JOURNAL OF BACTERIOLOGY, Sept. 1983, p. 1279–1287
0021-9193/83/091279-09$02.00/0

Vol. 155, No. 3

# In Vivo 5′ Terminus and Length of the mRNA for the Proton-Translocating ATPase (*unc*) Operon of *Escherichia coli*

HELEN M. JONES, CONNIE M. BRAJKOVICH, AND ROBERT P. GUNSALUS*

*Department of Microbiology and the Molecular Biology Institute, University of California, Los Angeles, California 90024*

Received 2 May 1983/Accepted 27 June 1983

The promoter for the proton-translocating ATPase (*unc*) operon of *Escherichia coli* was localized by using a plasmid promoter-screening vector system. S1 nuclease analysis, using the appropriate single-stranded DNA probe from this promoter region and in vivo mRNA, revealed that the 5′ end of the in vivo *unc* mRNA initiates with a guanine residue 73 bases before the start of the proposed gene 1 or 474 bases before *uncB*. An in vivo *unc* mRNA species of approximately 7,000 nucleotides in length which initiates in the *unc* promoter region was shown to exist by RNA-DNA hybridization analysis. This *unc* mRNA species (based on DNA sequence analysis) is sufficient in length to contain all nine genes, gene 1 and *uncBEFHAGDC*. That gene 1 is cotranscribed with the *unc* genes was confirmed by using hybridization probes containing the promoter-proximal (gene 1) or -distal gene (*uncC*). No strong internal promoters within the *unc* operon were revealed with either the promoter-screening vector system or the RNA-DNA hybridization analysis. The 5′ terminus and the length of the *unc* mRNA were found to be identical in cells grown either aerobically or anaerobically. The level of *unc* operon expression, as assayed with the *unc* promoter plasmid, did not significantly differ when cells bearing the plasmid were grown either aerobically or anaerobically.

The *unc* operon of *Escherichia coli* is located at 83.5 min on the genetic map (2) and contains the genes encoding the eight individual polypeptide types comprising the proton-translocating ATPase complex (8, 15). Upon expression, the *unc* polypeptides assemble into a macromolecular complex which is bound to the cytoplasmic membrane. The ATPase carries out the oxidative phosphorylation reactions coupled to proton import for ATP synthesis under aerobic growth conditions. Under anaerobic growth conditions, *E. coli* employs the ATP hydrolysis reaction of the complex coupled to proton extrusion from the cell to maintain the proton gradient across the cytoplasmic membrane for other cellular functions, including solute transport and motility. The *unc* gene and polypeptide orders have been shown to be *uncBEFHAGDC* and acbδαγβε, respectively (7, 15). The DNA sequence in the *unc* region has been determined (12, 13, 18, 20–22, 26, 32), and protein sequence studies with polypeptides, synthesized both in vivo and in vitro, confirm the translational start sites predicted from DNA sequence analysis for the known eight *unc* polypeptides (10, 17; R. D. Simoni, personal communication). In addition, DNA sequence analysis revealed an open reading frame upstream of *uncB* which may encode a protein of approximately 14 kilodaltons. This hypothetical gene has been designated gene 1 (13). Evidence for the existence of this gene is presented in an accompanying paper (6), although the function of the protein is presently unknown.

Little is known about *unc* operon regulation at the transcriptional or translational level. Early genetic studies by Gibson and co-workers (14) suggested that there may be a single transcriptional unit for the four *unc* operon genes then known, *uncBADC*. These studies involved the creation of polar mutations using phage Mu insertion. These mutants showed the loss of a functional ATPase and energy-linked activities, both of which could be complemented by using partial diploids. Second, DNase footprinting (19), in vitro runoff transcriptions (19) and Tn*10* insertions (36) suggested that *unc* promoters existed upstream of and within gene 1 and the *unc* genes, but whether any of these promoters acted in vivo to form *unc* messages was unknown. Furthermore, the physiological role of the ATPase complex differs when *E. coli* is grown aerobically or anaerobically, and it is not known whether this is accompanied by a change in *unc* gene regulation. Finally, regulation of expression must occur at the transcriptional or

translational level to produce the proper amounts of the individual ATPase polypeptide subunits. These are present in the mature complex at the stoichiometries of 1, 10, 2, 1, 3, 1, 3, 1, for a, c, b, δ, α, γ, β, ε, respectively (4, 11). In vitro and in vivo plasmid-directed protein synthesis studies also demonstrate this unequal polypeptide synthesis (5).

We examined the transcriptional regulation of the *E. coli unc* operon. The promoter was localized upstream of gene 1 by cloning restriction fragments containing *unc* sequences into plasmid expression vectors. No strong internal promoters, which might explain the differential polypeptide stoichiometry, were revealed. S1 nuclease protection experiments precisely determined the site of *unc* mRNA initiation in vivo. RNA-DNA hybridization experiments revealed a single major *unc* mRNA species of approximately 7 kilobases (kb) which initiates in the region localized above. The identical mRNA species hybridized to promoter-proximal and -distal genes. These results demonstrated the existence of a single major *unc* mRNA species containing the nine genes, gene 1 and *uncBEFHAGDC*, of the *unc* operon. S1 nuclease and RNA-DNA hybridization analyses revealed an identical *unc* mRNA in aerobically and anaerobically grown cells. Finally, by using a plasmid expression vector system, expression from the *unc* promoter was found to be similar in aerobically and anaerobically grown cells.

## MATERIALS AND METHODS

**Materials.** D-[$^{14}$C]galactose (0.04 to 0.06 Ci/mmol), [γ-$^{32}$P]ATP (3 Ci/mol), and [α-$^{32}$P]dATP (0.4 Ci/mol) were purchased from Amersham Corp., Arlington Heights, Ill. T4 polynucleotide kinase was obtained from PL Biochemicals, Inc., Milwaukee, Wis. S1 nuclease was obtained from Miles Laboratories, Inc., Elkhart, Ind. All other chemicals were of the highest purity commercially available.

