Identification of an Intragenic Ribosome Binding Site That Affects Expression of the uncB Gene of the Escherichia coli Proton-Translocating ATPase (unc) Operon

SHARLENE R. MATTEN,† THOMAS D. SCHNEIDER, STEVEN RINGQUIST, AND WILLIAM S. A. BRUSILOW

Department of Biochemistry and Molecular Biology, Wayne State University School of Medicine, Detroit, Michigan 48201; Laboratory of Computational and Experimental Biology, National Cancer Institute, Frederick, Maryland 21702-1201; and Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, Colorado 80309

Received 7 August 1997/Accepted 1 June 1998

The uncB gene codes for the a subunit of the F o, proton channel sector of the Escherichia coli F 1 F o ATPase. Control of expression of uncB appears to be exerted at some step after translational initiation. Sequence analysis by the perceptron matrices (G. D. Stormo, T. D. Schneider, L. Gold, and A. Ehrenfeucht, Nucleic Acids Res. 10:2997–3011, 1982) identified a potential ribosome binding site within the uncB reading frame preceding a five-codon reading frame which is shifted one base relative to the uncB reading frame. Elimination of this binding site by mutagenesis resulted in a four- to fivefold increase in expression of an uncB-lacZ fusion gene containing most of uncB. Primer extension inhibition (toeprint) analysis to measure ribosome binding demonstrated that ribosomes could form an initiation complex at this alternative start site. Two fusions of lacZ to the alternative reading frame demonstrated that this site is recognized by ribosomes in vivo. The results suggest that expression of uncB is reduced by translational frameshifting and/or a translational false start at this site within the uncB reading frame.

The Escherichia coli unc (also called atp) operon is comprised of nine genes which encode the subunits of the proton-translocating F o sector of the ATP synthase (1). There are at least two ways to encode a subunit, and there is evidence for at least three initiation sites (2, 3, 9). The subunit a is encoded by two genes: uncA and uncB. The subunit c is encoded by uncC, which in addition to being a typical Shine-Dalgarno ribosome binding site has a second, alternate Shine-Dalgarno site 13 codons upstream (3, 14). Translational initiation at position 285 of uncB results in a translational frameshift that directionalizes the initiation codon to a stop codon. The initiating codon for the alternate reading frame is shifted one base to the right of the AUG (3, 10). The a and c subunits are synthesized in significant different amounts (2, 19). However, our previous studies on expression of a series of uncB-lacZ fusion genes (25) demonstrated that translational initiation of uncB is apparently as efficient as initiation of uncE (c subunit). Early fusions, containing 10 to 15% of uncB fused to lacZ, produced as much β-galactosidase activity as an uncE-lacZ fusion. A late uncB-lacZ fusion, carrying about 95% of uncB, however, was expressed 10 times less well than the early fusions. Also, Lang et al. (9) showed that even though the a and c subunits were synthesized in vitro at significantly different rates, ribosome binding to uncB appeared to be very similar to that of uncE, especially when the ratio of mRNA to ribosomes was low. The uncB reading frame might therefore contain some posttranslational initiation signals which decrease synthesis of the a subunit. The present study analyzed the uncB reading frame for alternative ribosome recognition sites and tested the hypothesis that uncB expression is influenced by the presence of such a site.

MATERIALS AND METHODS

Plasmids. Plasmids pDKWH103, pKS104, and pKS105 were described previously (25). Plasmid pSRM106 was identical to pKS104 except for the T→C mutation at position 285 of uncB, as shown below (see Fig. 2). The 617-bp BamHI fragment within uncB and cloned into M13 and mutagenized to change the start codon to ACG, by using the mutagenic primer 5‘-GGACGCGCAGGATCCAGCTTTCT-3’. The entire fragment was sequenced to ensure that there were no additional mutations. The mutagenized fragment was then used to construct pSRM106 just as the wild-type fragment had been used to construct pKS104 (25). Plasmid pSRM114 was constructed by mutagenizing the 617-bp BamHI fragment so that the termination codon of the false start initiation frame was eliminated, and the false start reading frame in frame to the uncB reading frame. The mutagenic primer 5‘-GGACGCGCAGGATCCAGCTTTCT-3’ results in the deletion of the first two bases, TG, of the stop codon for the alternative reading frame. Plasmid pWSB52 consists of lacZ fused directly to the false start initiation codon. An NcoI site was added to the lac fusion vector pMLB1034 (24) by digesting the plasmid with EcoRI and ligating.

