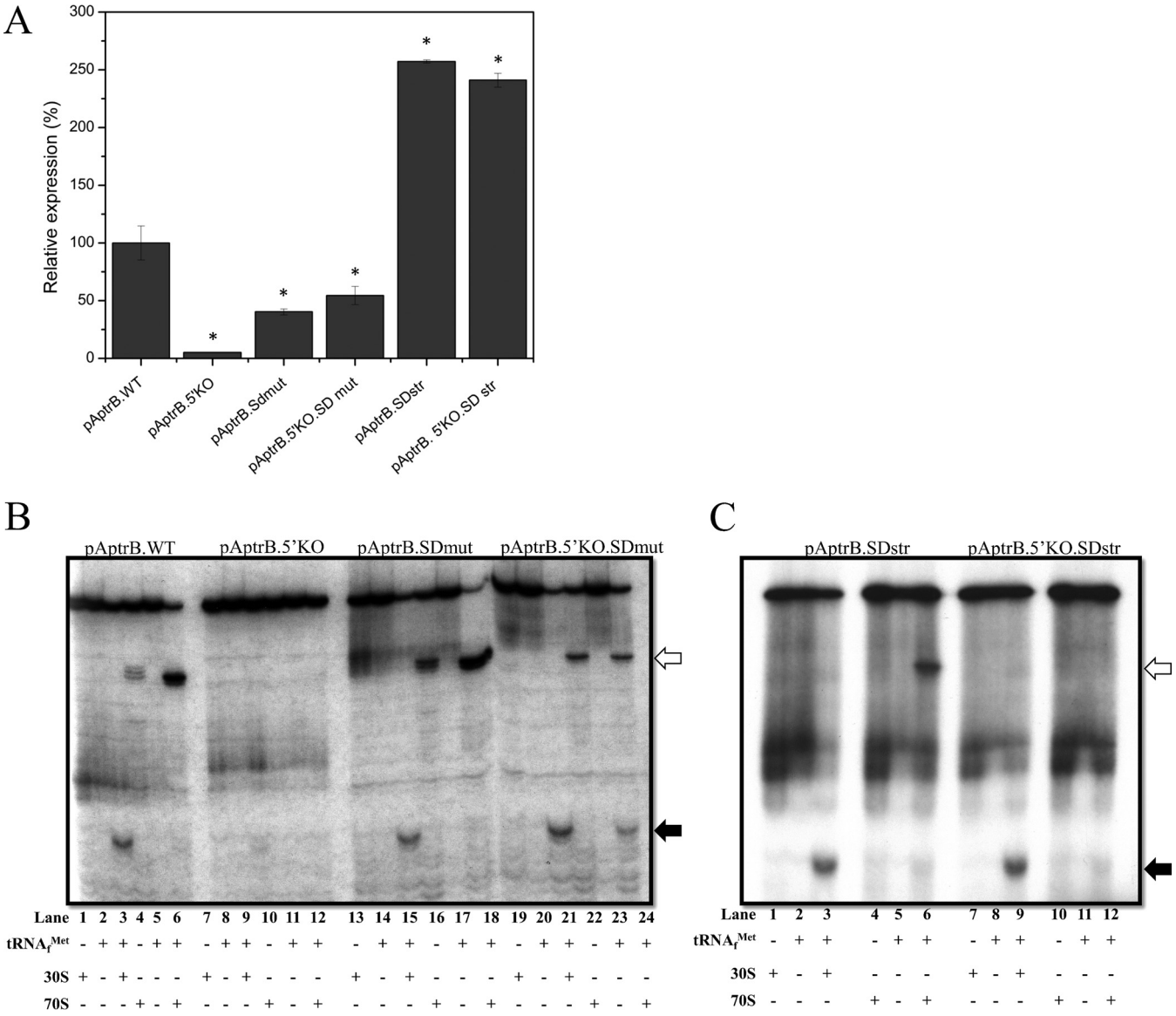

FIG 2 Role of the SD sequence in *ptrB* regulation. (A) Expression from the *ptrB* CDS fused to *lacZ* in the presence of the various SD sequence mutations (Table 1), shown as percentages of the *ptrB* wild-type (WT) CDS (pAptrB.WT) fused to *lacZ* (100%; 5,700 Miller units). *, statistically significant difference from pAptrB.WT ($P < 0.001$). (B) Primer extension inhibition (toeprint) reaction mixtures containing mRNA, 30S subunits, or 70S ribosomes and initiator tRNA, as indicated by the + or – symbols. Predicted positions of toeprint signals (+15) to the 5′-uAUG (open arrow) and downstream AUG (closed arrow) are indicated. (C) Toeprint assay with *ptrB* mRNA containing strengthened SD sequence mutations (Table 1). Reaction components are indicated.

supporting the conclusion that translational coupling does not contribute to *ptrB* regulation. This also reinforces the notion that it is the 5′-uAUG itself that is necessary for downstream *ptrB* CDS expression rather than uORF translation.

Predicted SD sequence suboptimal in controlling *ptrB* expression. The 26-nucleotide-long *ptrB* 5′-UTR contains a predicted SD sequence located 9 nucleotides upstream of the initiation codon with imperfect complementarity to the anti-SD sequence (Fig. 1). In the presence of the 5′-uAUG mutation, expression of the downstream *ptrB* CDS was drastically reduced despite the fact that the predicted SD sequence was still intact (23). To address the SD sequence's role in *ptrB* CDS expression, the SD sequence was mutated to its complement (5′-CTTTC-3′) to disrupt its complementarity with the 16S rRNA anti-SD sequence. This change significantly reduced *ptrB* CDS expression by approximately 60% (Fig. 2A). However, the mutation of an SD sequence in a canonical *E. coli* mRNA would typically result in complete loss of