**Bacterial strains and plasmids.** *E. coli* N100 (*galK recA pro lac*$^+$) was used for plasmid transformations, plasmid preparations, and galactokinase assays (25). *E. coli* LE392 (F$^-$ *hsdR514*($r_k^-m_k^-$) *supE44 supI58 lacYI or* Δ(*lacIZY*)*6 galK2 galT22 metB1 trpR55* lambda$^-$) (23), and a derivative of *E. coli* A19 (*rns* X$^-$ Δ(*tonB-trpAE*)*1 trpR* Δ (*uncB-D*), where X represents an unknown growth requirement satisfied by 0.05% acid hydrolyzed casein) (38) carrying a chromosomal deletion in *unc* were used for mRNA isolation. The *unc* deletion strain was a gift of W. S. A. Brusilow and was created by a crossover between an *unc* plasmid carrying a 5-kb *PstI* deletion between *uncB* and *uncD* and the chromosome. Plasmids pRPG26, pRPG27, pRPG53, pRPG54, (15), pKO-4, pKO-6, pKG1800 (25), and pUC9 (Bethesda Research Laboratories, Inc., Gaithersburg, Md.) were used as indicated.

**Construction of plasmids.** Endonuclease digestions, phosphatase treatment, ligation, analysis of restriction fragments, cell transformations, and DNA preparations were carried out as described previously (16).

**Galactokinase assay of extracts from plasmid-bearing cells.** Galactokinase assays were performed as described previously (25) with the following modifications. Cell-free extracts were prepared by harvesting 20 ml of a mid-log phase culture (60 Klett turbidity units; red filter), washing once (5 ml), and sonicating in 2 ml of 20 mM Tris-chloride (pH 8.3). Appropriately diluted cell-free extract (55 μl) was added to 220 μl of assay mix (12.5 mM Tris [pH 8.3], 0.682 mM galactose at 4.1 mCi/mmol, 6.25 mM ATP, 4 mM NaF, 1.25 mM dithiothreitol, 5 mM $MgCl_2$). The mixture was incubated at 30°C, and 50-μl portions were removed to Whatman DE81 filter paper disks (Whatman, Clifton, N.J.) which were resting on dry ice. The filters were washed five times with 1 ml of water, with 0.45-μm membrane filters (Millipore Corp., Bedford, Mass.) for support, dried, and counted in 5 ml of scintillation cocktail (Amersham ACS; Amersham Corp.). The specific activity of the label was determined by counting 2-μl portions of the assay mix in vials containing a DE81 filter. The galactokinase units were determined from the slope of the curve resulting from the plot of nanomoles of D-galactose-1-phosphate formed per milligram of protein versus time. Protein concentration was determined by using Bradford reagent (3) with bovine serum albumin as a standard.

**Preparation of in vivo mRNA.** Wild-type *E. coli* mRNA from aerobically and anaerobically grown cells and *E. coli* mRNA from an aerobically grown *unc* deletion strain were prepared as described previously (31).

**S1 protection experiments and DNA sequencing analysis.** Phosphatase-treated DNA fragments (the *Sau*3A-*Hin*dIII$_{241}$ fragment of pHJ.UP11 and the *Hin*dIII-*Bam*HI$_{240}$ fragment of pHJ.UP3) were treated with kinase by the method of Nichols (B. Nichols, personal communication). Briefly, DNA was dissolved in 5 μl of denaturation buffer (5 mM Tris-chloride [pH 9.5], 1 mM spermidine, 0.1 mM EDTA), heated at 95°C for 2 min, and cooled in ice water for 1 min. Kinase buffer (1 μl; 10× (0.5 M glycine-NaOH [pH 9.5], 0.1 M $MgCl_2$, 0.05 M dithiothreitol, 35% glycerol), 4 μl of [γ-$^{32}$P]ATP (10 mCi/ml), and 0.5 μl of T4 polynucleotide kinase (5 U/μl) were added to the DNA, and the mixture was incubated at 37°C for 1 h. The DNA was then ethanol precipitated, and the pellet was washed three times with 75% ethanol and dried. Strand separation and DNA sequencing of the labeled fragments were performed as described previously (24). mRNA-DNA hybridizations were performed basically as described previously (27). Briefly, total bacterial mRNA (80 to 100 μg; 3 mg/ml) was mixed with a 5′ $^{32}$P-labeled single-stranded DNA fragment (0.8 to 1.0 μg) (described above) in 50 μl of 10 mM Tris (pH 7.2)–1 mM EDTA–1 M NaCl. The mixture was heated at 75°C for 10 min and then incubated at 68°C for 30 min. The hybridization mixtures were chilled, 500 μl of S1 buffer (0.25 M NaCl, 0.03 M sodium acetate [pH 4.5], 1 mM $ZnSO_4$, 5% glycerol), and 5,000 U of S1 nuclease were added, and the resulting mixture was incubated at 15°C for 2 h. The mixture was ethanol precipitated, washed three times with 75% ethanol, and dried. DNA pellets from the sequencing and S1 protection experiments were suspended in formamide buffer and electrophoresed on an 8% polyacrylamide–8 M urea sequencing gel as described previously (24).

**RNA-DNA hybridization analysis.** The nick-translat-

ed probes (plasmids pRPG26, pRPG27, pRPG53, pRPG54, pBR322, and pRPG5 DNA) were prepared as described previously (29). After the reaction (14°C for 2 h), the DNA was extracted once with phenol-chloroform (saturated with 10 mM Tris, 1 mM EDTA [pH 7.5]), and the resulting aqueous phase was extracted two times with 50% water-saturated ether. The reaction mixture was not run over a Sephadex G-50 desalting column as suggested by Rigby et al. (29) since no difference in the hybridization results was found when this step was omitted. Total *E. coli* mRNA was denatured with glyoxal as described previously (34) and electrophoresed on a 1% agarose gel (12.5 by 19 by 0.6 cm) in 1× MAE buffer (20 mM MOPS [morpholinepropanesulfonic acid], 5 mM sodium acetate, 1 mM EDTA [pH 7.0]) at 40 mA for approximately 4 h. The buffer was continuously recirculated to maintain the pH at 7.0. Gel lanes to be stained with ethidium bromide were removed from the rest of the gel, treated with 50 mM NaOH for 45 min, neutralized in 0.5 M Tris (pH 7.5)–1 μg of ethidium bromide per ml for 20 min, and photographed. The remaining gel was immediately transferred to a Genatran membrane (Genatran 45; DL4510RG; D+L Filter Corp., Woburn, Mass.) as described previously (37). The resulting "blot" was baked, pretreated, and hybridized essentially as described previously (34) with the following exception: the hybridizations were carried out at 42°C for 24 or more h in a mixture containing $2 \times 10^6$ cpm of [α-$^{32}$P]dATP-labeled probe (described above) in 5 to 10 ml of hybridization buffer (without dextran sulfate) per three to four lanes of mRNA. The membranes were washed and autoradiographed as described previously (34).

