*Proc. Natl. Acad. Sci. USA*
Vol. 75, No. 5, pp. 2190–2194, May 1978
Biochemistry

# Homologous nucleotide sequences between prokaryotic and eukaryotic mRNAs: The 5′-end sequence of the mRNA of the lipoprotein of the *Escherichia coli* outer membrane

(RNA sequence/ribosome-binding site/brome mosaic virus)

ROBERT M. PIRTLE, IRMA L. PIRTLE, AND MASAYORI INOUYE

Department of Biochemistry, State University of New York at Stony Brook, Stony Brook, New York 11794

*Communicated by R. B. Setlow, February 13, 1978*

ABSTRACT **The sequence of the first 89 nucleotides at the 5′ end of the mRNA for the lipoprotein of the *Escherichia coli* outer membrane is: GCUACAUGGAGAUUAACUCAAUCU-AGAGGGUAUUAAUAAUGAAAGCUACUAAACUGGUACU-GGGCGCGGUAAUCCUGGGUUCUACUCUG. The sequence of the first 72 nucleotides was established by direct sequencing methods and was extended to 89 residues on the basis of the known sequences of oligonucleotides obtained from complete digestion of the mRNA by ribonuclease $T_1$ or A and the known amino acid sequence of the prolipoprotein. The mRNA has an untranslated region of 38 residues before the initiation codon, AUG. A unique feature of the 5′-end sequence of the mRNA is that the sequence of 12 nucleotides (GUAUUAAUAAUG) prior to, and including, the initiation codon is the same as that found at the ribosome-binding site for 80S ribosomes in brome mosaic virus RNA4, a eukaryotic mRNA [Dasgupta, R., Shih, D., Saris, C. & Kaesberg, P. (1975) *Nature* 256, 624–628].**

The lipoprotein of the *Escherichia coli* outer membrane, an abundant and extensively studied structural protein, provides an excellent model system for the study of membrane biogenesis (see review, ref. 1). Studies of the product synthesized *in vivo*, as directed by the mRNA, indicated faithful translation in an *E. coli* cell-free system (2), as well as in a eukaryotic (wheat germ) cell-free system (3). The complete amino acid sequence of the precursor (prolipoprotein) synthesized in the *E. coli* cell-free system demonstrated that the prolipoprotein consisted of 78 amino acid residues (4). Because we had previously estimated by nucleotide sequence analysis (5) that the lipoprotein mRNA was 360 ± 10 nucleotides in length, we would expect to find approximately one-third of the total number of nucleotides in the noncoding regions of the mRNA molecule.

In this paper, the sequence of the first 89 nucleotides at the 5′ terminus of the mRNA is presented. It was found that there are 38 nucleotides preceding the initiation codon (AUG) in the untranslated region at the 5′ end of the mRNA. Remarkably, the nine nucleotides (GUAUUAAUA) preceding the initiation codon are identical to the nine nucleotides preceding the initiation codon in brome mosaic virus RNA4, a eukaryotic mRNA (6).

## MATERIALS AND METHODS

**Materials.** Polynucleotide kinase was obtained from PL Biochemicals. Calf intestine alkaline phosphatase was obtained from Boehringer Mannheim. The tetratriethylammonium salt of [$\gamma$-$^{32}$P]ATP was a product of ICN. Electrophoresis-grade acrylamide from Bio-Rad Laboratories was used for polyacrylamide gels. For autoradiography, Kodak Blue Brand medical x-ray film (BB-5) was used. Blue Max 2 (General Electric) x-ray screens were utilized for intensification of the radiation. The sources for other materials have been previously listed (5).

**Purification of the Nonradioactive Lipoprotein mRNA.** Purification of the nonradioactive lipoprotein mRNA was achieved by a procedure developed in this laboratory (2), except that the sucrose density gradient centrifugation step was omitted. Because highly purified mRNA was necessary for nucleotide sequence analysis, the mRNA-containing fraction from the Sephadex G-200 column was further purified by using reverse-phase column chromatography on RPC-5 (7).

**$^{32}$P-Labeling of the 5′ Terminus of the mRNA.** For dephosphorylation of the lipoprotein mRNA, 2 μg of RNA (85% pure) from the RPC-5 column was dissolved in 13.5 μl of 50 mM Tris·HCl (pH 8.2) containing 0.014 unit of calf intestine alkaline phosphatase and incubated at 45° for 45 min (8, 9). The reaction mixture was diluted with 40 μl of water and extracted with phenol/chloroform (10).

