Transcribed sequences of the *Escherichia coli* btuB gene control its expression and regulation by vitamin B₁₂
(attenuation/gene fusions/post-transcriptional regulation/transport)

MICHAEL D. LUNDRIGAN*, WOLFGANG KÖSTER†, AND ROBERT J. KADNER‡
Department of Microbiology, School of Medicine, University of Virginia, Charlottesville, VA 22908

Communicated by John Roth, November 13, 1990

ABSTRACT The *Escherichia coli* btuB gene product is an outer membrane protein required for the active transport of vitamin B₁₂ and other cobalaminis. Synthesis of BtuB is repressed when cells are grown in the presence of cobalaminis. Mapping of the 5′ end of the btuB transcript revealed that a 240-nucleotide transcribed leader precedes the coding sequence. Point mutations causing increased expression under repressing conditions were isolated by use of a btuB-lacZ gene fusion. Mutations at many sites within the leader region affected btuB-lacZ regulation, whereas some base changes upstream of the start of transcription affected the absolute level of expression but not its repressibility. Analysis of btuB-phoA gene fusions and btuB-lacZ operon and gene fusions of various lengths showed that sequences within the btuB coding region (between nucleotides +250 and +350) had to be present for proper expression and transcriptional regulation. Sequences within the leader region (up to +250) conferred regulation of translational fusions. These results indicate that btuB expression is controlled at both the transcriptional and translational levels and that different but possibly overlapping sequences in the transcribed region, including the coding region for the transport protein itself, mediate these two modes of regulation.

The outer membrane protein BtuB plays an essential role in the transport of vitamin B₁₂ (cyanocobalamin, CN-Cbl) and other cobalamins in *Escherichia coli*. BtuB binds cobalamins with high affinity and, in conjunction with the tonB product, catalyzes their active transport across the outer membrane (1–3). Specific transport systems in the outer membrane are required for uptake of cobalamins and iron-siderophore complexes, which are too large to diffuse effectively through the porin channels. Synthesis of these outer membrane transport proteins is regulated in response to the availability of their substrates. Expression of the ferric siderophore transport proteins is repressed when cells are grown in the presence of excess iron. This regulation depends on the regulatory gene fur, encoding a repressor protein which binds in the presence of various divalent cations to specific operator sequences in regulated promoters (4, 5).

The amount of BtuB in the outer membrane is repressed by growth with cobalamins (6). An unusual feature of btuB regulation is that repression seems to be as effective when the btuB gene is carried on a multicopy plasmid as when in single copy. This was found both for the intact gene by Heller et al. (7) and for a btuB-lacZ gene fusion by Aufrere et al. (8), who suggested that the apparent lack of titration by excess copies of its target sequences could indicate that there is a large excess of the putative repressor in the cell. Efforts to obtain mutations in a gene for a cobalamin-responsive repressor have been unsuccessful. Using a btuB-lacZ fusion in which synthesis of β-galactosidase was regulated in the same manner as the intact btuB gene, Lundrigan et al. (9) obtained mutants that showed elevated β-galactosidase activity in the presence of CN-Cbl. Some of these mutants showed fully constitutive expression of btuB and carried mutations at btuR, unlinked to the btuB locus. The BtuR phenotype is recessive to wild type and affects all btuB genes in the cell, as expected for a defective repressor. However, further analysis indicated that the btuR product is actually involved in the metabolic conversion of CN-Cbl to adenosylcobalamin (Ado-Cbl) (10). Thus, btuR mutants are completely defective in the synthesis of Ado-Cbl but not other intracellular cobalamin derivatives. Exogenous Ado-Cbl, but not other cobalamins, is still fully effective at repressing synthesis of BtuB in a btuR mutant.