**Anaerobic cell growth.** Cells were grown anaerobically in Erlenmeyer flasks placed within Brewer jars. The atmosphere, composed of carbon dioxide and hydrogen, was generated by a GasPak (BBL Microbiology Systems, Cockeysville, Md.) with a palladium catalyst. Methylene blue oxygen indicator strips (BBL) confirmed the absence of oxygen in the chamber. The apparatus was assembled to allow cell samples to be removed for optical density measurements without disturbing the anaerobic environment.

**Media.** Vogel-Bonner minimal medium (35) supplemented with 0.5% glucose, 1% Casamino Acids, 30 μg of tryptophan per ml, 60 μg of methionine per ml, 1 μg of thiamine per ml, and 0.01 mM $FeCl_3$ was used in the experiments described in Tables 2 and 3. Luria broth plates (16) and MacConkey-galactose (1%) plates (Difco Laboratories, Detroit, Mich.) were used for the initial isolations and screenings. Ampicillin (Amp), obtained from Sigma Chemical Co., St. Louis, Mo., was used at 30 μg/ml.

## RESULTS

**Location of the *unc* transcriptional control region.** We have cloned restriction fragments containing various regions of the *unc* operon into the pKO *galK* expression vectors (25). Transformation of *E. coli* N100 with these plasmids leads to expression of the downstream *galK* gene dependent on the newly inserted promoter region. Levels of the *galK* gene product (galactokinase) present in cell-free extracts prepared from these transformed strains then indicate the presence and strength of a promoter site in the newly inserted sequence. The plasmid constructions containing the various *unc* operon restriction fragments are described in Fig. 1. Briefly,

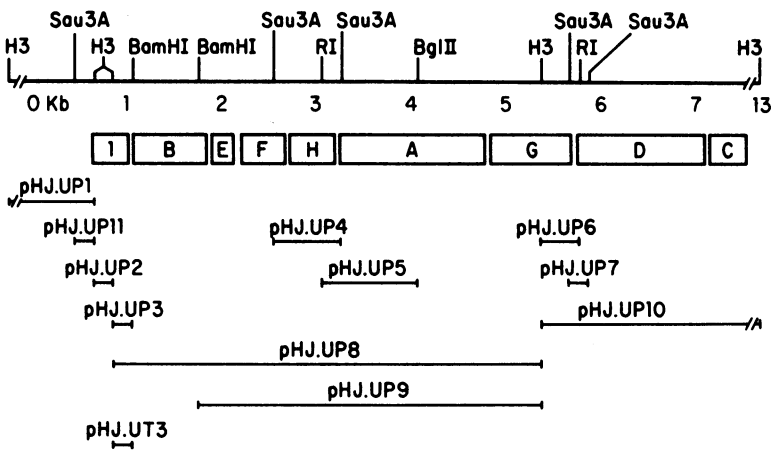


FIG. 1. Restriction map of galactokinase expression vectors containing various regions of the *unc* operon. The upper line shows the location of the pertinent restriction endonuclease recognition sites with distance given in kb. The location of the *unc* genes within this region is boxed. The operon is transcribed from left to right as drawn. The ends of the lines indicate the location and size of the individual restriction fragments within the operon used for cloning into the appropriate promoter (pKO) and terminator (pKG1800) screening vectors. The vector used for each plasmid, the restriction fragment and size, and the plasmid size in kb are shown in Table 1. All plasmids confer Amp resistance. Plasmids were constructed by standard methods (16). The restriction endonuclease *Hin*dIII and *Eco*RI sites are abbreviated as H3 and RI, respectively. Plasmid pHJ.UT3 was constructed through an intermediate vector, pUC9, which provided the restriction sites necessary for cloning into pKG1800.

TABLE 1. Promoter test plasmids used for in vivo galactokinase assay

| Plasmid | *unc* fragment[a] | Genes encoded[b] | Plasmid vector | Total size (kb) |
|---|---|---|---|---|
| pHJ.UP1 | *Hin*dIII$_{1300}$ | Gene 1′ | pKO-4 | 5.3 |
| pHJ.UP2 | *Hin*dIII$_{199}$ | Gene 1′ | pKO-4 | 4.2 |
| pHJ.UP3 | *Hin*dIII/*Bam*HI$_{240}$ | Gene 1′*uncB*′ | pKO-4 | 4.2 |
| pHJ.UP4 | *Sau*3A$_{662}$ | *uncF*′*HA*′ | pKO-4 | 4.7 |
| pHJ.UP5 | *Eco*RI/*Bgl*II$_{807}$ | *uncA*′ | pKO-4 | 4.8 |
| pHJ.UP6 | *Hin*dIII/*Eco*RI$_{440}$ | *uncG*′*D*′ | pKO-6 | 4.4 |
| pHJ.UP7 | *Sau*3A$_{356}$ | *uncG*′*D*′ | pKO-4 | 4.3 |
| pHJ.UP8 | *Hin*dIII$_{4300}$ | Gene 1′*uncBEFHAG*′ | pKO-4 | 8.3 |
| pHJ.UP9 | *Bam*HI/*Hin*dIII$_{3580}$ | *uncB*′*EFHAG*′ | pKO-6 | 7.6 |
| pHJ.UP10 | *Hin*dIII$_{8500}$ | *uncG*′*DC* | pKO-4 | 12.5 |
| pHJ.UP11 | *Sau*3A/*Hin*dIII$_{241}$ | Gene 1′ | pKO-6 | 4.2 |
| pHJ.UT3 | *Hin*dIII*Bam*HI$_{240}$ | Gene 1′*uncB*′ | pKG1800 | 5.0 |