The $^{32}$P-labeling reaction was done by transferring the dephosphorylated mRNA [dissolved in 20 μl of 50 mM Tris·HCl, pH 8.0/10 mM $MgCl_2$/10 mM dithiothreitol/10% glycerol (vol/vol)/10 μg bovine serum albumin per ml/3.5 units of T4 polynucleotide kinase] into a siliconized test tube containing 160 pmol (1934 Ci/mmol) of lyophilized [$\gamma$-$^{32}$P]ATP (8–10). The reaction mixture was incubated at 37° for 30 min, diluted with 1.1 ml of 0.25 M sodium acetate (pH 5.2) followed by 2 μl of 0.1 M ATP (pH 7.0) and 20 μl of *E. coli* B tRNA (10 mg/ml) as carrier, and precipitated (5).

To determine if the untreated mRNA was phosphorylated, 1 μg of the mRNA from the RPC-5 column was subjected to the same $^{32}$P-labeling procedure using 80 pmol of [$\gamma$-$^{32}$P]ATP for the labeling reaction but without prior treatment with alkaline phosphatase to remove terminal phosphate.

The two mRNA samples were applied to a gel as described in the legend of Fig. 1. The $^{32}$P-labeled mRNA band was excised and extracted essentially as described (5) except that 175 μg of *E. coli* B carrier tRNA was added and the acid treatment was replaced with the following procedure. The RNA pellet obtained after the final ethanol precipitation was dissolved in 1.5 ml of 20 mM Tris·HCl, pH 7.5/0.1 M NaCl/1 mM EDTA. This solution was extracted twice, first with 3 ml and then with 1.5 ml of phenol/chloroform (1:1). The combined organic phases were back-extracted with 1 ml of fresh buffer and the aqueous phase was re-extracted with 2 ml of phenol/chloroform (1:1). The combined aqueous phases (about 2.5 ml) were subjected to two successive ethanol precipitations (5). The resulting RNA pellet was washed with 1 ml of cold ethanol, recentrifuged, and dried at reduced pressure over $P_2O_5$ before use.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

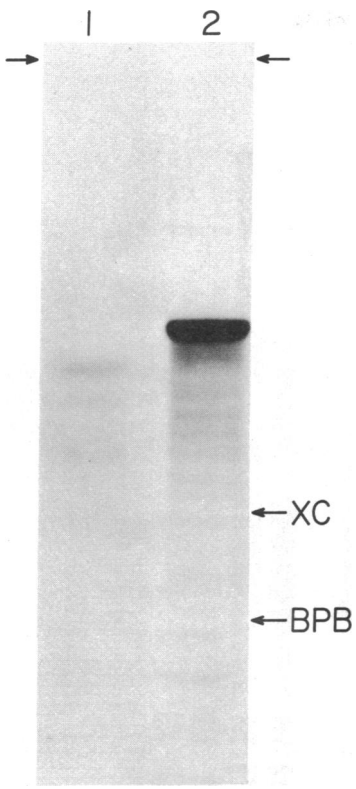



FIG. 1. Autoradiogram of [5′-$^{32}$P]mRNA after electrophoresis on a 5% polyacrylamide slab gel. Lane 1 contained 1 μg of mRNA subjected to the $^{32}$P-labeling procedure without prior treatment with alkaline phosphatase. Lane 2 contained 2 μg of mRNA that was dephosphorylated before labeling with [γ-$^{32}$P]ATP and polynucleotide kinase. The samples were dissolved in 40 μl of a 50% (wt/vol) urea solution containing 10 mM Tris·HCl (pH 7.5), 25 mM NaCl, and 4 mM EDTA; heated at 60° for 3 min; rapidly cooled; and mixed with 40 μl of dye solution containing 50 mM citric acid (pH 3.5), 0.04% (wt/vol) xylene cyanol, 0.04% (wt/vol) bromphenol blue, 20% (wt/vol) sucrose, and 7 M urea. The pH of the solutions was adjusted to 3.5 by addition of 20 μl of 0.25 M citric acid, and the samples were applied to the gel as described (5). Electrophoresis was carried out at 4° and 200 V for 15 hr in 25 mM citric acid (pH 3.5). The buffer in the lower reservoir contained 25 mM citric acid (pH 3.5) and 7 M urea. XC and BPB represent the respective positions of the xylene cyanol and bromphenol blue markers; the small arrows indicate the origin.