This paper describes the properties of cis-acting regulatory mutations obtained by screening for altered expression of the btuB-lacZ fusion and which are linked to btuB. These mutations are shown to lie within the transcribed leader region upstream of the coding sequence. In addition, the regulatory behavior of various btuB fusions to reporter moieties revealed a marked difference between transcriptional and translational fusions. The apparent requirement for a substantial portion of the btuB transcript in control of its own expression and regulation suggests that post-transcriptional events involving the leader and btuB coding region influence both transcriptional readthrough and translation initiation, perhaps determined by the RNA secondary structure. Involvement of transcribed regions in regulation has been thoroughly documented for attenuation control of many amino acid biosynthetic pathways in bacteria (11). The unusual features of btuB regulation are that important regulatory sites are located within the btuB coding sequence and that this regulation affects both transcription and translation.

MATERIALS AND METHODS

Bacterial Strains and Plasmids. The *E. coli* K-12 strains used in this study are derivatives of strain RK5173 and have been previously described (9, 10, 12). Strains CC118 and CC202 [CC118/F42 lacI3 50::TnphoA] were from Manoil and Beckwith (13).

Plasmid pH3-5 carries the intact btuB gene as a 2.2-kilobase (kb) partial Sau3A fragment in the BamHI site of pBR322 (14). Plasmid pML395 was constructed by inserting the 316-base-pair (bp) EcoRI–HindIII fragment containing the btuB promoter into the replicative form of M13gt130.

Abbreviations: CN-Cbl, cyanocobalamin or vitamin B₁₂; Ado-Cbl, adenosylcobalamin; X-Gal, 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside.

*Present address: Department of Microbiology, University of Mississippi School of Medicine, Jackson, MS 39216.
†Present address: Mikrobiologie II, Universität Tübingen, D7400 Tübingen, F.R.G.
‡To whom reprint requests should be addressed.
§The sequence reported in this paper has been deposited in the GenBank data base (accession no. M10112).
followed by cloning the insert as an EcoRI-BamHI fragment in plasmid pRS414. This construction encodes a hybrid gene containing the first five codons of \textit{btuB} (Met-Ile-Lys-Lys-Ala) and three codons (Trp-Asp-Pro) from the multiple cloning site fused in frame to the ninth codon of \textit{lacZ}. Plasmid pML261 carries a 2.4-kb \textit{btuB}\textsuperscript{+} \textit{Cla} I fragment in pUC8 (12); plasmid pAGI was derived from pML261 by replacement of the \textit{HindIII} site in the multiple cloning region with an \textit{Xho} I site (15). The \textit{lac} operon and gene fusion vector plasmids pRS415 and pRS414 were obtained from R. Simons (16).

Minimal growth medium was medium A supplemented with 20 mM glucose and required supplements (ref. 17, p. 432).

**Enzyme Assays.** Cells were grown in medium A supplemented with 20 mM glucose, 0.5% Difco Casamino acids, thiamin (1 \textmu g/ml), and methionine (100 \textmu g/ml) (10). \beta-Galactosidase and alkaline phosphatase activities were measured by the increase in absorbance at 420 nm of 2 mM solutions of o-nitrophenyl \beta-D-galactopyranoside in Z buffer (ref. 17, p. 352) or p-nitrophenyl phosphate in 1 M Tris-HCl, pH 8.0, respectively. \beta-Lactamase activity was similarly measured using the chromogenic substrate CENTA (Calbiochem-Behring). Units of activity are reported as the change in absorbance \times \min^{-1} \times (2 \times 10^9 \text{cells})^{-1}. For cells carrying fusions on plasmids, enzyme activities were normalized to \beta-lactamase activity to correct for variations in plasmid copy number.

**Transcript Start Site.** The 5' terminus of \textit{btuB} mRNA was determined by protection against S1 nuclease hydrolysis, using as probe the 316-nucleotide EcoRI-HindIII fragment labeled with \textsuperscript{32}P at the 5' end generated by \textit{HindIII} (18). RNAs were isolated from cells of strains RK5173 (\textit{btuB}\textsuperscript{+}), RK4793 (\textit{dtbuB}), and RK4793/pKH3-5. The products of S1 nuclease protection were electrophoresed next to products of dideoxy sequencing reactions using M13mp18 carrying the EcoRI-HindIII fragment as template and the Nae I-HindIII fragment of \textit{btuB} as primer.