[a] Subscripts are fragment sizes (bp).
[b] A prime (′) indicates a partial gene.

the individual restriction fragments containing *unc* sequences were ligated into the indicated *galK* vectors (Table 1), and the resulting plasmids were transformed into *E. coli* N100 and plated on L agar-Amp plates. Isolated colonies were then screened for *galK* expression by replica plating onto MacConkey-galactose-Amp indicator plates. Colonies containing plasmids pKG1800, pHJ.UP1, and pHJ.UP11 appeared red due to acid production from galactose utilization and indicated the presence of an inserted promoter sequence. In the positive control strain containing pKG1800, the *gal* promoter is present and transcribes the downstream *galK* gene responsible for the first step in galactose utilization catalyzed by galactokinase. Colonies bearing plasmids pHJ.UP3, pHJ.UP8, and pHJ.UP9 appeared pink, indicating a low or intermediate level of *galK* expression. Colonies bearing pHJ.UP2, pHJ.UP4, pHJ.UP5, pHJ.UP6, pHJ.UP7, and pHJ.UP10 appeared white on the MacConkey-galactose indicator medium, as did the strain containing the parental pKO-4 vector which contains no promoter sequence for *galK* expression. Plasmid DNA was prepared from colonies picked from the nonselective L agar-AMP plates. Restriction enzyme analysis was performed to confirm the presence of the *unc* inserts and their respective orientation within the newly constructed expression test vectors. The cloned *unc* fragment sizes as determined by agarose and polyacrylamide gel analyses are shown in Table 1 and are consistent with those expected from published DNA restriction and sequence data (12, 13, 15, 18, 20–22, 26, 32). For plasmids pHJ.UP3 and pHJ.UP11, the inserted sequence was confirmed by DNA sequence analysis (see below).

**Assay of *galK* expression in promoter test plasmid-bearing strains.** Galactokinase (*galK*) expression was determined by measuring the activities in cell extracts prepared from each of the transformed strains. The strains were grown aerobically to the mid-exponential growth phase and harvested by centrifugation, and cell-free extracts were prepared. The galactokinase levels in cell-free extracts from the *unc* promoter test plasmid relative to the control plasmid-containing strains are shown in Table 2. The strong *gal* operon promoter (pKG1800) gives 414.3 galactokinase units relative to the promoter-negative pKO-4 vector (6.7 U) under the assay conditions employed. The *unc* fragment-containing vector pHJ.UP1 shows strong (274.7 U) galactokinase expression, as was initially indicated by the red color on the MacConkey-galactose indicator plates. These results suggest that the putative *unc* promoter resides on the 1.3-kb *Hin*dIII fragment of pHJ.UP1. The location of this promoter was more precisely determined by subcloning the *Sau*3A/*Hin*dIII$_{241}$ frag-

TABLE 2. Galactokinase activity of extracts prepared from the *unc* promoter test plasmid strains

| Plasmid | Galactokinase (U)[a] |
|---|---|
| pHJ.UP1 | 274.7 |
| pHJ.UP11 | 284.7 |
| pHJ.UP2 | 13.9 |
| pHJ.UP3 | 18.0 |
| pHJ.UP4 | 16.5 |
| pHJ.UP5 | 9.3 |
| pHJ.UP6 | 9.7 |
| pHJ.UP7 | 9.2 |
| pHJ.UP8 | 26.3 |
| pHJ.UP9 | 27.0 |
| pHJ.UP10 | 5.0 |
| pKO-4 | 6.7 |
| pKG1800 | 414.3 |
| pHJ.UT3 | 395.9 |

[a] Nanomoles of D-galactose-1-phosphate formed per minute per milligram of protein.

ment of pHJ.UP1 into pKO-6, creating pHJ.UP11. Cell extracts of pHJ.UP11-transformed N100 exhibited galactokinase levels (284.7 U) comparable to that of the pHJ.UP1-containing strain (274.7 U). Evaluation of *unc* sequences downstream from this region within the operon using the *galK* expression test vectors did not reveal any additional strong promoter-containing sequences (Fig. 1 and Table 2). For example, the 240-base pair (bp) *Hin*dIII-/*Bam*HI restriction fragment containing the first known structural gene (*uncB*) of the operon and adjoining regions just upstream, located on pHJ.UP3, shows little, if any, promoter activity (18.0 U) in the galactokinase assay, as was suggested by the pink color on the indicator plates. This putative promoter region, if effectively used for initiation of *unc* transcription, would account for less than 4% *unc* message production relative to the apparent *unc* promoter contained on pHJ.UP11.

**Analysis of *unc* sequences for termination signals.** To determine whether the strong promoter contained on pHJ.UP11 represents the exclusive gene 1 promoter, we screened for the presence of a transcriptional stop signal between gene 1 and *uncB*. The *Hin*dIII/*Bam*HI$_{240}$ fragment from pHJ.UP3, which contains the end of gene 1 and the beginning of *uncB*, was cloned into the pKG1800 terminator screening vector, creating pHJ.UT3 (Fig. 1 and Table 1). The insertion of a terminator-containing fragment into pKG1800 between the *gal* operon promoter and the *galK* gene would result in a decreased level of galactokinase expression with respect to the parental plasmid in transformed strains. Cells bearing this plasmid (pHJ.UT3) appeared red on MacConkey-galactose indicator plates, and restriction enzyme analysis confirmed the plasmid construction. Galactokinase assay of pHJ.UT3-transformed cells (Table 2) revealed high levels of galactokinase (395.9 U), comparable to that of the parent pKG1800-containing cells (414.3 U), thereby indicating the lack of a terminator signal sequence between gene 1 and *uncB*.