* Corresponding author. Mailing address: Department of Biochemistry and Molecular Biology, Wayne State University School of Medicine, Scott Hall, 540 E. Canfield Ave., Detroit, MI 48201. Phone: (313) 577-6659. Fax: (313) 577-2765. E-mail: wbrusilow@med.wayne.edu.
† Present address: U.S. Environmental Protection Agency, Washington, DC 20460.
‡ Present address: Sidney Kimmel Cancer Center, San Diego, CA 92121.
in the presence of an EcoRI-Ncol adaptor, 5′-AATTCCTCATGGG-3′. The resultant plasmid, pWSB53, contained an Ncol site, the ATG for which is in frame with lacZ. The false start initiation codon was then cloned into this vector by digesting pKS104 with PstI and Ncol and cloning the resultant fragment into pWSB53 which had been digested with PstI and Ncol. The resultant plasmid, pWSB52, carries the unc promoter, uncI, and uncB up to the false start initiation codon, the ATG for which is part of the Ncol site, and is in frame with the lacZ gene of pMLB1034.

Assays of β-galactosidase activity. As described by Solomon et al. (25), these fusions were constructed in plasmids and then recombined into λ and finally into the λ at site to create single-copy fusions in the chromosome of MC1000 Δ(uncI-uncC). An E. coli strain deleted for both unc and lac (1), β-Galactosidase activities produced by each fusion gene in single-copy lysogens were assayed as described by Miller (17). The values for the β-galactosidase activities produced by single copies of the fusions in pDKWH103 and pKS105 are from the work of Solomon et al. (25).

Toeprint analysis. Primer extension inhibition (toeprint) analysis of the atrogenic false start site within uncB was performed. The extent of ribosome binding in the initiation region of protein synthesis was determined by primer extension inhibition. RNA was synthesized in vitro from PCR-generated DNA. Primers containing the 77 polymerase promoter sequence and DNA flanking the false start site were used to mutagenize the sequence. T7 polymerase was then used to synthesize RNA. The transcript was purified on a 6% acrylamide gel and annealed to a 3′-end-labeled oligonucleotide primer (see Fig. 3); this was followed by RNA sequencing reactions and primer extension inhibition (toeprinting) reactions without and with 30S ribosomes and tRNAfMet, as described by Hartz et al. (8), except that gel-purified RNA was used. Toeprinting reaction mixtures contained 100 nm 30S subunits and 500 nm tRNAfMet. Preincubation was done for 10 min at 37°C; this was followed by primer extension for 15 min at 37°C with Moloney murine leukemia virus reverse transcriptase (200 U/reaction mixture).

Immunoblots of fusion proteins. E. coli MC1000 Δ(uncI-uncC) (1) carrying either pKS104, pSRM106, or pWSB52 was grown in LB-ampicillin medium to an optical density of 0.60 at 0.4 to 0.7 and then chilled, pelleted by centrifugation, resuspended in 3 ml of 10 mM MOPS (morpholinepropanesulfonic acid)–10 mM MgCl2 (pH 7), and lysed in a French press at 16,000 lb/in2. The pKS104 culture was grown in 25 ml of medium; the other two were grown in 250 ml. Undosed cells were removed by centrifugation, and the supernatant fractions were loaded onto a sodium dodecyl sulfate (SDS)–7% polyacrylamide gel. Electrophoresis and blotting onto a nitrocellulose membrane were done as described previously (3).

RESULTS

Potential ribosome binding sites of the entire unc sequence, including uncB, were distinguished from other sites by using three perceptron weight matrices developed previously, w101, w71, and w51 (26). Each perceptron weight matrix represents a ribosome as it scans the mRNA. Translational initiation occurs with some fixed probability at any site along the RNA message. This probability is related to the value given by the perceptron weight matrix when positioned at a given site. Regions with a proper initiation codon and Shine-Dalgarno sequence are given positive values by the perceptron matrices. A higher weight matrix score indicates that a particular sequence fitting within the bounds of the chosen matrix is more likely to be a ribosome binding site. The results of this analysis of the unc operon are shown in Table 1. All three matrices identified ribosome binding sites (i.e., produced positive values) at the initiation codons for uncB, -H, -A, -D, and -C. One matrix identified the initiation codon for uncE. None of the matrices identified ribosome binding sites preceding uncI or uncF. Other potential ribosome binding sites are indicated by positive values in the table. All three matrices identified a site within uncB (at base 1307) as being a potential ribosome binding site. Figure 1 shows the location of this potential false start and how the resultant reading frame is five codons long and shifted one base compared to the uncB reading frame. Except for this site and the true initiation codons for five of the ATPase genes, no other site in the operon was given a positive score by all three matrices.