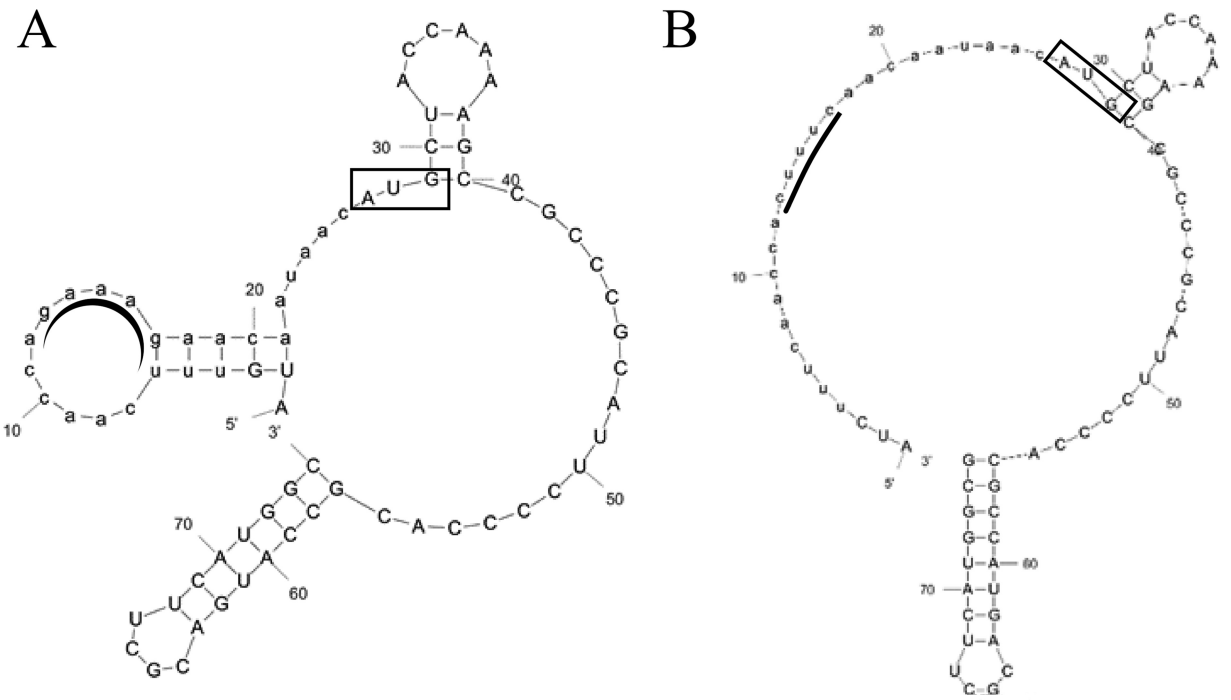

FIG 3 The *ptrB* mRNA structure may play a secondary role in expression. The predicted secondary structure of *ptrB* WT mRNA (A) and *ptrB* mRNA with the 5′-uAUG mutated to AUC and the SD sequence mutated to its complement (pAptrB.5′KO.SDmut [Table 1]) (B) via computational modeling (31). The predicted SD sequence is underlined, and the *ptrB* CDS initiation codon is boxed.

expression, because the SD sequence is the primary ribosome binding signal (28–30). Also, in agreement with the expression observed with the *ptrB* SD sequence mutant (Table 1, pAptrB.SDmut), a 30S subunit binding signal was found to be present at the internal start codon in the primer extension inhibition assay (toeprint assay) (Fig. 2B, lane 15).

Next, the SD sequence was strengthened such that its sequence had absolute complementarity to the anti-SD sequence (i.e., 5′-GAAAG-3′ → 5′-GGAGG-3′). The mutation increased the downstream expression level by 2.5-fold (Fig. 2A). To investigate the relationship between the 5′-uAUG and the SD sequence, the 5′-uAUG was mutated to 5′-AUC in concert with the strengthened SD sequence (Table 1, pAptrB.5′KO.SDstr). No change in *ptrB* CDS expression was observed as a result of the 5′-uAUG mutation in the presence of the strengthened SD sequence (Fig. 2A). The loss of 5′-uAUG dependence as a result of the strengthened SD sequence was supported by the findings from the toeprint assay, which displayed a loss of 70S binding to the 5′ terminus due to the 5′-uAUG mutation but maintained a strong internal 30S binding signal (Fig. 2C).

The 5′-uAUG and SD sequence appear to act through distinct regulatory mechanisms; therefore, we tested whether *ptrB* expression would be completely abolished by disrupting both signals. In the presence of both the 5′-uAUG → AUC mutation and the SD sequence mutated to its complement (Table 1, pAptrB.5′KOSDmut), *ptrB* CDS expression was reduced by 45% (Fig. 2A). This finding was supported by the ribosome binding data as well, which exhibited a band corresponding to 30S subunit binding to the *ptrB* CDS start codon in the presence of the tandem mutation (Fig. 2B).

Due to this unexpected result, we employed secondary structure prediction software (31) to analyze changes in the mRNA secondary structure that resulted from the sequence mutations. The wild-type secondary structure includes a hairpin loop within the 5′-UTR (Fig. 3A), and mutation of the 5′-uAUG also exhibits a secondary structure in the 5′-UTR. However, both SD sequence mutants (pAptrB.SDmut and pAptrB.5′KO.SDmut) (Table 1) resulted in an open conformation with no steric hindrance to prevent