**Determination of the in vivo mRNA initiation site.** S1 nuclease analysis was carried out to determine precisely the initiation site of the *unc* mRNA in vivo. The *Sau*3A/*Hin*dIII$_{241}$ fragment of pHJ.UP11 was 5′ end labeled, the two strands were separated, and each strand was sequenced to identify the strand complementary to the *unc* mRNA. The appropriate strand was hybridized to excess in vivo mRNA, S1 nuclease treated, and electrophoresed adjacent to a DNA sequencing "ladder" for this strand. mRNA prepared from both aerobically and anaerobically grown cells was used for this analysis (Fig. 2). S1 nuclease protection of the DNA up to the C residue corresponding to the G residue in the

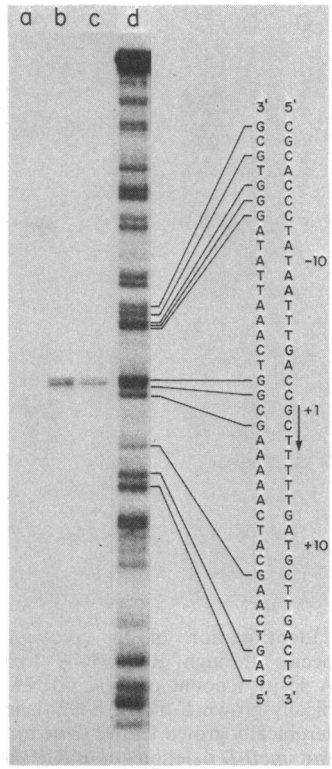


FIG. 2. Location of the in vivo 5′ start for *unc* mRNA. The single-stranded *Sau*3A/*Hin*dIII$_{241}$ fragment of pHJ.UP11 labeled at the 5′ terminus was hybridized to mRNA and treated with S1 nuclease. Lanes b and c are hybridizations with mRNA prepared from an aerobically and an anaerobically grown wild-type *E. coli* strain, LE392, respectively. Lane a represents treatment in the absence of mRNA. In lane d is the G sequencing reaction from the *Sau*3A/*Hin*dIII$_{241}$ fragment.

mRNA indicated the 5′ terminus of the *unc* mRNA to be 73 bases before the initiator codon for gene 1. The protected fragment ran slower than the corresponding band in the sequencing gel since S1 cleavage yielded a 3′-hydroxyl group on the protected terminal G residue, whereas the chemical degradation in the Maxam and Gilbert sequencing reaction destroyed the G residue, leaving a terminal 3′-phosphate group (1, 9). The *unc* mRNA start site was identical in cells grown aerobically and anaerobically (Fig. 2, lanes b and c). Similar S1 nuclease protection experiments were performed with the *Hin*dIII/*Bam*HI$_{240}$ fragment of pHJ.UP3, which contains minor promoter activity (see above). Greater than 95% of the radioactivity was found in the fully protected species, consistent with the indication that the principal promoter is on the pHJ.UP11-*unc* fragment.

**RNA-DNA hybridization analysis of in vivo *unc***

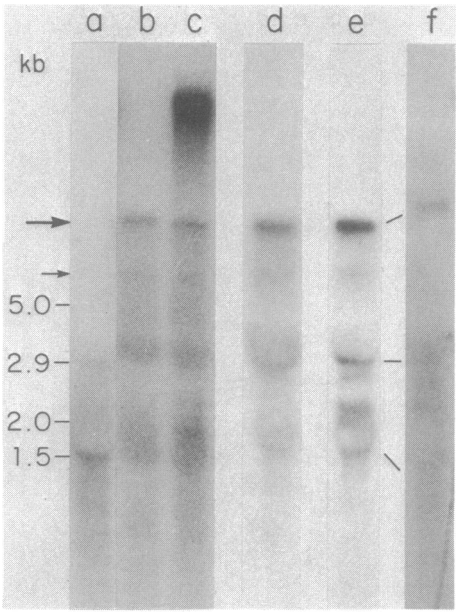


FIG. 3. Identification of in vivo *unc*-specific mRNA species through RNA-DNA hybridization analysis. A 15-μg amount of total mRNA prepared from aerobically grown *E. coli* LE392 (lanes b, d, e, and f), anaerobically grown LE392 (lane c), and aerobically grown *uncB-D* deletion strain A19 (lane a) was denatured with glyoxal, fractionated on a 1% agarose gel, and transferred to a Genatran membrane. Sections of the membrane were hybridized with the nick-translated DNA probes pRPG54, pRPG27, pRPG53, and pRPG26, which are specific to *unc* DNA. Lanes a, b, and c, mRNAs described above probed with pRPG54, which contains approximately 6.7 kb of *unc* DNA, including genes 1'*BEFHAGDC*; lane d, mRNA probed with pRPG27, which contains 199 bases of DNA sequence internal to gene 1; lane e, mRNA probed with pRPG53, which contains a 1.18-kb DNA fragment including *uncD'C*; lane f, mRNA probed with pRPG26, which contains the *unc* promoter and only 122 bases of DNA which can hybridize to an *unc* message. Numbers to the left of the panel are approximate procaryotic (*E. coli*) and eucaryotic (chicken erythrocyte) rRNA size standards. Light hybridization to the procaryotic 23S and 16S rRNAs is seen with *unc*-specific and nonspecific, i.e., pBR322, probes. The intensity of the bands varies with the particular message preparation and its age. The upper arrow indicates the major full-length *unc* mRNA species, and the lower arrow indicates a minor, shortened *unc* transcript. The high-molecular-weight band in lane c is chromosomal DNA in the mRNA preparation. This band is not present in DNase-treated mRNA preparations.