Effect of mutating the false start ribosome binding site on expression of the late uncB′-lacZ fusion gene. As discussed above, even though the differential expression of most of the ATPase genes appears to be controlled at the level of translational initiation, past studies suggest that synthesis of the a subunit is controlled at some step after initiation. We tested the role of the alternative reading frame identified by perceptron analysis in decreasing the expression of the pKS104 uncB′-lacZ fusion gene, which contains most of the uncB gene (25) (Fig. 2). The CAU corresponding to histidine-95 (Fig. 1) was changed to a CAC (also histidine), creating plasmid pSRM106. This mutation converted the false-start AUG to ACG, thereby destroying the putative ribosome binding site. The score resulting from perceptron w101 analysis of this region changed from +61 to −69 as a result of this single change. The scores resulting from analysis by the other matrices were even more negative. Eliminating this false start ribosome binding site resulted in a four- to fivefold increase in expression of the uncB′-lacZ fusion gene from 10 to 12 U of β-galactosidase for the unmuted single-copy KS104 construction to 45 to 50 U for the mutated single-copy SRM106 fusion gene. Mutation of the site identified by perceptron analysis increases expression of uncB (Fig. 2).

<table>
<thead>
<tr>
<th>Nucleotide position</th>
<th>Weight matrix score</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>w51</td>
<td>w71</td>
<td>w101</td>
</tr>
<tr>
<td>uncI</td>
<td>374</td>
<td>−76</td>
</tr>
<tr>
<td>uncB</td>
<td>623</td>
<td>−199</td>
</tr>
<tr>
<td>uncE</td>
<td>708</td>
<td>−53</td>
</tr>
<tr>
<td>uncF</td>
<td>1024</td>
<td>13</td>
</tr>
<tr>
<td>uncG</td>
<td>1307</td>
<td>36</td>
</tr>
<tr>
<td>uncD</td>
<td>1868</td>
<td>131</td>
</tr>
<tr>
<td>uncE</td>
<td>1901</td>
<td>40</td>
</tr>
<tr>
<td>uncF</td>
<td>2187</td>
<td>326</td>
</tr>
<tr>
<td>uncG</td>
<td>2261</td>
<td>−62</td>
</tr>
<tr>
<td>uncD</td>
<td>2672</td>
<td>57</td>
</tr>
<tr>
<td>uncI</td>
<td>2801</td>
<td>−128</td>
</tr>
<tr>
<td>uncG</td>
<td>2919</td>
<td>24</td>
</tr>
<tr>
<td>uncE</td>
<td>3143</td>
<td>−50</td>
</tr>
<tr>
<td>uncF</td>
<td>3218</td>
<td>365</td>
</tr>
<tr>
<td>uncG</td>
<td>3443</td>
<td>113</td>
</tr>
<tr>
<td>uncD</td>
<td>3736</td>
<td>−208</td>
</tr>
<tr>
<td>uncE</td>
<td>4524</td>
<td>5</td>
</tr>
<tr>
<td>uncG</td>
<td>4810</td>
<td>6</td>
</tr>
<tr>
<td>uncD</td>
<td>5700</td>
<td>329</td>
</tr>
<tr>
<td>uncF</td>
<td>7103</td>
<td>175</td>
</tr>
</tbody>
</table>

The ECOUNC entry (accession no. J01594) was obtained from GenBank and searched with the three perceptron matrices (26). The matrices had not been trained on this sequence.

Nucleotide position from ECOUNC for each ATPase gene initiation codon or any sequence which produced a positive score in the perceptron matrix analysis.

Scores produced by analysis of each nucleotide position with each of the three matrices, w51, w71, and w101.

ATPase gene whose initiation codon corresponds to the given position.

—, position identified by all three matrices but not corresponding to any one of the known translational initiation sites.

FIG. 1. Location of translation false start within uncB. All three perceptron weight matrices identified this region as a translation initiation region. The bases between positions 271 and 303 of the uncB reading frame are shown. The amino acids of the uncB product coded for by those bases are shown above them. The false start translation initiation site identified by the perceptron matrices, followed by a short reading frame, is indicated below the bases.

![Image](https://example.com/image.png)
Toeprint analysis of the false start ribosome binding site.

To determine whether this false start region was actually capable of forming an initiation complex with ribosomes and initiator tRNA, we carried out primer extension inhibition experiments, also called toeprint analysis (8). In a toeprint experiment, reverse transcriptase is allowed to begin transcribing an mRNA from a primer downstream of a putative ribosome binding site. Without ribosomes in the reaction mixture, the reverse transcriptase should pass across the site. When ribosomes are included and allowed to bind, they block the transcriptase, and a distinct band forms 15 bases downstream of the initiation codon. Such analysis has been done previously for the ribosome binding sites of the genes of the unc operon (22).