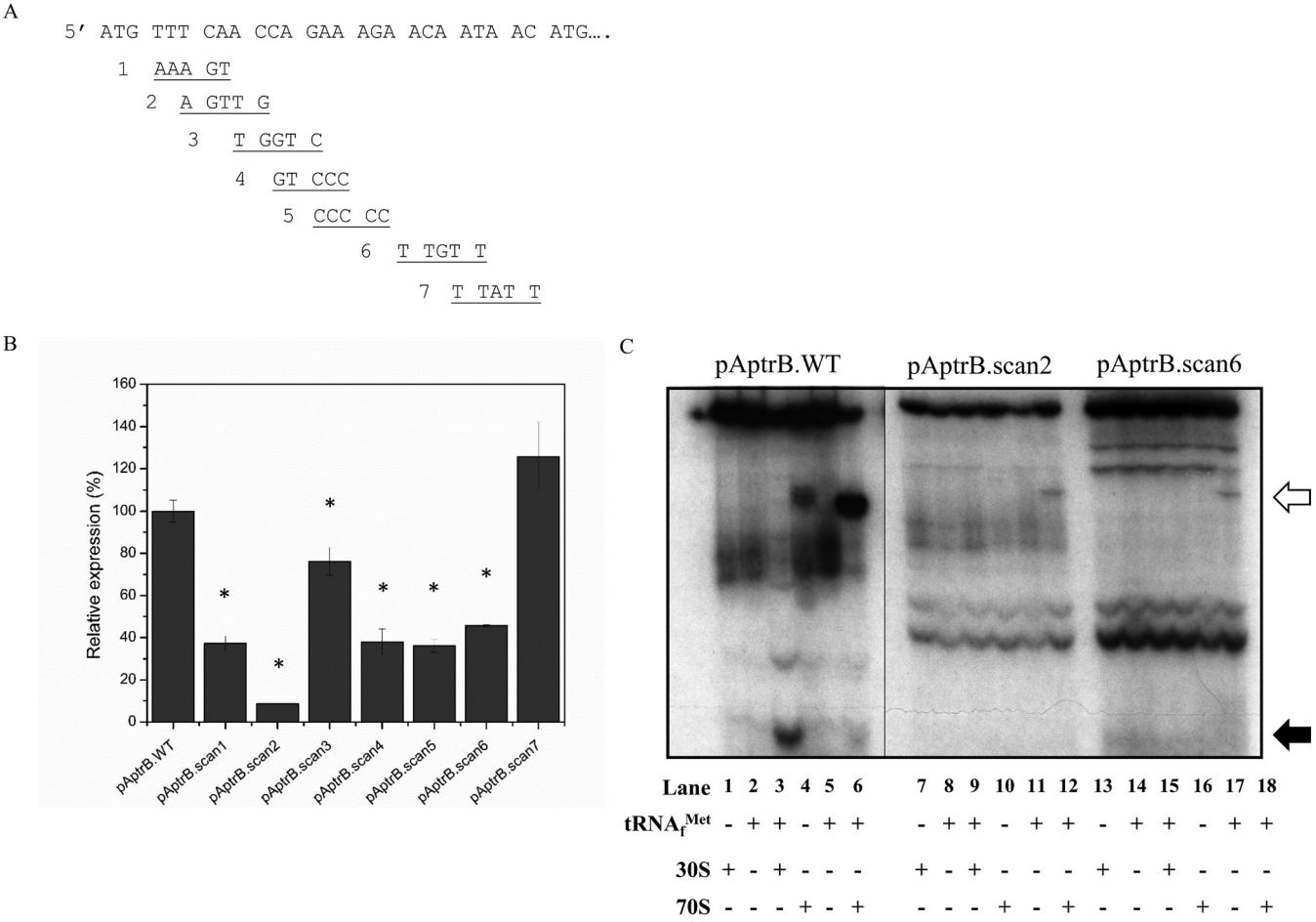

FIG 4 Scanning mutagenesis of the *ptrB* 5′-UTR. (A) Schematic of regions within the 5′-UTR with mutations of the DNA sequence, corresponding to lines 1 to 7 beneath the *ptrB* sequence. (B) Expression from the *ptrB* CDS fused to *lacZ* in the presence of the various 5′-UTR mutations (numbers 1 to 7), shown as a percentage of the *ptrB* WT CDS (pAptrB.WT) fused to *lacZ* (100%; 11,000 Miller units). *, statistically significant difference compared to pAptrB.WT ($P < 0.05$). (C) Toeprint assay results for pAptrB.WT mRNA compared to pAptrB.scan2 and pAptrB.scan6 mRNAs. The image is a composite of one gel, with unrelated lanes removed. Reaction components are indicated. Predicted positions of toeprint signals (+15) to the 5′-uAUG (open arrow) and downstream AUG (closed arrow) are indicated.

ribosome binding (Fig. 3B). The absence of a secondary structure is sufficient to initiate translation in mRNAs lacking an SD sequence (32); therefore, the complete loss of secondary structure in the 5′-UTR of the SD mutants could allow for SD sequence-independent initiation.

Signals present within the *ptrB* 5′-UTR influence *ptrB* CDS expression. To examine whether there are other regions of importance within the 5′-UTR, scanning mutagenesis was conducted and spanned the length of the *ptrB* 5′-UTR with overlapping mutations (Fig. 4A). The mutations changed the wild-type sequence to its complement, except in the case of the SD region, in which the nucleotides were all mutated to cytosine. The majority of the mutations significantly disrupted *ptrB* CDS expression (Fig. 4B). Mutations within the region of ribosomal coverage when bound to the 5′ terminus of the mRNA, overlapping to approximately position +19 (33), caused a severe decrease in *ptrB* CDS expression (Fig. 4B). The mutation out of the range of ribosomal coverage resulted in no change in *ptrB* CDS expression (Fig. 4B). A toeprint assay was also conducted to analyze the ribosome binding pattern in two of the scanning mutant constructs (pAptrB.scan2 and pAptrB.scan6) (Table 1). In both mutants, 70S binding to the 5′ terminus and internal 30S binding were greatly reduced compared to the *ptrB* wild type, even in the presence of an intact 5′-uAUG (Fig. 4C). These data suggest that there are signals within the 5′-UTR that impact ribosome binding to the mRNA 5′ terminus and influence *ptrB* CDS expression.

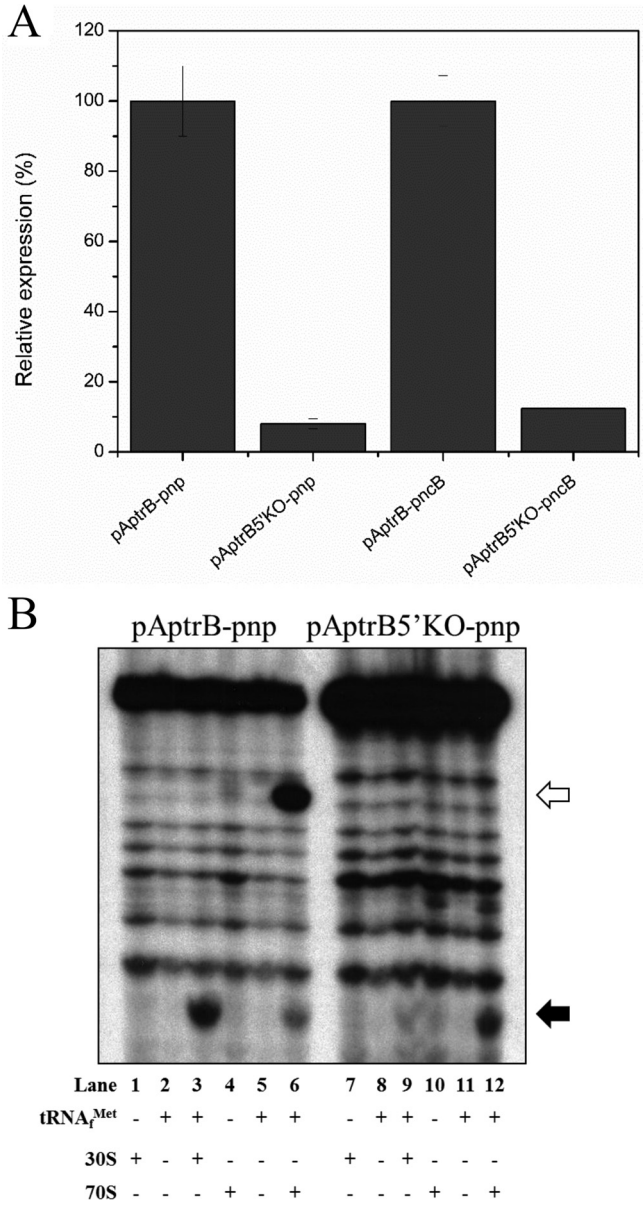

FIG 5 Heterologous regulation by the 5′-UTR of *ptrB*. (A) LacZ activity expressed from the *ptrB* 5′-UTR (5′-uAUG or AUG → AUC knockout [KO]) fusions to *pnp* or *pncB* coding sequences fused to *lacZ* (Table 1) (100%; 25,000 Miller units and 8,500 Miller units, respectively). (B) Toeprint reactions with mRNA containing the *ptrB* 5′-UTR (5′-uAUG or AUG → AUC KO) fused to the *pnp* coding sequence. Reaction components are indicated. Predicted positions of toeprint signals (+15) to the 5′-uAUG (open arrow) and downstream *ptrB* CDS AUG (closed arrow) are indicated.