**Digestion of the 5′-End $^{32}$P-Labeled mRNA with Nuclease $P_1$.** In order to determine the 5′-end nucleotide of the mRNA and to examine the homogeneity of the mRNA, the following was done. Labeled mRNA (about $6.0 \times 10^4$ cpm) was completely digested by incubation for 1 hr at 37° in 12 μl of 50 mM ammonium acetate (pH 6.0) containing 1 μg of nuclease $P_1$ per 5 μg of RNA. A portion of this reaction mixture was fractionated by two-dimensional thin-layer chromatography (9) on a 20 × 20 cm cellulose plate (Brinkman CEL MN 300). Another portion was spotted on Whatman 3MM paper and electrophoresed in 5% (vol/vol) acetic acid adjusted to pH 3.5 with $NH_4OH$.

**Partial Nuclease $P_1$ Digestion of the [5′-$^{32}$P]mRNA for "Mobility Shift" Analysis.** [5′-$^{32}$P]mRNA (about $7.5 \times 10^5$ cpm) containing 39 μg *E. coli* B tRNA as carrier was incubated with 5.84 ng of nuclease $P_1$ in 50 μl of 50 mM ammonium acetate (pH 5.3) at room temperature; 10-μl aliquots were removed at 1, 2, 5, 10, and 20 min. The aliquots were mixed with 1.5 μl of 30 mM EDTA (pH 7.4), boiled for 3 min, rapidly cooled, pooled, and lyophilized to dryness (8). One-fourth of the digest was dissolved in 3 μl of water for electrophoresis–homochromotography as described in Fig. 2.

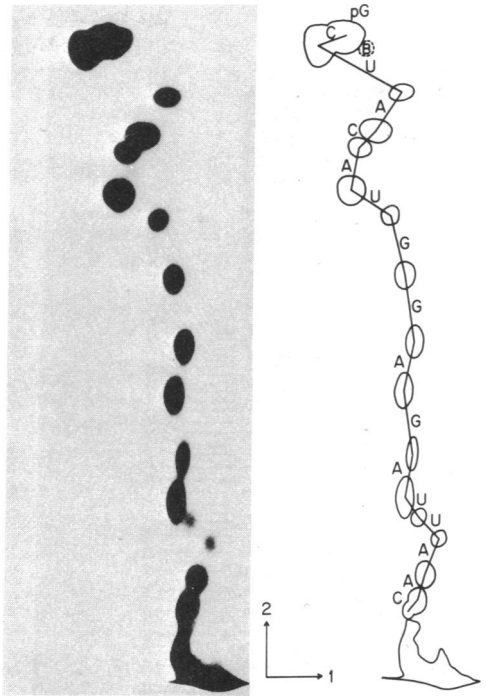



FIG. 2. Two-dimensional separation of a partial nuclease $P_1$ digest of [5′-$^{32}$P]mRNA. This "mobility shift" fingerprint was obtained by two-dimensional electrophoresis–homochromatography. First dimension (12): electrophoresis at pH 3.5 on cellulose acetate in 5% acetic acid/5 mM EDTA/7 M urea at 14°. Second dimension (9): thin-layer homochromatography on a 20 × 40 cm DEAE-cellulose plate at 65° using 20 mM KOH-strength "homomix" (9) (pH 4.7). An autoradiogram of the homochromatogram is shown on the left; a schematic representation is shown on the right. The location of the blue dye marker is indicated by the letter B.

**Determining Positions of Adenines, Guanines, and Pyrimidines by the Polyacrylamide Gel Sequencing Method.** The [5′-$^{32}$P]mRNA was partially digested with ribonucleases $U_2$ and $T_1$ and by limited alkaline hydrolysis essentially as described by H. Donis-Keller *et al.* (11). Details of these procedures are described in the legends of Figs. 3 and 4.

## RESULTS

**Purification of the $^{32}$P-Labeled mRNA on a 5% Polyacrylamide Gel.** Fig. 1 shows an autoradiogram of the gel used for purification of mRNA labeled at the 5′ end with $^{32}$P. The fact that no significant amount of labeled mRNA was observed for the untreated sample (lane 1 in Fig. 1) suggests that this mRNA has a phosphate moiety. About $1.7 \times 10^7$ cpm of the [5′-$^{32}$P]mRNA was recovered from the gel, corresponding to a labeling efficiency of about 31%.

**Identification of the 5′-End Nucleoside Monophosphate.** On the autoradiograms of the thin-layer chromatography plate and the electrophoresis paper used to fractionate portions of the complete nuclease $P_1$ digest, the spot containing 94% of the radioactive material coincided with the position of the pGuo marker (data not shown).