In our initial studies on the mRNA produced by the pKS104 plasmid, which consists of the true unc promoter, uncI, and most of uncB fused in frame to lacZ (Fig. 2), we detected no toeprint at the false start. Computer analysis of the mRNA around the false start suggested the existence of a strong secondary structure (Fig. 3) which, in the in vitro toeprint experiment, might prevent or interfere with ribosome binding. A smaller stem-loop structure immediately following the larger stem-loop might also interfere with the primer extension reaction used to create the toeprint (Fig. 3). Using mutagenic PCR primers, we replaced 21 bases of one side of the stem-loop with 11 adenosine residues, and we made three single base changes in the smaller stem-loop to minimize its effect on the primer extension reactions (Fig. 3). Ribosome binding to the resultant mRNA revealed a strong toeprint 115 bases from the false start initiation codon, precisely where the perceptron analysis predicted it (Fig. 4). Therefore, this region is capable of forming an initiation complex with ribosomes once the mRNA secondary structure is disrupted.

The RNA synthesized for the toeprint analysis was too short in sequence to be analyzed by any of the perceptron weight matrices. To evaluate how these mutations might affect the perceptron analysis, we assumed the presence of 100 adenosine residues on each end of the sequence, analyzed those sequences with the perceptron matrices, and found that the changes which allowed us to observe a toeprint were changes which raised the perceptron score resulting from analysis with the w51 matrix but lowered the scores resulting from analysis with the other two matrices. It is therefore unlikely that these mutations were creating a new ribosome binding site. Interestingly, analysis of the mutated sequence by the w51 matrix gave a positive score to a site seven bases upstream of the false start.

FIG. 2. β-Galactosidase activities produced by in-frame uncB-lacZ fusions. Locations of the unc promoter (Punc), uncI, uncB, and uncE are shown at the top of the figure. The amounts of unc DNA fused in frame to lacZ in plasmids pDKWH103, pKS104, and pKS105 are indicated by the horizontal lines. Plasmid pSRM106 is identical to pKS104 except for the T→C mutation at position 285 of uncB, which is indicated by an asterisk. β-Galactosidase activities produced by each fusion gene in single-copy lysogens were assayed as described by Miller (17). The values for the β-galactosidase activities produced by single copies of the fusions in pDKWH103 and pKS105 are from the work of Solomon et al. (25).

FIG. 3. Putative mRNA secondary structure around translational false start. The sequence between bases 247 and 334 of uncB are shown folded into a secondary structure predicted by the Mfold program, version 2.0 (10, 11, 30). The false-start AUG and the primer used for sequencing and for primer extension in the toeprint experiment are indicated. To obtain a toeprint of this site, the region indicated by the arrows was replaced by 11 adenosine residues, and a single-base deletion and two single-base mutations were constructed in the small loop, as indicated.
site, and the primer extension inhibition analysis (Fig. 4) revealed a minor toeprint at that site.

Translational initiation at the false start. We fused lacZ to the false start reading frame two ways (Fig. 5). First, we deleted the first two bases, TG, of the stop codon of the false start reading frame so that this reading frame, instead of ending, was fused to the remainder of the uncB reading frame. The resultant plasmid, pSRM114, is therefore identical to pKS104 except for a mutation which disrupts the true uncB reading frame at the stop codon for the false start reading frame. Second, we fused lacZ in frame directly to the false start initiation codon to create pWSB52. β-Galactosidase activity produced by either of these two constructions probably results from translation initiation at the false start or ribosomal frameshifting at the false start. We compared the β-galactosidase activities produced in uncB-deleted cells carrying pKS104, pSRM114, or pWSB52. The true uncB-lacZ fusion plasmid pKS104 produces between 200 and 600 U of activity in cells grown on minimal medium containing antibiotic. Measurements of activity produced by high-copy-number plasmids produce a much wider range of activities than those produced by single-copy lysogens.) Cells carrying plasmid pSRM114 produce 10 to 20 U, and cells carrying pWSB52 produce 50 to 60 U under the same conditions. When the false start codon in pWSB52 was mutagenized to an ACG, the β-galactosidase activity produced from the resultant multicopy plasmid dropped by an average of 70%. Although this decrease confirms that translation is initiated at this site, the fact that activity was not completely abolished indicates that additional factors contribute substantially to ribosomal frameshifting and/or initiation at this site.