**Construction of \textit{btuB-lac} and \textit{btuB-phoA} Fusions.** Operon and gene fusions to \textit{lacZ} were constructed in the plasmid vectors pRS413 and pRS414 (16), respectively, by inserting EcoRI-BamHI fragments from plasmid pAGI derivatives with 6-bp BamHI linker insertions in \textit{Hpa} II sites of the \textit{btuB} gene or its upstream region, as previously described (15).

The \textit{btuB-phoA} fusions were isolated by the method of Manoil and Beckwith (13). The F' \textit{zfr-TnphoA} plasmid from strain CC202 was introduced by conjugation into strain CC118 carrying monomeric species of the \textit{btuB}\textsuperscript{+} plasmid pML261 (18); plasmid transfer was selected on medium containing ampicillin (25 \mu g/ml) and kanamycin (25 \mu g/ml). Cells in which \textit{TnphoA} transposed onto plasmid pML261 were selected on medium containing ampicillin (25 \mu g/ml) and high levels of kanamycin (300 \mu g/ml). PhoA\textsuperscript{+} colonies were identified by their blue color on minimal medium containing 5-bromo-4-chloro-3-indolyl \beta-D-galactopyranoside (X-Gal)-containing minimal medium or on MacConkey-lactose indicator medium, both with 5 \mu M CN-Cbl.

**RESULTS**

**Transcription Start Site.** The 5' end of \textit{btuB} mRNA was identified by S1 nuclease protection assay (Fig. 1). RNA was extracted from haploid cells and from cells harboring the \textit{btuB}-containing plasmid pKH3-5 that were grown in the absence and presence of 5 \mu M CN-Cbl. In all cases, cellular RNA protected a single species against S1 nuclease digestion. The size of this protected fragment indicated that the 5' terminus of \textit{btuB} mRNA corresponded to the C residue located 240 nucleotides upstream of the start of translation, in exact agreement with the finding of Aufrere \textit{et al.} (8). The sequence of nucleotides -60 to +340, relative to the \textit{btuB} transcription start site, is presented in Fig. 2. There is a likely promoter sequence which matches the consensus at 9 of 12 positions, with a spacing of 18 nucleotides between the -35 and -10 regions. Some regions of extensive, punctuated dyad symmetry are also indicated and discussed below.

The level of \textit{btuB} mRNA was substantially lower in haploid \textit{btuB}\textsuperscript{+} cells than in pKH3-5-bearing cells, and the mRNA was absent from cells carrying a \textit{btuB} deletion, confirming that the protected species was specific for \textit{btuB}. The amount of

**Isolation of Regulatory Mutations.** Plasmid pML95 was mutagenized with hydroxyamine as described (19). For bisulfite mutagenesis (20), the viral strand of M13tg130 carrying the EcoRI-BamHI fragment of \textit{btuB} was annealed with an excess of denatured double-stranded M13tg130. Treatment with sodium bisulfite affected only the single-stranded insert. After transcription and recloning of the EcoRI-BamHI fragment in pRS414, transformants were screened for altered \textit{btuB-lac} expression by the color of colonies on 5-bromo-4-chloro-3-indolyl \beta-D-galactopyranoside (X-Gal)-containing

**Fig. 1.** Mapping of the 5' end of \textit{btuB} mRNA by S1 nuclease protection. The restriction map of the 5' region of \textit{btuB} and the location of the various probes are shown at the top; the filled bar indicates the location of the \textit{btuB} coding region. The arrow indicates the transcription start site. The 316-bp EcoRI-HindIII fragment was labeled at the 5' end of the HindIII site with polynucleotide kinase and [\gamma-\textsuperscript{32}P]ATP. The probe was hybridized to 150 \mu g of total RNA from strains RK4793/pKH3-5 grown in the absence (lane 5) or presence (lane 7) of 5 \mu M CN-Cbl, RK4793 (\textit{dtbuB}) (lane 6), or RK5173 (\textit{btuB}\textsuperscript{+}) (lane 8). The sequencing ladder (lane 1-4) was the product of dideoxy termination reactions using an M13mp18 virus with the same EcoRI-HindIII fragment as template and the Nae I-HindIII fragment of \textit{btuB} as primer, which allows exact correspondence with the protected species. The upper band in lanes 5-8 with S1-digested samples represents re-formed double-stranded probe; the lower band is the protected species. The sequence in the vicinity of the transcription start site is shown at the bottom.
btuB mRNA in pKH3-5-bearing cells was substantially reduced by growth with repressing levels of CN-Cbl, suggesting that repression affects mRNA synthesis or stability.