**Nucleotide Sequence of the 5′ End of the mRNA by "Mobility Shift" Analysis.** Partial digestion of the 5′-end-labeled mRNA followed by two-dimensional electrophoresis–homochromatography yielded the "mobility shift" pattern depicted in Fig. 2. From Fig. 2 and the fact that Guo is at the 5′ terminus, the sequence of the first 17 nucleotides from the 5′ end is shown to be GCUACAUGGAGAUUAAC. Residues 2–8 correspond to the sequence of ribonuclease $T_1$ fragment T20a; residues 8–13 correspond to RNase A fragment A31 (unpublished data).

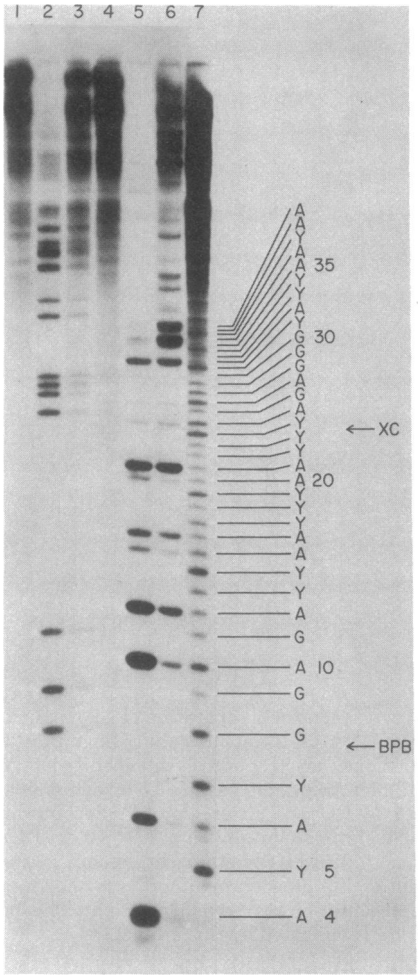



FIG. 3. Autoradiogram of a 20% polyacrylamide sequencing-gel used to identify the positions of all the adenines and guanines among the first 39 nucleotides at the 5′ terminus of the mRNA. Partial digestions with RNase $T_1$ and $U_2$ were carried out in 20 mM Na citrate, pH 5.0/1 mM EDTA/7 M urea/0.025% xylene cyanol/0.025% bromphenol blue at 50° for 15 min (11). Limited alkaline hydrolysis was carried out at pH 9.0 in 50 mM $NaHCO_3/Na_2CO_3$ and 1 mM EDTA (11) at 90° for 15 min. A 20 μl solution of 10 M urea/0.05% xylene cyanol/0.05% bromphenol blue was added to the 20-μl hydrolysis mixture after digestion. A portion (4–7 μl) of each of these digests was applied to the gel as follows (where E = units of enzyme and R = μg of RNA): lane 1, undigested [5′-$^{32}$P]mRNA; lane 2, $T_1$ digest, E/R = $2 \times 10^{-2}$; lane 3, $T_1$ digest, E/R = $2 \times 10^{-3}$; lane 4, $T_1$ digest, E/R = $2 \times 10^{-4}$; lane 5, $U_2$ digest, E/R = $2 \times 10^{-1}$; lane 6, $U_2$ digest, E/R = $2 \times 10^{-2}$; and lane 7, limited alkaline hydrolysis. The buffer used for electrophoresis was 50 mM Tris/borate (pH 8.3) containing 1 mM EDTA. The 20% polyacrylamide gel (33 × 40 × 0.15 cm) was prepared as described by Donis-Keller *et al.* (11). Electrophoresis was carried out at 1000 V (about 10 mA) for 12 hr. The positions of the xylene cyanol and bromphenol blue markers are indicated by XC and BPB, respectively. The letter Y represents the position of pyrimidines in the array.

Residues 12–17 (AUUAAC) contain six nucleotides at the 5′ end of the RNase $T_1$ fragment T33 (AUUAACUCAAUCUAG).

**Extension of the 5′-End Sequence by Using 20% Polyacrylamide Sequencing Gels.** To establish the nucleotide sequence of a longer segment at the 5′ end of the mRNA, the recently developed RNA-sequencing gel technique (11) was employed. Resolution of the three sets of partial products by size, and hence by locus, from the common 5′ terminus is shown in the autoradiograms of the polyacrylamide sequencing gels in Figs. 3 and 4.