Ability of *ptrB* 5′-UTR regulation to be transplanted. There appear to be multiple regions of importance contained within the *ptrB* 5′-UTR that work cooperatively to influence downstream *ptrB* CDS expression. To determine if the *ptrB* 5′-UTR could similarly influence other genes' CDSs, gene fusions were constructed linking the *ptrB* 5′-UTR to two other the ORFs of two other genes, *pncB* and *pnp* (pABptrB-pnp and pABptrB-pncB) (Table 1). To examine the 5′-uAUG regulatory effect on the other genes, the *ptrB* 5′-uAUG was again mutated to AUC to disrupt ribosome binding and 5′-uORF translation. The change caused the expression levels of both *pncB* and *pnp* CDSs to be significantly reduced by more than 90% (Fig. 5A), following the trend that was seen previously for the *ptrB* CDS. The ribosome binding pattern of the *ptrB-pnp* fusion supported the expression data, with the 5′-uAUG → AUC mutation resulting in the loss

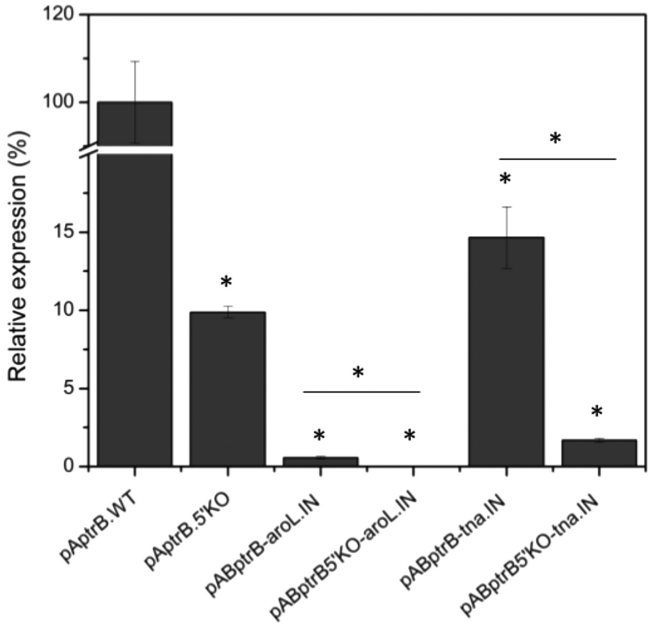

FIG 6 The *ptrB* 5′-UTR is insufficient to stimulate expression of internal RNA fragments. Expression levels of internal RNA fragments (*aroL*.IN/*tna*.IN) fused to *lacZ* with or without a *ptrB* 5′-uAUG mutation (Table 1). All results were compared to the *ptrB* WT CDS (pAptrB.WT) fused to *lacZ* (100%; 11,600 Miller units). *, statistically significant difference compared to pAptrB.WT ($P < 0.001$). A horizontal line above two bars denotes a statistical significance between those two constructs ($P < 0.001$).

of binding signals at both the 5′ terminus and the internal start codon (Fig. 5B). Therefore, the *ptrB* 5′-UTR regulatory signals can exert their effects on other *E. coli* genes when transferred upstream of their CDSs. Interestingly, although a loss of 70S binding at the 5′ terminal was observed in the pABptrB5′KO-pnp mutant (Table 1), an internal 70S toeprint was present (Fig. 5A, lane 12). Internal 70S binding has been reported in previous studies (23); however, the mechanism and cause for this binding are unclear.

Initiation signals present within a bona fide CDS necessary for efficient translation. We next sought to determine if the addition of the *ptrB* 5′-UTR would allow for translation to initiate from an internal fragment of RNA. Presumably, an internal fragment would be devoid of any translation initiation signals but would still have the ability to be translated due to its coding potential. Internal fragments of two *E. coli* genes, *tna* and *aroL*, were used. An internal in-frame methionine, 381 nucleotides and 126 nucleotides from the native *tna* and *aroL* initiation codons, respectively, was selected to act as an initiation codon. Each truncated ORF was then fused to the *ptrB* 5′-UTR (pABptrB-tna.IN and pABptrB-aroL.IN) (Table 1). This fusion allowed the *ptrB* 5′-UTR to regulate expression of the downstream ORF to produce a truncated version of the Tna or AroL protein, respectively.

The *ptrB* 5′-UTR was not able to confer efficient expression of the internal *tna* or *aroL* ORFs (Fig. 6). Each fusion produced relatively low levels of expression. Interestingly, in both the *tna* and *aroL* fusions, when the *ptrB* 5′-uAUG was mutated to AUC, it produced a similar trend to that seen with the bona fide coding sequences of *ptrB*, *pncB*, and *pnp* (Fig. 6). In each case, the 5′-uAUG → AUC mutation caused a drastic reduction in expression of approximately 90% (Fig. 2A, 5A, and 6). Therefore, there are signals within translation initiation regions of bona fide coding sequences that are important for translation, and the *ptrB* 5′-UTR is not able to confer efficient expression independently. However, the 5′-uAUG still acts as a regulatory signal, even for low levels of translation.

Nucleotides within bona fide CDSs contribute to expression. To further examine the role of CDS regions, we utilized *pncB*, which relies on two nucleotides within its CDS for expression (34). Wild-type *pncB* has a weak SD sequence, does not possess a

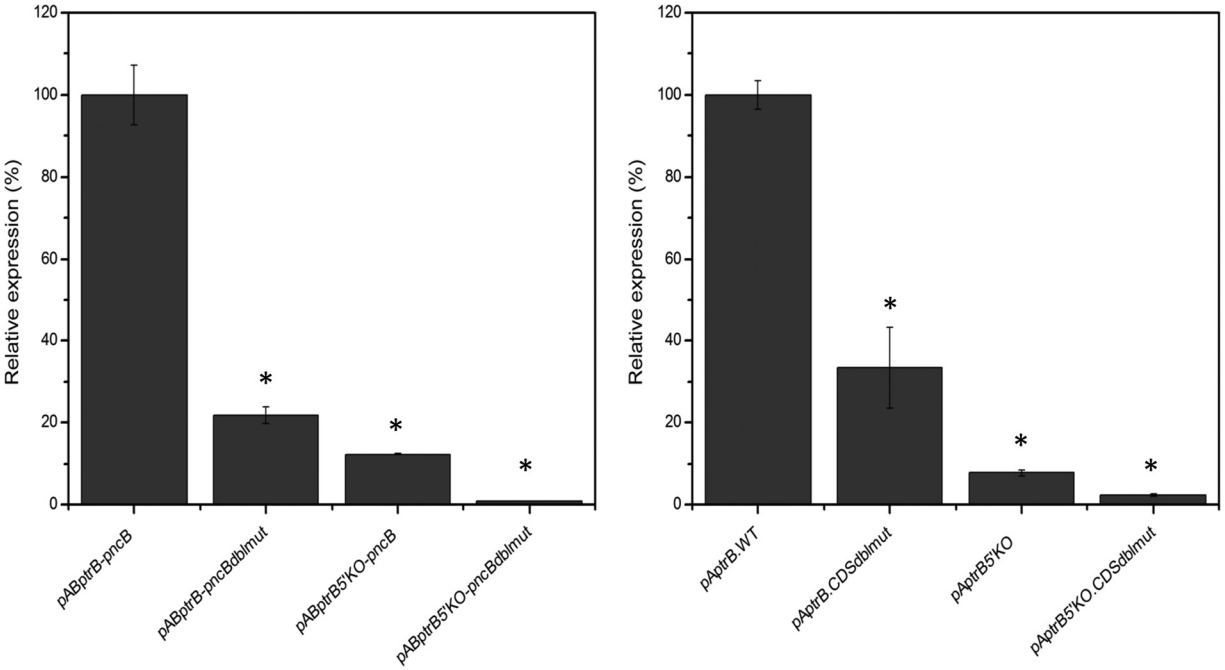

FIG 7 Expression is regulated by both upstream and downstream signals. (Left) Expression values, presented as percentages of the *ptrB* 5′-UTR fused to the *pncB* CDS, with the *pncB* CDS in frame with *lacZ* (100%; 8,500 Miller units). (Right) Expression values, presented as percentages of the *ptrB* WT CDS in frame with *lacZ* (100%; 7,000 Miller units). *, statistically significant difference compared to pABptrB-pncB or pAptrB.WT, respectively ($P < 0.001$).