When the fusion in pWSB52 was moved into λ to create single-copy lysogens, it produced very low but non-zero activity in our normal assay, approximately 5 to 10% of the activity produced by the pKS104 lysogen. The single-copy activity produced from the SRM114 construction was too low to measure.

As a further demonstration that ribosomes recognize the false start site in vivo, we analyzed whole-cell lysates of MCI000 Δ(unc1-uncC) carrying the three fusion plasmids pKS104, pSRM114, and pWSB52 by immunoblotting with anti-β-galactosidase (Promega). The results are shown in Fig. 6. Lane 1 contains the products of pKS104. The full-length uncB-lacZ product is clearly visible as the highest-molecular-weight band. As is often observed with β-galactosidase fusion proteins, all three lysates contain a significant amount of the β-galactosidase moiety alone, which probably results from proteolysis at or near the fusion joint. Lane 2 contains the products of pSRM114. The protein visible as the top band, which is the same size as the fusion protein produced by pKS104, could be produced only by translational initiation at the true uncB start codon and ribosomal frameshifting in the vicinity of the false start reading frame. Compared to pKS104, pSRM114 also produces an additional fusion protein (lane 2). This additional protein is the correct size for a fusion protein that was initiated at the false start initiation codon. Since this band is not visible at all in the products of pKS104, it is unlikely to be a proteolysis product of the full-length fusion protein. The fusion protein made from pWSB52 (lane 3) would be the same size as the β-galactosidase moiety produced by cleavage at the fusion joint of the other full-length fusion proteins, if translation were to initiate at the false start. The higher-molecular-weight protein produced by pWSB52 (lane 3) is the correct size for the product of the true uncB translational start site frameshifted at the false start. The relatively small difference in molecular weight between this frameshifted protein produced by pWSB52 and the protein initiating at the false start in pSRM114 would not be detectable on these gels (∼5,500 difference in proteins with
of plays a posttranslation initiation role in controlling expression resulting only from a frameshift at the false start of ribosomes which had initiated. The top band in lane 2 and 3 therefore contained 20 times as much total protein as lane 1. The top band in lane 1 is the full-length translation protein coded for by pKS104; it consists of most of the a subunit fused in frame to pWSB52, 250 g (lane 3). Lanes 2 and 3 probably represent the ribosomal frameshifting at the false start. The second band (arrow) in lane 2 represents a lac fusion protein initiated at the false start site. The top band (arrow) in lane 3 is the proper size for a protein initiated at the true uncB translational start site and fused to lacZ by frameshifting at the false start. The second bands in lanes 1 and 2 represent the uncB-galactosidase moiety (BG) derived from proteolysis of each fusion protein at the fusion joint. The equivalent band in lane 3 probably represents the fusion protein derived from translational initiation at this site due to either the frameshifted ribosome binding site, the mRNA secondary structure, or both. No footprint is observed on RNA carrying the wild type sequence, so it is unlikely that unbound ribosomes initiate translation at this site. The low rate of expression of lacZ fusions to the false start ribosome binding site also indicates that ribosomes frameshift and initiate translation at this site poorly. The effect of this site on uncB expression may involve a relationship between translation and mRNA breakdown. The uncB mRNA has been shown to be more susceptible to degradation than other cistrons in the operon (13, 16, 20, 21). It has been demonstrated by Chevrier-Miller et al. (4) that uncoupling of transcription and translation leads to differential mRNA half-lives of lacZ RNA. These authors proposed that inefficient ribosome loading might unwind the RNA without protecting it, making such RNA less stable than untranslated RNA. In studies where RNaseE cleavage of mRNA, Gross (7) has speculated that RNase recognition of mRNA requires some undefined interactions of mRNA with ribosomes, so that only translated mRNA is susceptible to degradation. It is therefore possible that the translational false start identified in these experiments might affect either the number of ribosomes that complete the translation of uncB, the sensitivity of the uncB message to endonucleolytic degradation, or both.

ACKNOWLEDGMENTS

This research was supported by National Science Foundation grant DMB-9096159 and by American Heart Association grant-in-aid 93007730. S.R.M. was supported by grant 7020010 from the American Heart Association, Maryland Affiliate.

We thank Heven Sze of the Department of Botany and Tomas Kempe of the Protein and Nucleic Acid Synthesis Laboratory, both at the University of Maryland. We also thank Larry Gold of the Department of Molecular, Cellular, and Developmental Biology at the University of Colorado for the use of his laboratory for the toeprinting and Robert Traut of the University of California, Davis, for generously providing purified 30S ribosomes.

REFERENCES

RIBOSOME BINDING SITE AFFECTS E. COLI uncB EXPRESSION