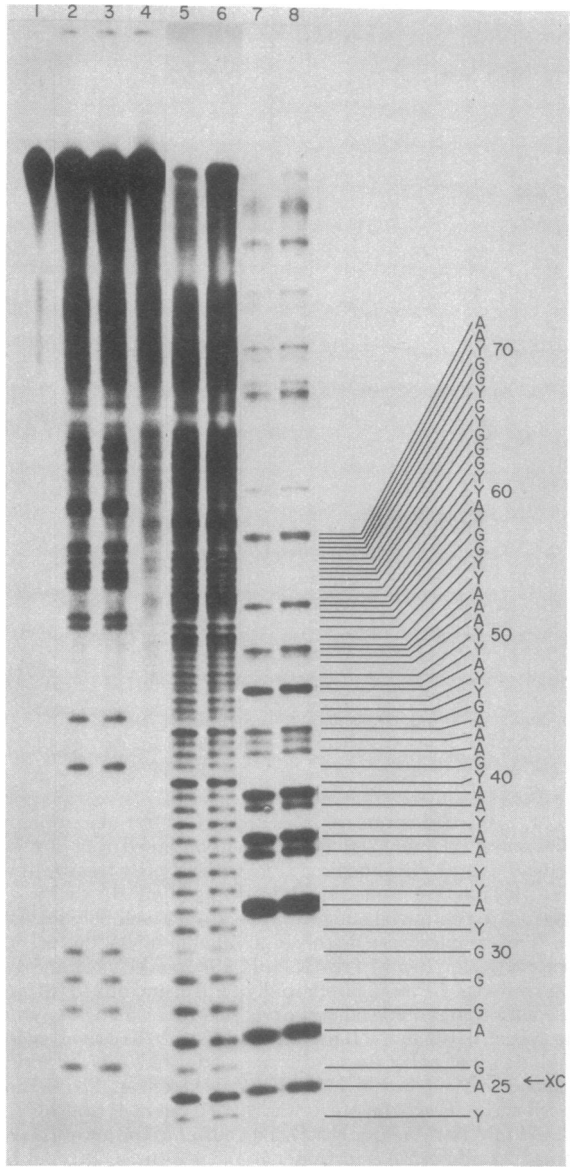



FIG. 4. Autoradiogram of a 20% polyacrylamide sequencing-gel used to identify the positions of all the adenines and guanines between residues 24 and 72 at the 5′ terminus of the mRNA. Partial digestions with RNase $T_1$ and $U_2$ were carried out as indicated in the legend to Fig. 3. Limited alkaline hydrolysis was carried out at pH 9.0 in 50 mM $NaHCO_3/Na_2CO_3$ and 1 mM EDTA at 90° for 15 and 30 min. A portion (4–7 μl) of each of these digests was applied to the gel as follows (where E = units of enzyme and R = μg of RNA): lane 1, undigested [5′-$^{32}$P]mRNA; lane 2, $T_1$ digest, E/R = $1.5 \times 10^{-2}$; lane 3, $T_1$ digest, E/R = $1.0 \times 10^{-2}$; lane 4, $T_1$ digest, E/R = $5.5 \times 10^{-3}$; lane 5, limited alkaline hydrolysis for 15 min; lane 6, limited alkaline hydrolysis for 30 min; lane 7, $U_2$ digest, E/R = $1.5 \times 10^{-1}$; and lane 8, $U_2$ digest, E/R = $1.0 \times 10^{-1}$. Electrophoresis was carried out at 1000 V for 24 hr on the same type of gel as described in Fig. 3. XC indicates the position of the xylene cyanol marker and Y represents pyrimidines.

The array of bands on the sequencing gel indicates that the sequence of residues 4–39 at the 5′ end is as shown in Fig. 3, where Y represents U or C. This was determined by comparison of lanes 2, 3, and 4 (cleavage at internal Gs) and lanes 5 and 6 (cleavage at internal As) with lane 7 (nonspecific cleavage by alkali). The sequence of residues 4–17 corresponds to the sequence previously derived from the "mobility shift" pattern in Fig. 2.

Since a C residue cannot be distinguished from a U residue on the sequencing gel, the information gained from "mobility

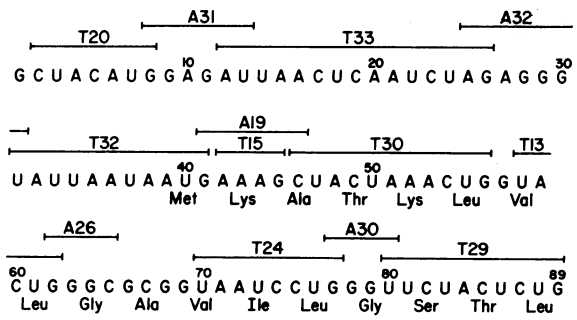



FIG. 5. Nucleotide sequence of the 5′ terminus of the lipoprotein mRNA. The numbers above the nucleotides represent the positions of the nucleotides relative to the 5′ terminal nucleotide. Corresponding oligonucleotides from RNase $T_1$ and RNase A digests (unpublished data) are indicated by the letters T and A, respectively, followed by an identification number. The $NH_2$-terminal sequence of the prolipoprotein (4) is written beneath the corresponding codons of the 5′-terminal region of the lipoprotein mRNA.