**mRNA.** Plasmids pRPG26, pRPG27, pRPG54, and pRPG53 containing DNA regions encoding the amino terminus of gene 1 protein, the carboxy terminus of gene 1 protein, *uncBEFHAGDC*, and *uncD'C*, respectively, were used as probes for RNA-DNA hybridization analysis. In vivo mRNA was prepared from both a wild-type *E. coli* strain, LE392, and an *E. coli* strain (A19) carrying an *uncB-D* deletion. In addition, mRNA was prepared from the anaerobically grown wild-type LE392 strain. The hybridization results are shown in Fig. 3. A single major mRNA species approximately 7,000 nucleotides in length appears in both the wild-type aerobic and anaerobic mRNA (Fig. 3, upper arrow, lanes b and c). This species is absent in mRNA prepared from the *unc* chromosomal deletion strain (Fig. 3, lane a). The same mRNA species was observed when probed with promoter-proximal (pRPG26 and pRPG27; Fig. 3, lanes f and d, respectively) or -distal (pRPG53; Fig. 3, lane e) genes. These data are consistent with the existence of a major *unc* mRNA species which initiates at the S1 nuclease predicted site before gene 1, continues through *uncBEFHAGDC*, and terminates after *uncC*. This *unc* mRNA species appears to be identical in cells grown aerobically and anaerobically by the criteria of mRNA start site (see above) and message length. A minor mRNA species of 5 to 6 kb appears in the mRNA lanes probed with pRPG27, pRPG54, and pRPG53, but not noticeably with pRPG26 (Fig. 3, lower arrow, lanes b to e). The distance on the chromosome between the *unc* DNA contained on pRPG27 and pRPG53 is approximately 5.7 kb. This suggests that a minor *unc* mRNA species may begin in the DNA region internal to gene 1 and end within or just after *uncD*. Nick-translated pBR322 (the vector for plasmids pRPG26, pRPG27, and pRPG53) did not hybridize to any mRNA species, and therefore the messenger species identified do not result from nonspecific binding (data not shown). mRNA prepared from the *unc* deletion strain when probed with pRPG5, which contains the *trpR* gene, showed an mRNA species of the size

TABLE 3. *unc*-directed galactokinase expression in aerobically versus anaerobically grown cells[a]

| Plasmid | Galactokinase (U)[b] | |
|---|---|---|
| | Aerobic | Anaerobic |
| pHJ.UP1 | 274.7 | 310.8 |
| pKO-4 | 6.7 | 7.4 |
| pKG1800 | 414.3 | 413.2 |
| pHJ.UP3 | 18.0 | 20.8 |

[a] Analyses were performed with plasmid-bearing strain N100 grown aerobically or anaerobically on Vogel-Bonner minimal medium supplemented with 0.5% glucose and 1% Casamino Acids as described in the text. The plasmids did not affect doubling time either aerobically (53 ± 3 min) or anaerobically (91 ± 3 min) significantly.

[b] Nanomoles of D-galactose-1-phosphate formed per minute per milligram of protein.

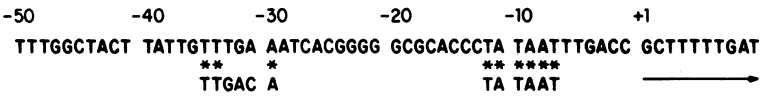


FIG. 4. Nucleotide sequence of the *unc* promoter region. Sequences are numbered relative to the 5′ terminus of the *unc* mRNA (see text). The guanine residue located 73 bp preceding the presumed site of gene 1 translation and 474 bp preceding the site of *uncB* translation is given as position +1. The consensus sequences of the RNA polymerase recognition sequences in the -35 and -10 regions are shown. Asterisks indicate homology with the consensus sequence (30, 33).

expected for the *trpR* transcript, confirming that the *unc* deletion strain mRNA preparation had not degraded (data not shown).

**Relative levels of *unc* transcription under different growth conditions.** The *unc* promoter-dependent *galK* expression plasmids described above provide a convenient way to measure the effect of cell growth conditions on *unc* expression at the transcriptional level. We examined galactokinase activity in the *unc* promoter-dependent, *galK* expression plasmid-bearing strain under aerobic and anaerobic growth conditions (Table 3). No significant difference in the galactokinase levels was seen in the N100 strain containing pHJ.UP1. Control experiments with the vector (pKO-4), the plasmid containing the *gal* operon promoter (pKG1800), and the plasmid containing the *unc* fragment just before *uncB* (pHJ.UP3) also did not show any significant variation in galactokinase expression in N100 cells grown aerobically and anaerobically. The expression of *galK* from these plasmids appears to remain independent of growth rate or growth mode.

## DISCUSSION

We examined the in vivo transcriptional regulation of the *E. coli unc* operon. The *unc* promoter region was initially localized by using a promoter-screening vector system in which the presence of a promoter within a given restriction fragment leads to the expression of the *galK* gene upon proper insertion into the vector. A single major promoter for the *unc* operon existed in a DNA fragment which lies upstream of *uncB* and which contains the first 49 bp of the 390-bp open reading frame designated gene 1. The precise promoter location was determined to be 73 bp before the start of gene 1 by S1 nuclease protection experiments performed with mRNA synthesized in vivo. The nucleotide sequence in the *unc* promoter region spanning the in vivo start of transcription is shown in Fig. 4. This determination of the *unc* mRNA initiation site with in vivo mRNA confirms the promoter location determined with in vitro transcription studies (19, 28), DNase footprinting (19) and S1 nuclease analysis performed with plasmid-directed mRNA (28). The promoter-screening vector system failed to localize any strong internal promoters for *unc* specific transcription (Table 2). However, we cannot exclude the possibility that several very weak internal *unc* promoters may exist, as evidenced by levels of galactokinase expression slightly above background. These results are consistent with the observation that Tn*10* insertions which affect *unc* operon expression lie within a 0.4-kb region upstream from *uncB* (36). Tn*10* analysis also suggested the presence of several very low-level internal promoters within gene 1 (36). Analysis for promoters in this region (Table 2) suggests that the relative strength of such sequences would be no more than 4% when compared to the major *unc* promoter located 73 bp before gene 1 (Table 2).

A single major in vivo *unc* mRNA species approximately 7,000 nucleotides in length was identified by RNA-DNA hybridization analysis. This mRNA species was shown to contain the initial gene 1 region and the eight known *unc* genes, *uncB* through *uncC*. Identical RNA-DNA hybridization patterns were seen when plasmids containing either the gene 1 region or the last gene in the operon, *uncC*, were used as probes. These results demonstrate that gene 1 is cotranscribed with the other eight *unc* genes and must be, as a result, under the same regulatory control. The length of this mRNA is just sufficient to contain all nine genes, gene 1 and *uncBEFHAGDC*, of the proton-translocating ATPase operon. The distance from the *unc* mRNA initiation site to the termination codon for the *uncC* gene product is 6,952 bases. The absence of a terminator signal between gene 1 and *uncB*, as shown by the terminator-screening vector analysis, and the full protection of this DNA region by *unc* mRNA from S1 nuclease digestion provide further evidence that gene 1 is transcribed with the *unc* message. The role of gene 1 and its relationship with the other *unc* genes is unknown at present. The gene 1 product may have a regulatory role in gene expression or may be somehow involved in the assembly or function of the ATPase complex (13). Although a polypeptide of the size expected for the gene 1 product has never been isolated with in vivo ATPase complex, evidence presented in an accompanying paper (6) suggests that such a gene 1 protein product is made. A minor *unc* mRNA