5′-uAUG, and is not influenced by its native 5′-UTR for expression. Instead, two nucleotides, both adenines, at positions +6 and +9 of the *pncB* CDS appear to influence translation. When these two adenines are mutated to guanines, *pncB* expression is completely abolished (34). A construct was made that replaced the *pncB* 5′-UTR with the *ptrB* 5′-UTR (pABptrB-pncB) (Table 1) to determine if the positive effects of the *ptrB* 5′-UTR could overcome the negative impact of the *pncB* CDS adenine mutations. In the presence of the *ptrB* 5′-UTR, the *pncB* CDS mutations (pABptrB-pncBdblmut) (Table 1) caused *pncB* expression to be significantly reduced by approximately 80% (Fig. 7, left) rather than completely abolished, suggesting that the signals within the *ptrB* 5′-UTR partially rescued the negative impact of the *pncB* CDS mutations on expression. The *ptrB* 5′-uAUG-to-AUC mutation resulted in an even more dramatic reduction in *pncB* CDS expression than the *pncB* CDS mutations in the presence of the *ptrB* 5′-UTR. In this case, expression was reduced by approximately 88% (Fig. 7, left), which follows the trends previously seen with *ptrB* 5′-uAUG mutations (Fig. 2A). These data indicate that, in this background, the 5′-uAUG has more of an impact on *pncB* CDS expression than the influential regions within the *pncB* CDS itself. Finally, when the *ptrB* 5′-uAUG and the two nucleotides within the *pncB* CDS were mutated in tandem (pABptrB5′KO-pncBdblmut) (Table 1), *pncB* CDS expression was further reduced by >99% (Fig. 7, left), supporting the notion that the 5′-uAUG and the two nucleotides within the CDS are the signals that control *pncB* CDS expression in this context.

Interestingly, the *ptrB* CDS also has adenines at positions +6 and +9 (with the A of the *ptrB* CDS start codon at +1). When both adenines were mutated to guanines in *ptrB*, CDS expression was reduced by approximately 70% (Fig. 7, right). In addition, when the 5′-uAUG and the two nucleotides within the *ptrB* CDS were mutated in tandem, expression was reduced by >98% (Fig. 7, right). These results are concordant with the expression data derived from similar mutations in the *ptrB-pncB* construct (Fig. 7, left). Taken together, these data suggest that specific upstream and downstream sequence elements are responsible for *ptrB* CDS expression in this novel regulation mechanism that compensates for the weak SD sequence.

DISCUSSION

Our results suggest that translation initiation of *E. coli ptrB* mRNA is controlled via its 5′-terminal AUG which, in cooperation with the *ptrB* 5′-UTR, represents a novel form of regulation that can act as an autonomous control element independent of gene context. We propose a model to explain the mechanism of *ptrB* regulation, in which the 5′-uAUG acts as a recognition signal and binding platform to attract ribosomes to the mRNA. The 5′-uAUG increases the local concentration of ribosomes to allow for more efficient binding and increased *ptrB* CDS initiation. Ribosomes initially load at the 5′ terminus, with sequences in the 5′-UTR stabilizing the ribosome until it can transition down the mRNA. Additional nucleotides within the *ptrB* CDS then help to compensate for the weak SD sequence to position the ribosome on the *ptrB* CDS initiation codon, thereby defining the reading frame. These upstream and downstream sequences function cooperatively as a regulatory mechanism, utilizing previously established recognition signals in a novel combination. This model is in agreement with the previously described cumulative specificity initiation mechanism hypothesis of Nakamoto (35), in which multiple preferred upstream and downstream bases work cumulatively, but independently of each other, to regulate translation initiation.

Evidence presented here supports the model, because the 5′-uAUG is required for efficient downstream translation and acts in a novel fashion as a binding/recognition signal rather than acting as an initiation codon. Instead, uORF translation appears to be repressed (data not shown) to allow *ptrB* to take advantage of the 5′-terminal AUG, which strongly binds ribosomes without the added energy requirements of uORF translation. Therefore, the *ptrB* CDS is regulated independently of uORF translation, which is distinct from previously studied uORF regulation mechanisms (16–22). Binding of the ribosome to the 5′-uAUG without initiation of translation is also supported by the presence of a tRNA-independent ribosome binding signal in the *ptrB* toeprints (Fig. 2B, lanes 4, 16, and 22). This indicates that ribosomes can stably bind the 5′ terminus without the formation of a ternary complex and suggests that ribosomes do not proceed into elongation due to the absence of initiator tRNA binding.

Since a local single-stranded region is needed for efficient ribosome binding (5–8), the 5′-uAUG might be necessary to counteract a secondary structure within the 5′-UTR. Binding of the ribosome at the 5′ terminus could resolve the secondary structure at the RBS due to the short nature of the *ptrB* 5′-UTR. It is unclear how the 70S ribosome is loaded onto the 5′ terminus of leaderless mRNA; however, it is thought that the 5′ end is pulled through the channel between the 30S and 50S subunits (36) until it is positioned in the P-site through interactions with the initiator tRNA (26). The tunnel is only, on average, 15 Å in diameter (37, 38), so it cannot accommodate a secondary structure. This suggests that the structure is eliminated as the mRNA is fed through the tunnel, such that the region of mRNA protected by the ribosome becomes single stranded. Since the *ptrB* 5′-UTR is only 26 nucleotides long, the loaded ribosome may cover the majority of the 5′-UTR, thereby eliminating a secondary structure.

Additionally, single-stranded regions upstream of the RBS have been shown to act as standby sites to allow for temporary ribosome binding until the inhibitory secondary structure opens (9, 39, 40), with subsequent scanning-like movement downstream (17, 39, 41). Energy-independent 70S scanning initiation has been demonstrated recently, with an emphasis on the requirement of an SD sequence for proper positioning (18). In the case of *ptrB*, the 5′-uAUG may act as a standby-like site for initial ribosomal loading, after which the ribosome can then scan down the mRNA using additional sequences for alignment to compensate for the weak SD sequence. Evidence presented here suggests that the ribosome does initially bind to the 5′ terminus, because nucleotides within the region of ribosomal occupancy (+1 to +19) (33) all impact *ptrB* CDS expression, whereas the nucleotides outside the region do not (Fig. 4B). In support of the scanning hypothesis, increasing the distance that the ribosome must travel leads to a decrease in *ptrB* CDS initiation efficiency. This was demonstrated by adding 9, 21, and 33 nucleotides of intervening sequence to the 5′-UTR, immediately following the 5′-uAUG

(pAptrB.9ntadd, pAptrB.21ntadd, and pAptrB.33ntadd) (Table 1). The addition of 9 nucleotides had no impact on downstream expression. However, as the region between the 5′-uAUG and the *ptrB* CDS initiation codon was extended, *ptrB* CDS expression was reduced in a linear fashion by 25% and 57%, respectively (Fig. S3). The same effect occurs in translation reinitiation, in which longer intercistronic gaps lead to reduced downstream initiation in operons due to ribosome drop-off (42).