shift" analysis (Fig. 2) and nucleotide sequence analysis of all the RNase $T_1$ and A oligonucleotides (unpublished data) was used to determine the identity of the pyrimidines. The A and G positions from residues 2–8, 12–26, and 31–41 correspond to three unique RNase $T_1$-derived oligonucleotides: T20a (CUACAUG), T33 (AUUAACUCAAUCUAG), and T32 (UAUUAAUAAUG). Residues 8–13 and 25–31 correspond to two RNase A-derived oligonucleotides: A31 (GGAGAU) and A32 (AGAGGGU). Thus, the sequence of the first 39 nucleotides is as indicated in Fig. 5.

To determine the sequence beyond residue 39, a second electrophoresis of digests of [5′-$^{32}$P]mRNA similar to that in Fig. 3 was done. In this case, however, the electrophoresis was continued until the xylene cyanol dye marker had reached the bottom of the gel, resulting in products of digestion less than 24 nucleotides in length (from the 5′ end) being eluted from the gel (Fig. 4). By a comparison of lanes 2 and 3 (cleavage at internal Gs) and lanes 7 and 8 (cleavage at internal As) with lanes 5 and 6 (nonspecific cleavage by alkali), it is possible to deduce the sequence shown in Fig. 4. Residues 25–31 and residues 31–41 contain the RNase A-derived fragment A32 (AGAGGGU) and the RNase $T_1$-derived fragment T32 (UAUUAAUAAUG) previously discussed. Other RNase A- and $T_1$-derived fragments occur at the following loci: residues 41–46 (A19, GAAAGC), residues 46–56 (T30, CUACUAAACUG), and residues 58–62 (T13a, UACUG) (unpublished data).

The AUG initiation codon, corresponding to residues 39–41, and the succeeding residues, read in triplets, code for the amino acids of the prolipoprotein as shown in Fig. 5. Residues 56–58 (GGY) correspond to a RNase A-derived fragment, GGU. Residue 58 is known to be U because it is the 5′-end nucleotide of a RNase $T_1$-derived oligonucleotide (T13a, UACUG) that corresponds to residues 58–62 in Fig. 4 and is the only RNase $T_1$-derived pentanucleotide that fits this locus. Residues 62–65 were determined to be GGGC for the following reason: There is only one GGGU (A30) and one GGGC (A26) in the RNase A digest of the mRNA (unpublished data). However, GGGU(A30) should be assigned to residues 77–80, since the U (residue 80) corresponds to the 5′ end of T29, as discussed below. Therefore, residue 65 must be C and residues 62–64 are GGG as seen in Fig. 4. Also, the pyrimidine at residue 67 was inferred to be C from the specificity of the genetic code for alanine (GCX).

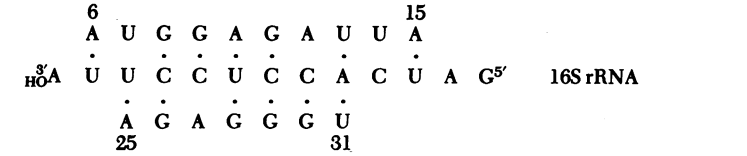
Although the resolution of the bands beyond residue 72 (in lanes 5 and 6 of Fig. 4) diminishes, the amino acid sequence of the prolipoprotein (4) is known and we can identify two unique RNase $T_1$ fragments, T24 (UAAUCCUG) and T29 (UUCUACUCUG) (unpublished data), which have been assigned to this portion of the amino acid sequence (Fig. 5). Residues 70–72 (UAA) shown in Fig. 4 correspond to three nucleotides at the 5′ end of T24. Thus, the sequence of residues 72–77 is as shown in Fig. 5. The three consecutive bands in lanes 2 and 3 (Fig. 4) just beyond the 72nd residue indicate that there is a GGG sequence in the vicinity of residues 77–79. The GGG sequence must correspond to the RNase A-derived fragment A30 (GGGU) at residues 77–80 as mentioned above. Because T29 is the only RNase $T_1$ fragment that codes for the Ser-Thr-Leu locus in the prolipoprotein, the nucleotide sequence can be extended to residue 89 as shown in Fig. 5. This is supported by the fact that the single A band expected for T29 appears in lanes 7 and 8 of Fig. 4. Thus, with the aid of the RNase $T_1$- and A-derived oligonucleotide sequences (unpublished data) and the amino acid sequence (4), the pyrimidines in Figs. 3 and 4 have been identified, and the sequence of 89 nucleotides at the 5′ end of the lipoprotein mRNA has been established.