species of 5 to 6 kb was identified which initiated in the DNA region internal to gene 1 (Fig. 3, lower arrow) and was sufficient in length, as determined by DNA sequence analysis, to span *uncBEFHAG* and to terminate within or after *uncD*. This minor mRNA species may arise as a stable degradation product of the full-length *unc* mRNA species or, alternatively, may be a separate transcript initiating from a low-level promoter within gene 1. As determined with the promoter test plasmids, the DNA region internal to gene 1 expressed galactokinase at levels less than 4% of those of the *unc* promoter but still slightly above background levels. The role of this minor *unc* message is presently unclear.

The demonstration of a single major *unc* message confirms the early genetic work by Gibson et al. (14) which proposed the existence of a single transcriptional unit for the then known *unc* genes, *uncBADC*.

The existence of a single major *unc* mRNA species and the corroborative lack of strong internal *unc* promoters suggest that the differential stoichiometries seen with the polypeptide components of the ATPase complex cannot be the result of transcriptional regulation. Rather, translational regulation, possibly in the form of differing ribosomal binding site affinities may explain the *unc* polypeptide ratios of 1, 10, 2, 1, 3, 1, 3, and 1 for the ATPase complex components a, c, b, δ, α, γ, β, ϵ, respectively (4, 11). An alternative explanation for the differential translational expression of the eight *unc* genes could involve the selective decay of the full-length *unc* mRNA to stable intermediates containing the genes most highly expressed (*uncE*, *A*, and *D*, which encode c, α, and β polypeptides, respectively). Since only a very minor secondary message was found which could encode at least six of the nine *unc* genes, this alternative explanation seems less likely, considering the order and locations of the abundantly expressed genes.

The *unc*-specific mRNA transcript is identical in *E. coli* cells grown aerobically and anaerobically. This is based on the criteria of 5′ start site and total message length as determined by S1 nuclease analysis and RNA-DNA hybridization analysis, respectively. The expression of the *unc* operon, at the transcriptional level, does not differ significantly when cells bearing the *unc* promoter plasmid are grown either aerobically or anaerobically. This assay of *unc* promoter expression was performed with the multicopy *galK* plasmid system, and the possibility cannot be ruled out that any *trans*-acting regulatory proteins, which may alter *unc* expression in these growth conditions, may have been titrated out. However, when the *trp* and *aroH* operon promoter-containing expression plasmids were assayed in the presence and absence of L-tryptophan, galactokinase levels, which reflect the promoter operator control, were observed to vary 5- to 10-fold. These results suggest that the *unc* operon is not differentially regulated under aerobic or anaerobic conditions at the transcriptional level in plasmid-bearing cells. Studies of in vivo *unc* mRNA levels under both growth conditions must be done to determine this conclusively.


## ACKNOWLEDGMENTS

We thank W. Brusilow and R. D. Simoni for the *E. coli unc* deletion strain and A. Tobin for rRNA size standards. We thank W. Romig and D. Nierlich for reviewing the manuscript. We also thank George Carlone for his advice and the use of his anaerobic cell culture apparatus.

This work was supported by Public Health Service grant GM-29456 from the National Institute of General Medical Sciences and the UCLA Biomedical Research Program. H.M.J. was supported by cellular and molecular biology training grant GM-07185 from the National Institutes of Health.


## LITERATURE CITED


1. **Aiba, H.** 1983. Autoregulation of the *Escherichia coli crp* gene: CRP is a transcriptional repressor for its own gene. Cell **32**:141–149.
2. **Bachmann, B. J., and K. B. Low.** 1980. Linkage map of *Escherichia coli* K-12, edition 6. Microbiol. Rev. **44**:1–56.
3. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–254.
4. **Bragg, P. D., and C. Hou.** 1975. Subunit composition, function, and spatial arrangement in the $Ca^{2+}$- and $Mg^{2+}$-activated adenosine triphosphatases of *Escherichia coli* and *Salmonella typhimurium*. Arch. Biochem. Biophys. **167**:311–321.
5. **Brusilow, W. S. A., D. J. Klionsky, and R. D. Simoni.** 1982. Differential polypeptide synthesis of the proton-translocating ATPase of *Escherichia coli*. J. Bacteriol. **151**:1363–1371.
6. **Brusilow, W. S. A., A. C. G. Porter, and R. D. Simoni.** 1983. Cloning and expression of *uncI*, the first gene of the *unc* operon of *Escherichia coli*. J. Bacteriol. **155**:●●●–●●●.
7. **Downie, J. A., G. B. Cox, L. Langman, G. Ash, M. Becker, and F. Gibson.** 1981. Three genes coding for subunits of the membrane sector ($F_0$) of the *Escherichia coli* adenosine triphosphatase complex. J. Bacteriol. **145**:200–210.
8. **Downie, J. A., F. Gibson, and G. B. Cox.** 1979. Membrane adenosine triphosphatases of prokaryotic cells. Annu. Rev. Biochem. **48**:103–131.
9. **Dunn, R., J. McCoy, M. Simsek, A. Majumdar, S. H. Chang, U. L. RajBhandary, and H. G. Khorana.** 1981. The bacteriorhodopsin gene. Proc. Natl. Acad. Sci. U.S.A. **78**:6744–6748.
10. **Dunn, S. D., L. A. Heppel, and C. S. Fullmer.** 1980. The $NH_2$-terminal portion of the α-subunit of *Escherichia coli* $F_1$ ATPase is required for binding the δ-subunit. J. Biol. Chem. **255**:6891–6896.
11. **Foster, D. L., and R. H. Fillingame.** 1982. Stoichiometry of subunits in the $H^+$-ATPase complex of *Escherichia coli*. J. Biol. Chem. **257**:2009–2015.
12. **Gay, N. J., and J. E. Walker.** 1981. The *atp* operon: nucleotide sequence of the region encoding the α-subunit of *Escherichia coli* ATP-synthetase. Nucleic Acids Res. **9**:2187–2194.
13. **Gay, N. J., and J. E. Walker.** 1981. The *atp* operon: nucleotide sequence of the promoter and the genes for the