This 5′-uAUG regulation mechanism has the ability not only to influence *ptrB* CDS expression but also to regulate other *E. coli* CDSs, independent of the gene context (Fig. 5). However, the inefficient use of internal RNA fragments (Fig. 6) suggests that signals within a genuine CDS must nonetheless be present to direct the ribosome downstream for start site selection. Our results demonstrate a positive correlation between the presence of adenines in *ptrB* and *pncB* and an increased level of expression (Fig. 7). Adenine-rich regions have been shown to enhance expression by increasing the rate and amount of ternary complex formation during initiation (34), likely due to ribosomal protein S1's high affinity for polypyrimidines (43, 44). S1-mediated initiation (45, 46) occurs in the absence of an SD sequence; therefore, it is possible that S1 also plays a role in *ptrB* regulation.

Overall, these data suggest novel regulation of translation initiation that is controlled by a 5′-uAUG and which varies from the canonical SD sequence mechanism. There are numerous examples of translation events occurring in the absence of an SD sequence (reference 47 and references therein), with bioinformatic analyses revealing the absence of a canonical SD sequence in a significant percentage of genes across prokaryotic genomes (48). One study suggested that approximately one-half of prokaryotic mRNAs lack an SD sequence altogether (49), demonstrating the importance of understanding mechanisms of translation in which the SD sequence is not the primary ribosome binding signal. Our previous study demonstrated the abundance of 5′-uAUGs within *E. coli* transcripts (23); therefore, it is possible that this mechanism of regulation is widespread in *E. coli* and may contribute to regulation of mRNA with weak or absent SD sequences.

Furthering our understanding of noncanonical translational regulation mechanisms will be beneficial in synthetic biology and genetic engineering to finely tune protein production. Other studies exploring optimization of recombinant gene expression have discovered upstream and downstream sequence elements that work cooperatively to influence expression; however, those elements identified worked in a gene-dependent manner (50). In contrast, the sequence elements identified in this study function in a gene-independent manner, suggesting that they may be more easily adapted to control gene expression.

MATERIALS AND METHODS

Bacterial strains and reagents. *E. coli* DH5α (New England BioLabs [NEB]) was used as the host for all plasmid DNA manipulations. *E. coli* RFS859 (F^- *thr-1 araC859 leuB6* Δ*lac74 tsx-274 gyrA11* λ^- *recA11 relA1 thi-1*) (51) was used as the host for the expression and assay of *lacZ* fusion mRNA constructs. Chromosomal DNA from *E. coli* K-12 was used as a template for initial PCR amplifications of the *ptrB* gene. Two plasmids were used in the study, pM1108 (for the incorporation of a mutated *lac* promoter) and pA906 (to allow for the *lacZ* gene fusion) (52).

Radiolabeled [γ-^{32}P]ATP (6,000 Ci/mmol, 150 mCi/ml) was purchased from PerkinElmer. Restriction endonucleases, T4 DNA ligase, T4 polynucleotide kinase (PNK), and T7 RNA polymerase were purchased from NEB and used according to the manufacturer's instructions. RNase-free DNase I (Roche), avian myeloblastosis virus reverse transcriptase (Life Sciences), and *Pfu* DNA polymerase (Stratagene) were used according to the manufacturers' specifications. DNA oligonucleotides were purchased (IDT).

Recombinant DNA procedures. PCR amplifications were performed using specific oligonucleotides (IDT) to incorporate specific mutations into *E. coli* K-12 chromosomal DNA or purified plasmids containing the gene of interest. Purified DNA amplicons were then digested with restriction endonucleases and ligated into plasmid pA906 containing an ampicillin resistance marker and the mutated *lac* promoter (AATAAT) to fuse the regions of interest in frame with the fifth codon of the *lacZ* reporter gene.

Gene fusions. Fusions of regions from two separate genes (Table 1) were constructed using trimolecular ligations. The primers at the amplicons' junction were phosphorylated with T4 PNK, T4 PNK buffer (2 M Tris-HCl [pH 7.6], 1 M $MgCl_2$, 1 M dithiothreitol [DTT]) and ATP (25 mM) for 30 min at 37°C. The reaction mixtures were then heat inactivated for 20 min at 65°C. The regions within the genes used for the fusions were amplified using PCR with the phosphorylated primers, purified, and digested to

allow for incorporation into pM1108. The two amplicons and the plasmid were then ligated using T7 ligase and 10× ligation buffer (NEB), and mixtures were incubated at 16°C overnight. The ligation product was then used as a template for a subsequent PCR amplification with a 5′ primer specific for the plasmid and a 3′ primer specific to the downstream amplicon to ensure proper ligation. The resulting amplicon was then purified and subcloned in pA906, as previously described for constructing *lacZ* fusions.

β-Galactosidase assay. β-Galactosidase assays were performed as previously described (53).

***In vitro* synthesis of RNA.** The cloned plasmids were used as the templates in PCR amplifications by utilizing a primer to incorporate the T7 RNA polymerase promoter sequence (5′-TAATACGACTCACTAT AG-3′). This produced DNA fragments containing a T7 promoter sequence, allowing for *in vitro* transcription with T7 RNA polymerase and production of RNA used for toeprint reactions. RNAs were synthesized and purified as described previously (54). RNAs used in toeprint assays were synthesized by combining purified PCR amplicons (constructs with *lacZ* fusions containing a T7 promoter) and T7 RNA polymerase in 1× buffer (40 mM Tris [pH 7.8], 25 mM $MgCl_2$, 1 mM spermidine, 0.01% Triton X-100, 5 mM each nucleoside triphosphate, and 30 mM DTT). Transcription reaction mixtures were incubated for approximately 4 h at 37°C, and 40 mM EDTA was added. Samples were treated with DNase (Roche) for 15 min at 37°C. RNA was ethanol precipitated and suspended in RNA loading dye (50% formamide, 0.05% bromophenol blue, 0.05% xylene cyanol). Samples were subjected to PAGE (6% acrylamide, 7 M urea), and full-length products were excised using UV shadowing. Gel slices were incubated overnight at room temperature in elution buffer (300 mM NaO-acetate [NaOAc; pH 5.2], 0.1% SDS, 1 mM EDTA) with gentle rocking. The supernatant was phenol extracted and ethanol precipitated.