Isolation of Mutations Affecting btuB Regulation. To identify sequences linked to btuB that are important in its regulation, the 316-bp EcoRI– HindIII (BamHI) DNA fragment (carrying the region from −60 to +253) was mutagenized with sodium bisulfite or hydroxylamine, then fused to lacZ in the vector pRS414. These constructs were then introduced into strain RK5173 by transformation and variants exhibiting increased or decreased β-galactosidase production in the presence of 5 μM CN-Cbl were isolated. Expression of β-galactosidase from pML95, in which the wild-type EcoRI–HindIII (BamHI) fragment is linked to lacZ, was repressed about 10-fold by 5 μM CN-Cbl.

A set of 36 variant plasmids was analyzed, in which a single base change was associated with altered expression of β-galactosidase on X-Gal-containing medium under repressing conditions. The mutations were all G→C→A→T transversions, as expected from the mutagenesis procedures used. There were 28 unique mutations, since two or three independent isolates contained the same change at nucleotides +49, +92, +106, +189, or +215. Fig. 2 shows the location of these mutations and Fig. 3 presents the relative levels of β-galactosidase under nonrepressing and repressing conditions, plotted against the site of the mutation. All mutations that affected repressibility lay within the transcribed region. Mutations at 14 sites throughout the promoter-proximal portion of the leader region between nucleotides +27 and +159 conferred approximately wild-type levels of β-galactosidase under nonrepressing conditions. Under repressing conditions, expression in these mutants was on average 4 times higher than parental, although no more than 50% of the nonrepressed level. Five mutations in this region displayed reduced activity (<50% of parental level). Three mutations affecting the distal portion of the leader (nucleotides +179, +189, and +215) resulted in substantial elevation of btuB expression under both repressing and nonrepressing conditions, and the base substitution at +253 resulted in greatly increased expression under both conditions (23.2 and 0.6 units, respectively). No mutations resulted in fully constitutive behavior, suggesting that different sequences in the leader act in independent manner to control gene expression.

Mutation that resulted in greatly decreased (>10-fold) btuB-lacZ expression were found at sites which could affect recognition of the promoter by RNA polymerase (nucleotides −40 or −13), or the Shine–Dalgarno sequence (+234), which interfere with translation initiation. Synthesis of β-galactosidase in these mutants was repressed by CN-Cbl, although the degree of regulation could not be accurately determined owing to the low activity. Mutations at −43, −30, +69, +86, or +136 had no apparent effect on btuB-lacZ expression or repression. Thus, no mutations that affected repressibility were found upstream of the transcription start site, as expected for an operator site. Instead, sequences...

![Fig. 3. Effect of point mutations on btuB expression. Strain RK5173 bearing derivatives of the btuB-lacZ plasmid pML95 with single point mutations in the 316-bp EcoRI–HindIII fragment were grown in the absence (circles) or presence (bars) of 5 μM CN-Cbl. The graph presents the ratio of β-galactosidase to β-lactamase activities, plotted against the site of the mutation, relative to the start of transcription at +1. The dotted and dashed horizontal lines indicate the activity expressed by parental plasmid pML95 under nonrepressing and repressing conditions, respectively.](image)

throughout the transcribed leader were crucial for repression and also influenced the derepressed level of btuB expression.