membrane proteins, and the δ-subunit of *Escherichia coli* ATP-synthase. Nucleic Acids Res. **9**:3919–3926.
14. **Gibson, F., J. A. Downie, G. B. Cox, and J. Radik.** 1978. Mu-induced polarity in the *unc* operon of *Escherichia coli*. J. Bacteriol. **134**:728–736.
15. **Gunsalus, R. P., W. S. A. Brusilow, and R. D. Simoni.** 1982. Gene order and gene-polypeptide relationships of the proton-translocating ATPase operon (*unc*) of *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. **79**:320–324.
16. **Gunsalus, R. P., G. Zurawski, and C. Yanofsky.** 1979. Structural and functional analysis of cloned deoxyribonucleic acid containing the *trpR-thr* region of the *Escherichia coli* chromosome. J. Bacteriol. **140**:106–113.
17. **Hoppe, J., H. U. Schairer, and W. Sebald.** 1980. The proteolipid of a mutant ATPase from *Escherichia coli* defective in $H^+$-conduction contains a glycine instead of the carbodiimide-reactive aspartyl residue. FEBS Lett. **109**:107–111.
18. **Kanazawa, H., T. Kayano, K. Mabuchi, and M. Futai.** 1981. Nucleotide sequence of the genes coding for α-, β-, and γ-subunits of the proton-translocating ATPase of *Escherichia coli*. Biochem. Biophys. Res. Commun. **103**:604–612.
19. **Kanazawa, H., K. Mabuchi, and M. Futai.** 1982. Nucleotide sequence of the promoter region of the gene cluster for proton-translocating ATPase from *Escherichia coli* and identification of the active promoter. Biochem. Biophys. Res. Commun. **107**:568–575.
20. **Kanazawa, H., K. Mabuchi, T. Kayano, T. Noumi, T. Sekiya, and M. Futai.** 1981. Nucleotide sequence of the genes for $F_0$ components of the proton-translocating ATPase from *Escherichia coli*: prediction of the primary structure of $F_0$ subunits. Biochem. Biophys. Res. Commun. **103**:613–620.
21. **Kanazawa, H., K. Mabuchi, T. Kayano, F. Tamura, and M. Futai.** 1981. Nucleotide sequence of genes coding for dicyclohexylcarbodiimide-binding protein and the α-subunit of proton-translocating ATPase of *Escherichia coli*. Biochem. Biophys. Res. Commun. **100**:219–225.
22. **Mabuchi, K., H. Kanazawa, T. Kayano, and M. Futai.** 1981. Nucleotide sequence of the gene coding for the δ-subunit of proton-translocating ATPase of *Escherichia coli*. Biochem. Biophys. Res. Commun. **102**:172–179.
23. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 504. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labelled DNA with base specific chemical cleavages. Methods Enzymol. **65**:499–560.
25. **McKenney, K., H. Shimatake, D. Court, U. Schmeissner, C. Brady, and M. Rosenberg.** 1981. A system to study promoter and terminator signals recognized by *Escherichia coli* RNA polymerase, p. 383–415. *In* J. C. Chirikjian and T. S. Papas (ed.), Gene amplification and analysis, vol. II. Elsevier North-Holland, New York.
26. **Nielsen, J., F. G. Hansen, J. Hoppe, P. Friedl, and K. von Meyenberg.** 1981. The nucleotide sequence of the *atp* genes coding for the $F_0$ subunits a, b, c, and the $F_1$ subunit δ of the membrane bound ATP synthase of *Escherichia coli*. Mol. Gen. Genet. **184**:33–39.
27. **Osborne, T. F., R. B. Gaynor, and A. J. Berk.** 1982. The TATA homology and the mRNA 5′ untranslated sequence are not required for expression of essential adenovirus E1A functions. Cell **29**:139–148.
28. **Porter, A. C. G., W. S. A. Brusilow, and R. D. Simoni.** 1983. Promoter for the *unc* operon of *Escherichia coli*. J. Bacteriol. **155**:1271–1278.
29. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase 1. J. Mol. Biol. **113**:237–251.
30. **Rosenberg, M., and D. Court.** 1980. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13**:319–359.
31. **Salser, W., R. F. Gesteland, and A. Bolle.** 1967. In vitro synthesis of bacteriophage lysozyme. Nature (London) **215**:588–591.
32. **Saraste, M., N. J. Gay, A. Eberle, M. J. Runswick, and J. E. Walker.** 1981. The *atp* operon: nucleotide sequence of the genes for the γ-, β-, and ε-subunits of *Escherichia coli* ATP synthase. Nucleic Acids Res. **9**:5287–5296.
33. **Siebenlist, U., R. B. Simpson, and W. Gilbert.** 1980. *Escherichia coli* RNA polymerase interacts homologously with two different promoters. Cell **20**:269–281.
34. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. U.S.A. **77**:5201–5205.
35. **Vogel, H. J., and D. M. Bonner.** 1956. Acetylornithinease of *E. coli*: partial purification and some properties. J. Biol. Chem. **218**:97–106.
36. **von Meyenburg, K., B. B. Jorgensen, J. Nielsen, and F. G. Hansen.** 1982. Promoters of the *atp* operon coding for the membrane-bound ATP synthase of *Escherichia coli* mapped by Tn*10* insertion mutations. Mol. Gen. Genet. **188**:240–248.
37. **Wahl, G. M., M. Stern, and G. R. Stark.** 1979. Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl-paper and rapid hybridization by using dextran sulfate. Proc. Natl. Acad. Sci. U.S.A. **76**:3683–3687.
38. **Zalkin, H., C. Yanofsky, and C. L. Squires.** 1974. Regulated in vitro synthesis of *Escherichia coli* tryptophan operon messenger ribonucleic acid and enzyme. J. Biol. Chem. **249**:465–475.