## DISCUSSION

The mRNA of the lipoprotein from *E. coli* has been isolated without employing *in vitro* transcription or *in vivo* labeling techniques using deletion mutants or transducing phages. Furthermore, because the cistron for the lipoprotein mRNA is thought to be constitutive, the study of its mRNA sequence could provide additional insight into prokaryotic transcriptional and translational processes.

As shown in Fig. 5, there are 38 nucleotides preceding the initiation codon for the prolipoprotein. The most striking feature of the untranslated region at the 5′ end of the lipoprotein mRNA from *E. coli* is that the sequence of the 12 nucleotides prior to, and including, the initiation codon (GUAUUAAUAAUG, residues 30–41) is exactly the same as that found at the 5′ end of brome mosaic virus RNA4 (6), a eukaryotic mRNA. Since the probability of having the same nucleotide sequence of 12 nucleotides is small (1 in $4^{12}$), it is most unlikely that the homology between the lipoprotein mRNA and brome mosaic virus RNA4 is coincidental. Thus, this sequence may have functional or evolutionary implications. In addition, the sequence GCUA (residues 1–4) is homologous to the G$^{m}$CUA found in the corresponding position in all six reovirus mRNAs (13).

In the 38 nucleotides prior to the initiation codon, there are two possible ribosome-binding sites that could form a complementary interaction with the 3′ end of 16S rRNA (14, 15), as shown below, and one possible site with the 3′ end of 18S rRNA (16). The sequence AGAGGGU (residues 25–31) is situated seven nucleotides preceding the AUG initiation codon, in agreement with Shine and Dalgarno (14) and Steitz and Jakes (15).




However, the other potential ribosome-binding site, GGAG (residues 8–11), is located far from the initiation codon. It is interesting to speculate how these two potential ribosome-binding sites may be involved in the control of translation of the mRNA. The sequence UAAUAAU (residues 34–40) is complementary to the AUCAUUA portion of the 3′ end of 18S rRNA (16). This could explain why the lipoprotein mRNA is able to translate in a wheat germ cell-free protein synthesis

system (3). The sequence UAAU (residues 34–37 and also residues 37–40) is known to precede the initiation codon for several eukaryotic mRNAs (6, 13, 17) as well as prokaryotic mRNAs (18–21).

The 5′-end region of the lipoprotein mRNA has nonsense codons out-of-phase with respect to the AUG initiation codon. In a one-base phase shift, two tandem UAA terminators occur at residues 34–39; in a two-base phase shift, a UAA terminator is found at residues 14–16. In addition, the mRNA has an in-phase AUG triplet (residues 6–8) and an in-phase UAG nonsense terminator (residues 24–26) which together could result in the translation of a hexapeptide (Met-Glu-Ile-Asn-Ser-Ile). However, the requirements imposed by the Shine and Dalgarno hypothesis (14) indicate that the AUG (residues 6–8) would not likely serve as a functional initiation site.

A hairpin loop in the 5′-terminal region of the lipoprotein mRNA can be drawn with base pairs between residues 10–14 (AGAUU) and 20–24 (AAUCU) with a $\Delta G$ value of −0.4 kcal (22). Although this loop is nominally stable, the significance of such a structure is open to question. However, the potential ribosome-binding site from residues 25–31 (AGAGGU) lies to the 3′ side of this loop structure in a single-stranded conformation. This site would still remain open for base-pairing with the 3′ end of 16S rRNA. None of the other possible loop structures have even nominal stability. Inverted sequences at the 5′-end untranslated region occur in the lipoprotein mRNA between residues 8–13 (GGAGAU) and residues 24–29 (UAGAGG), residues 14–18 (UAAC) and 19–22 (CAAU), and residues 30–35 (GUAUUA) and 36–41 (AUAAUG) as shown in Fig. 5. These sequences may prevent a great amount of secondary structure at the 5′ end of the mRNA as suggested by Pieczenik *et al.* (21). It can also be pointed out that there are two repeat sequences of seven nucleotides in the 5′ end untranslated region of the lipoprotein mRNA as follows: residues 8–16 (GG*AG*AUUAA) and residues 28–36 (GG*GU*AUUAA).