Primer extension inhibition (toeprint) assay. DNA oligonucleotides were phosphorylated at the 5′ terminus with [γ-^{32}P]ATP (6,000 Ci/mmol, 150 mCi/ml; PerkinElmer) and T4 PNK in 1× kinase buffer for 30 min at 37°C and annealed to the 3′ termini of RNA as previously described (55). Annealed RNA was incubated with 30S subunits or 70S ribosomes with or without $tRNA^{fMet}$ for 15 min at 37°C. Reaction mixtures were transferred to ice, and reverse transcriptase was added. The reaction mixtures were incubated for 15 min at 37°C, and reactions were stopped by the addition of 0.3 M NaOAc and 100% ethanol and precipitated overnight at −80°C. Precipitated complexes were collected by centrifugation and dissolved in loading dye (80% deionized formamide, 10 mM NaOH, 1 mM EDTA, 0.5% bromophenol blue, and xylene cyanol), followed by heat treatment (95°C, 5 min) and PAGE (6% acrylamide, 7 M urea) in 1× Tris-borate-EDTA. Gels were visualized via autoradiography.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/JB.00091-17.

SUPPLEMENTAL FILE 1, PDF file, 0.3 MB.

ACKNOWLEDGMENTS

We thank Racheal Desmone and Sarah Steimer for stimulating conversations. We also thank Mitch Balish, Eileen Bridge, D. J. Ferguson, Xiao-Wen Cheng, Jack Vaughn, and Chris Sedlacek for their help in editing the manuscript.

This work was funded by a grant (GM065120) from the National Institutes of Health.

REFERENCES

1. Laursen B, Sorensen H, Mortensen K, Sperlink-Petersen H. 2005. Initiation of protein synthesis in bacteria. Microbiol Mol Biol Rev 6:101–123. https://doi.org/10.1128/MMBR.69.1.101-123.2005.
2. Shine J, Dalgarno L. 1974. The 3′ terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc Natl Acad Sci U S A 71:1342–1346. https://doi.org/10.1073/pnas.71.4.1342.
3. Ringquist S, Shinedling S, Barrick D, Green L, Binkley J, Stormo GD, Gold L. 1992. Translation initiation in *Escherichia coli*: sequence within the ribosome-binding site. Mol Microbiol 6:1219–1229. https://doi.org/10.1111/j.1365-2958.1992.tb01561.x.
4. McCarthy JE, Brimacombe R. 1994. Prokaryotic translation: the interactive pathway leading to initiation. Trend Genet 10:402–407. https://doi.org/10.1016/0168-9525(94)90057-4.
5. Draper DE, Gluick TC, Schlax PJ. 1998. Pseduoknots, RNA folding and translational regulation, p 415–436. *In* Simons RW, Grunberg-Manago M (ed), RNA structure and function, vol 35. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
6. Draper DE. 1987. Translational regulation of ribosomal proteins in *Escherichia coli*, p 1–26. *In* Ilan J (ed), Translational regulation of gene expression. Plenum Press, New York, NY.
7. Gualerzi CO, Pon CL. 1990. Initiation of mRNA translation in prokaryotes. Biochemistry 29:5881–5889. https://doi.org/10.1021/bi00477a001.
8. de Smit MH, van Duin J. 1990. Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. Proc Natl Acad Sci U S A 87:7668–7672. https://doi.org/10.1073/pnas.87.19.7668.
9. de Smit MH, van Duin J. 1994. Translational initiation on structured messengers. Another role for the Shine-Dalgarno interaction. J Mol Biol 235:173–184.
10. Gu W, Zhou T, Wilke CO. 2010. A universal trend of reduced mRNA stability near the translation-initiation site in prokaryotes and eukaryotes. PLoS Comput Biol 6:e1000664. https://doi.org/10.1371/journal.pcbi.1000664.
11. Storz G. 1999. An RNA thermometer. Genes Dev 13:633–636. https://doi.org/10.1101/gad.13.6.633.
12. Narberhaus F, Waldminghaus T, Chowdhury S. 2006. RNA thermometers. FEMS Microbiol Rev 30:3–16. https://doi.org/10.1111/j.1574-6976.2005.004.x.
13. Klinkert B, Narberhaus F. 2009. Microbial thermosensors. Cell Mol Life Sci 66:2661–2676. https://doi.org/10.1007/s00018-009-0041-3.
14. Sarabhai A, Brenner G. 1967. A mutant which reinitiates the polypeptide chain after chain termination. J Mol Biol 27:145–162. https://doi.org/10.1016/0022-2836(67)90357-9.
15. Min Jou W, Haegeman G, Ysebaert M, Fiers W. 1972. Nucleotide sequence of gene coding for the bacteriophage MS2 coat protein. Nature 237:82–88. https://doi.org/10.1038/237082a0.