Regulation of Transcriptional and Translational Fusions. To examine further the role of transcribed sequences in btuB regulation, increasing lengths of the btuB gene from the EcoRI site at −60 were placed in front of reporter genes. Transcriptional regulation was determined with btuB-lacZ operon fusions in which EcoRI–BamHI fragments from a series of BamHI linker insertions in btuB (18) were cloned in the lac operon fusion plasmid pRS415 (16). Regulation of transcription and translation was determined with btuB–phoA gene fusions (13). Different reporter groups were used because of the lethality engendered by export of β-galactosidase fusion proteins and because alkaline phosphatase activity requires the fusion of that reporter gene beyond the signal sequence of a secreted protein. Relative β-galactosidase or alkaline phosphatase activities in cells grown under nonrepressing or repressing conditions are plotted in Fig. 4 as a function of the distance of the fusion junction from the start of transcription.

All btuB–phoA gene fusions carrying btuB sequences from nucleotide +368 (corresponding to amino acid 22 of the mature BtuB polypeptide) to near nucleotide +1200 expressed similar levels of alkaline phosphatase activity (Fig. 4B). Fusions beyond nucleotide 1200 showed progressively lower levels of PhoA activity owing to release of the enzyme into the medium (data not shown). All btuB–phoA fusions had low PhoA activity when grown in the presence of CN-Cbl,...
sequences between +250 and +313, within the structural gene, are necessary for transcriptional repression.

To test for control at the translational level, btuB–lacZ operon and protein fusions with the same fusion junction were compared. These were formed by insertion of the btuB fragment from −60 to +256 into plasmid pRS415, to yield plasmid pML95 encoding an in-frame fusion protein, or into plasmid pRS415, to yield plasmid pML96 encoding an operon fusion. The levels of β-galactosidase, relative to β-lactamase, from the operon fusion plasmid pML96 grown under nonpressing and repressing conditions, were 42.1 and 31.8 units, respectively (repression ratio, 0.79). The corresponding values for the protein fusion plasmid pML95 were 0.68 and 0.04 unit (repression ratio, ∼0.06). The substantially lower β-galactosidase activity shown by the protein fusion could result from many factors, including decreased stability or specific activity of the hybrid BtuB–LacZ protein, or inefficient translation initiation owing to the use of the weak btuB Shine–Dalgarno sequence or to sequestration of the Shine–Dalgarno sequence by base pairing. However, the finding that the protein fusion at +256 is regulated by CN-Cbl, while the operon fusion is not, indicates that sequences before +256 are able to confer repressibility by CN-Cbl at the translational level but not at the transcriptional level. Sequences between +256 and +368 are needed for transcriptional repression.

**Regulation of btuB Expression in a btuB Mutant Strain.** The role of the btuB product in the regulation of its own synthesis is difficult to determine since BtuB is required for the efficient uptake of the repressing effector, CN-Cbl. We previously showed that expression of a btuB–lacZ fusion in a btuB host was not repressed by 10−6 M CN-Cbl, whereas full repression in a btuB+ host occurred at 10−8 M CN-Cbl (9). We report here that btuB–lacZ fusions in a ΔbtuB strain were repressed almost fully at external concentrations of 10−4 M CN-Cbl. Furthermore, the level of expression and regulation by CN-Cbl of a chromosomal btuB–lacZ fusion was not substantially affected by the presence of a multicopy plasmid carrying btuB+. Both of these results indicate that BtuB is not directly involved in its own regulation.

**DISCUSSION**

Single base changes in the 240-nucleotide btuB leader decreased but did not eliminate repression by CN-Cbl, showing the important role of the leader in gene regulation. Expression of reporter genes fused to increasing lengths of btuB confirmed that transcribed sequences are necessary for gene expression and proper regulation and that the intact btuB protein is not directly involved in its regulation. The pattern of expression from these fusions suggested that regulation might operate at both the transcriptional and translational levels and that different portions of the leader transcript participate in the two modes of regulation.