Because the lipoprotein mRNA has about 360 nucleotides (5), the fact that there are 38 residues in the 5′-end untranslated region and the fact that about 240 residues are needed to code for the prolipoprotein (4) indicate that there are about 80 nucleotides at the 3′-end untranslated region of the mRNA.

In comparison to the 5′-end nucleotide sequences of other *E. coli* mRNAs so far determined (20, 23–26), the lipoprotein mRNA appears to have a rather unique nucleotide sequence, especially with respect to homology with eukaryotic mRNAs. It would be of interest to learn whether the homology between brome mosaic virus RNA4 (6) and *E. coli* lipoprotein mRNA at the 5′ terminus is a special case, or whether further homologies between other prokaryotic and eukaryotic mRNAs will be discovered. The unique features of the lipoprotein mRNA may be associated with the fact that the mRNA is coding for a constitutive membrane protein. They also may be related to the stability of the mRNA (1) or the translational control mechanism of the mRNA.

The authors thank Helen Donis-Keller for a description of her RNA sequencing method prior to publication and Hisako Ohtsubo for demonstration of the technique for constructing and running the 20% polyacrylamide gel. Also, we would like to acknowledge U. L. RajBhandary for information on "mobility shift" analysis and Bernard Dudock for critical reading of the manuscript. This investigation was supported by grants from the U.S. Public Health Service (GM 19043) and the American Cancer Society (BC-67C) as well as by fellowships from the National Cancer Institute, Departments of Health, Education, and Welfare, to R.M.P. (CA 05809) and I.L.P. (CA 09121).

1. Inouye, M. (1975) in *Membrane Biogenesis*, ed. Tzagoloff, A. (Plenum, New York), pp. 351–391.
2. Hirashima, A., Wang, S. & Inouye, M. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 4149–4153.
3. Wang, S., Marcu, K. B. & Inouye, M. (1976) *Biochem. Biophys. Res. Commun.* **68**, 1194–1200.
4. Inouye, S., Wang, S., Sekizawa, J., Halegoua, S. & Inouye, M. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 1004–1008.
5. Takeishi, K., Yasumura, M., Pirtle, R. & Inouye, M. (1976) *J. Biol. Chem.* **251**, 6259–6266.
6. Dasgupta, R., Shih, D., Saris, C. & Kaesberg, P. (1975) *Nature* **256**, 624–628.
7. Pearson, R., Weiss, J. & Kelmers, A. (1971) *Biochim. Biophys. Acta* **228**, 770–774.
8. Lockard, R. & RajBhandary, U. L. (1976) *Cell* **9**, 747–760.
9. Silberklang, M., Prochiantz, A., Haenni, A. & RajBhandary, U. L. (1977) *Eur. J. Biochem.* **72**, 465–478.
10. Silberklang, M., Gillum, A. & RajBhandary, U. L. (1977) *Nucleic Acids Res.* **4**, 4091–4108.
11. Donis-Keller, H., Maxam, A. & Gilbert, W. (1977) *Nucleic Acids Res.* **4**, 2527–2538.
12. Sanger, F., Donelson, J., Carlson, A., Kössel, H. & Fischer, D. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 1209–1213.
13. Kozak, M. (1977) *Nature* **269**, 390–394.
14. Shine, J. & Dalgarno, L. (1975) *Nature* **254**, 34–38.
15. Steitz, J. & Jakes, K. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 4734–4738.
16. Hunt, J. (1970) *Biochem. J.* **120**, 353–363.
17. Rose, J. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3672–3676.
18. Sanger, F., Air, G., Barrell, B., Brown, N., Coulson, A., Fiddes, J., Hutchison, C., Slocombe, P. & Smith, M. (1977) *Nature* **265**, 687–695.
19. Sugimoto, K., Sugisaki, H., Okamoto, T. & Takanami, M. (1977) *J. Mol. Biol.* **110**, 487–507.
20. Musso, R., de Crombrugghe, B., Pastan, I., Sklar, J., Yot, P. & Weissman, S. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 4940–4944.
21. Pieczenik, G., Model, P. & Robertson, H. (1974) *J. Mol. Biol.* **90**, 191–214.
22. Tinoco, I., Borer, P., Dengler, B., Levine, M., Uhlenbeck, O., Crothers, D. & Gralla, J. (1973) *Nature New Biol.* **246**, 40–41.
23. Lee, N. & Carbon, J. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 49–53.
24. Steege, D. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4163–4167.
25. Squires, C., Lee, F., Bertrand, K., Squires, C., Bronson, M. & Yanofsky, C. (1976) *J. Mol. Biol.* **103**, 351–381.
26. Maizels, N. (1974) *Nature* **249**, 647–649.