16. Andre A, Puca A, Sansone F, Brandi A, Antico G, Calogero RA. 2000. Reinitiation of protein synthesis in *Escherichia coli* can be induced by mRNA cis-elements unrelated to canonical translation initiation signals. FEBS Lett 468:73–78. https://doi.org/10.1016/S0014-5793(00)01198-4.
17. Adhin MR, van Duin J. 1990. Scanning model for translational reinitiation in eubacteria. J Mol Biol 213:811–818. https://doi.org/10.1016/S0022-2836(05)80265-7.
18. Yamamoto H, Wittek D, Gupta R, Qin B, Ueda T, Krause R, Yamamoto K, Albrecht R, Pech M, Nierhaus KH. 2016. 70S-scanning initiation is a novel and frequent initiation mode of ribosomal translation in bacteria. Proc Natl Acad Sci U S A 113:E1180–E1189. https://doi.org/10.1073/pnas.1524554113.
19. Lovett PS, Rogers EJ. 1996. Ribosome regulation by the nascent peptide. Microbiol Rev 60:366–385.
20. Tenson T, Ehrenberg M. 2002. Regulatory nascent peptides in the ribosomal tunnel. Cell 108:591–594. https://doi.org/10.1016/S0092-8674(02)00669-4.
21. Nakatogawa H, Ito K. 2002. The ribosomal exit tunnel functions as a discriminating gate. Cell 108:629–636. https://doi.org/10.1016/S0092-8674(02)00649-9.
22. Cruz-Vera LR, Sachs MS, Squires CL, Yanofsky C. 2011. Nascent polypeptide sequences that influence ribosome function. Curr Opin Microbiol 14:160–166. https://doi.org/10.1016/j.mib.2011.01.011.
23. Beck HJ, Fleming IMC, Janssen GR. 2016. 5′-terminal AUGs in *Escherichia coli* mRNAs with Shine-Dalgarno sequences: identification and analysis of their roles in non-canonical translation initiation. PLoS One 11: e0160144. https://doi.org/10.1371/journal.pone.0160144.
24. Balakin AG, Skripkin EA, Shatsky IN, Bogdanov AA. 1992. Unusual ribosome binding properties of mRNA encoding bacteriophage λ repressor. Nucleic Acids Res 20:563–571. https://doi.org/10.1093/nar/20.3.563.
25. Udagawa T, Shimizu Y, Ueda T. 2004. Evidence for the translation initiation of leaderless mRNAs by the intact 70S ribosome without its dissociation into subunits in Eubacteria. J Biol Chem 279:8539–8546. https://doi.org/10.1074/jbc.M308784200.
26. Moll I, Hirokawa G, Kiel MC, Kaji A, Blasi U. 2004. Translation initiation with 70S ribosomes: an alternative pathway for leaderless mRNAs. Nucleic Acids Res 32:3354–3363. https://doi.org/10.1093/nar/gkh663.
27. Kanatani A, Masuda T, Shimoda T, Misoka F, Lin XS, Yoshimoto T, Tsuru D. 1991. Protease II from *Escherichia coli*: sequencing and expression of the enzyme gene and characterization of the expressed enzyme. J Biochem 110:315–320. https://doi.org/10.1093/oxfordjournals.jbchem.a123577.
28. de Boer HA, Hui A, Comstock LJ, Wong E, Vasser M. 1983. Portable Shine-Dalgarno regions: a system for a systematic study of defined alterations of nucleotide sequences within *E. coli* ribosome binding sites. DNA 2:231–235. https://doi.org/10.1089/dna.1983.2.231.
29. Hui A, de Boer HA. 1987. Specialized ribosome system: preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in *Escherichia coli*. Proc Natl Acad Sci U S A 84:4762–4766. https://doi.org/10.1073/pnas.84.14.4762.
30. Van Etten WJ, Janssen GR. 1998. An AUG initiation codon, not codon-anticodon complementarity, is required for the translation of unleadered mRNA in *Escherichia coli*. Mol Microbiol 27:987–1001. https://doi.org/10.1046/j.1365-2958.1998.00744.x.
31. Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31:3406–3415. https://doi.org/10.1093/nar/gkg595.
32. Scharff LB, Childs L, Walther D, Bock R. 2011. Local absence of secondary structure permits translation of mRNAs that lack ribosome-binding sites. PLoS Genet 7:e1002155. https://doi.org/10.1371/journal.pgen.1002155.
33. Huttenhofer A, Noller HF. 1994. Footprinting mRNA-ribosome complexes with chemical probes. EMBO J 13:3892–3901.
34. Brock JE, Paz RL, Cottle P, Janssen GR. 2007. Naturally occurring adenines within mRNA coding sequences affect ribosome binding and expression in *Escherichia coli*. J Bacteriol 189:501–510. https://doi.org/10.1128/JB.01356-06.
35. Nakamoto T. 2011. Mechanisms of the initiation of protein synthesis: in reading frame binding of ribosomes to mRNA. Mol Biol Rep 38:847–855. https://doi.org/10.1007/s11033-010-0176-1.
36. Moll I, Grill S, Gualerzi CO, Blasi U. 2002. Leaderless mRNAs in bacteria: surprises in ribosomal recruitment and translational control. Mol Microbiol 43:239–246. https://doi.org/10.1046/j.1365-2958.2002.02739.x.
37. Nissen P, Hansen J, Ban N, Moore PB, Steitz TA. 2000. The structural basis of ribosome activity in peptide body synthesis. Science 289:920–930. https://doi.org/10.1126/science.289.5481.920.
38. Takyar S, Hickerson RP, Noller HF. 2005. mRNA helicase activity of the ribosome. Cell 120:49–58. https://doi.org/10.1016/j.cell.2004.11.042.
39. Studer SM, Joseph S. 2006. Unfolding of mRNA secondary structure by the bacterial translation initiation complex. Mol Cell 22:105–115. https://doi.org/10.1016/j.molcel.2006.02.014.
40. Unoson C, Wagner EGH. 2007. Dealing the stable structures at ribosome binding sites: bacterial translation and ribosome standby. RNA Biol 4:113–117. https://doi.org/10.4161/rna.4.3.5350.
41. Petersen HU, Joseph E, Ullmann A, Danchin A. 1978. Formylation of initiator tRNA methionine in prokaryotic protein synthesis: in vivo polarity in lactose operon expression. J Bacteriol 135:453–459.
42. Cone KC, Steege DA. 1985. Functional analysis of *lac* repressor restart sites in translational initiation and reinitiation. J Mol Biol 186:733–742. https://doi.org/10.1016/0022-2836(85)90393-6.
43. Draper DE, von Hippel PH. 1978. Nucleic acid binding properties of *Escherichia coli* ribosomal protein S1. J Mol Biol 122:321–359. https://doi.org/10.1016/0022-2836(78)90193-6.
44. Subramanian AR. 1983. Structure and functions of ribosomal protein S1. Prog Nucleic Acid Res Mol Biol 28:101–142. https://doi.org/10.1016/S0079-6603(08)60085-9.
45. Boni IV, Artamonova VS, Tzareva NV, Dreyfus M. 2001. Non-canonical mechanism for translational control in bacteria: synthesis of ribosomal protein S1. EMBO J 20:4222–4232. https://doi.org/10.1093/emboj/20.15.4222.
46. Osterman IA, Evfratov SA, Sergiev PV, Donstova OA. 2013. Comparison of mRNA features affecting translation initiation and reinitiation. Nuc Acids Res 41:474–486. https://doi.org/10.1093/nar/gks989.
47. Nakamoto T. 2006. A unified view of initiation of protein synthesis. Biochem Biophys Res Commun 341:675–678. https://doi.org/10.1016/j.bbrc.2006.01.019.
48. Omotajo D, Tate T, Cho H, Choudhary M. 2015. Distribution and diversity of ribosome binding sites in prokaryotic genomes. BMC Genomics 16: 604–612. https://doi.org/10.1186/s12864-015-1808-6.
49. Chang B, Halgamuge S, Tang S. 2006. Analysis of SD sequences in completed microbial genomes: non-SD-led genes are as common as SD-led genes. Gene 373:90–99. https://doi.org/10.1016/j.gene.2006.01.033.
50. Berg L, Kucharova V, Bakke I, Valla S, Brautaset T. 2012. Exploring the 5′-UTR DNA region as a target for optimizing recombinant gene expression from the strong and inducible *Pm* promoter in *Escherichia coli*. J Biotechnol 158:224–230. https://doi.org/10.1016/j.jbiotec.2011.07.012.
51. Schleif R. 1972. Fine-structure deletion map of the *Escherichia coli* L-arabinose operon. Proc Natl Acad Sci U S A 11:3479–3484. https://doi.org/10.1073/pnas.69.11.3479.
52. Janssen GR, Bibb MJ. 1993. Derivatives of pUC18 that have BglII sites flanking a modified multiple cloning site and that retain the ability to identify recombinant clones by visual screening of *Escherichia coli* colonies. Gene 124:133–134. https://doi.org/10.1016/0378-1119(93)90774-W.
53. Miller JH. 1992. A short course in bacterial genetics: a laboratory manual for *Escherichia coli* and related bacteria, vol 1. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
54. Fredrick K, Noller HF. 2002. Accurate translocation of mRNA by the ribosome requires a peptidyl group or its analog on the tRNA moving into the 30S P site. Mol Cell 5:1125–1131. https://doi.org/10.1016/S1097-2765(02)00523-3.
55. Martin-Farmer J, Janssen GR. 1999. A downstream CA repeat sequence increases translation from leadered and unleadered mRNA in *Escherichia coli*. Mol Microbiol 31:1025–1038. https://doi.org/10.1046/j.1365-2958.1999.01228.x.