The btuB leader region contains substantial segments of punctuated dyad symmetry, although the existence and role of these potential RNA secondary structures in gene regulation will require further genetic and physical analysis. The sequences of several of these regions are presented in Fig. 5, and a tentative model for their participation in gene regulation, similar to transcription attenuation mechanisms (11), might help account for the results observed here. Analysis of portions of the first 400 nucleotides by the RNA-folding program of Zuker and Stiegler (22) suggested the possibility of alternative secondary structures. Complementary segments C and C′, as defined in Fig. 2, might modulate translation initiation by sequestration of the Shine–Dalgarno sequence contained in C′. Supporting evidence for this process is the fact the C → G mutation at position 234 in region C′ greatly reduced btuB–lacZ hybrid gene expression,
whereas the C → G mutation at position 215 in region C resulted in increased expression that was repressed only 2-fold.

Furthermore, analysis of the fusions showed that sequences between +260 and +313 participate in transcriptional regulation. This region contains segments D and D', whose transcript could form a 10-bp G+C-rich stem and 7-nucleotide loop followed by a run of 4 U residues, typical of rho-independent transcription terminators. This D-D' region from +258 to +284 overlaps segment E', from +259 to +309, which has extensive base-pairing possibilities with segment E at the 5' end of the leader. Thus, pairing of region E' with E would prevent formation of the D-D' stem and thereby allow btuB transcription to continue. Segment E itself is composed of complementary regions A and A'. Estimation of the stability of these secondary structures using the parameters of Tinoco et al. (21) suggested that E-E' is somewhat more stable than A-A' + D-D', by -7.9 kcal/mol. It is possible that the binding of a regulatory protein to A-A' in an Ado-Cbl-dependent manner could prevent E-E' from forming and thereby favor transcription attenuation.

This model can explain the different pattern of expression when the btuB fragment from −60 to +256 was coupled to lacZ as an operon or a protein fusion. High and unregulated activity from the operon fusion could result from the absence of the putative transcription terminator between +258 and +289. The fact that expression from this operon fusion is not as high as in the next longer fusion might indicate that the C-C' stem and loop formed by three U residues at +210 to +237 (the Shine-Dalgarno region and its complement) has weak terminator activity. The low but repressible expression of the btuB-lacZ gene fusion, despite the high levels of transcription as indicated by the operon fusions, suggests that sequestration of the Shine-Dalgarno sequence may play a significant role in btuB expression.

Many of the point mutations that affected btuB regulation occurred in the region corresponding to the E segment. Other mutations affecting regulation lay between region E and region C, suggesting that secondary structures or specific interactions may include additional regions between E and E'. One region of dyad symmetry, indicated B and B' in Fig. 2, included several mutations that decreased repressibility and increased basal gene expression. It must be noted, however, that all these mutations were characterized in a fusion construct that lacked at least one major regulatory region—including the transcription attenuator. Hence, their significance must not be overemphasized beyond the conclusion that transcribed sequences are involved in regulation. Our search for mutants defective in a trans-acting factor for repression made use of the same lacZ gene fusion at nucleotide +256. As shown here, this construct is not properly expressed, even though subject to regulation. Use of longer gene fusions should be informative.

In conclusion, a model for btuB gene regulation is proposed in which secondary structures in the transcript determine both transcription and translation of that gene. Few examples exist of the participation of part of a structural gene in regulation of its synthesis. One precedent is the binding of the galactose-responsive repressor protein, GalR, to an internal operator sequence within the galE structural gene and to a similar sequence centered 60 bp before the start of transcription (23). We are not aware of examples where regulation by transcription attenuation involves sequences within the structural gene.

We are indebted to M. J. Friedrich for sequencing of the constructs containing point mutations. We thank C. Manoil, J. Beckwith, and R. Simons for necessary strains and plasmids; and A. Beyer, R. Gunsalus, S. French, K. Heller, and M. Smith for interest and comments. This work was supported by Research Grant R01 GM19078 from the National Institute of General Medical Sciences. M. D. L. was a postdoctoral trainee on National Research Service Award Grant T32 CA09109 from the National Cancer Institute; W. K. was supported by a stipend from the Deutsche Forschungsgemeinschaft.

Fig. 5. Structures of some regions of dyad symmetry potentially involved in btuB regulation. The locations of these segments are identified in Fig. 2. Stability constants were calculated by using the method and parameters of Tinoco et al. (21); 1 kcal = 4.18 kJ. n't's, Nucleotides.