Proc. Natl. Acad. Sci. USA
Vol. 77, No. 5, pp. 2621–2625, May 1980
Biochemistry

DNA sequence encoding the NH_2-terminal peptide involved in transport of λ receptor, an *Escherichia coli* secretory protein

(*lamB* gene/signal peptide/*in vitro* protein synthesis/amino acid sequence)

JOE HEDGPETH*, JEAN-MARIE CLEMENT, CHRISTIAN MARCHAL, DAVID PERRIN, AND MAURICE HOFNUNG

Unitè de Genie Gènétique, Institut Pasteur, 28 Rue du Docteur Roux, 75015 Paris, France

Communicated by Walther Stoeckenius, February 25, 1980

ABSTRACT ***lamB*** **encodes the λ receptor of** ***Escherichia coli,*** **an outer membrane protein. We have identified the beginning of the** ***lamB*** **gene by correlating DNA nucleotide sequence with a partial sequence of the primary translation product of** ***lamB.*** **We show that λ receptor is synthesized as a precursor containing an extra 25 amino acids at its NH_2 terminus. These amino acids are predominately hydrophobic and probably comprise a structure required for initiation of transport of λ receptor from the cytoplasm to the outer membrane.**

We are interested in identifying the factors that determine that certain proteins are exported from their site of synthesis in the cytoplasm to the exterior of the cell. Exported proteins in both eukaryotic and prokaryotic cells have been shown, with one exception (1), to be synthesized with an NH_2-terminal peptide of 15–30 amino acids which is not found in the mature protein (2, 3). Although these extra NH_2-terminal sequences differ from one protein species to another, they all contain a high proportion of hydrophobic amino acids. It has been suggested that this hydrophobic peptide, termed "signal peptide," initiates the export process by interacting with the cytoplasmic membrane (2). Signal peptides probably also play a role in membrane assembly because some integral proteins of cytoplasmic membranes are synthesized with signal peptides (3, 4).

In *Escherichia coli*, exported proteins are found in the periplasmic space or in the outer membrane. Few, if any, are released from the cell. The protein product of *lamB* (λ receptor) of *E. coli* K-12 is an outer membrane protein that serves in transport of maltodextrins, in transport of maltose at low maltose concentration, and as the receptor for adsorption of phage λ (5–7). λ receptor is synthesized on membrane-bound polyribosomes and the initial product contains an extra peptide not found in mature λ receptor (8). Using an *in vitro* protein-synthesizing system programmed with *lamB*-containing DNA, we have recently shown that λ receptor is, indeed, synthesized as a precursor that contains an extra peptide of about 25 amino acids at its NH_2 terminus (9). Two lines of evidence strongly suggest that this NH_2-terminal sequence is a factor that determines that λ receptor is transported from the cytoplasm. First, fusion of the "early" (NH_2-terminal) region of *lamB* with *lacZ* produced a hybrid protein, containing the NH_2-terminal region of λ receptor and a COOH-terminal region of β-galactosidase. A considerable proportion of this protein is found in the outer membrane, whereas β-galactosidase is normally a cytoplasmic protein (10). Second, mutations in the NH_2-terminal region of *lamB* in this hybrid gene preclude transport of the hybrid protein to the outer membrane (11).

Determining the primary structure of the NH_2-terminal region of *lamB* is an obvious step towards understanding how this part of the protein functions in transport of λ receptor. We present here the nucleotide sequence of the NH_2-terminal region of *lamB*. From the partial amino acid sequence of the precursor protein presented here and the NH_2-terminal sequence of the mature protein (12) we demonstrate that pre-λ receptor contains a very hydrophobic peptide at its NH_2 terminus which is removed apparently during transport of the protein to the outer membrane.

MATERIALS AND METHODS

Bacteria and Phage Strains. Strain pop2044 (13) was used for plasmid propagation. Transformation with plasmid DNA has been described (14). Phage λap*malB13* is a plaque-forming λ*mal* transducing phage that carries the NH_2-terminal third of *lamB* (15). Phage were grown and purified as described (15).

Plasmids. Plasmid pHC6 (Fig. 1) is one of a series of plasmids produced by deleting different segments of DNA from the parent plasmid, pHC1. pHC1 contains an *Eco*RI fragment of the *malB* region, including the COOH-terminal four-fifths of *malK* and the entirety of *lamB* (15), inserted at the *Eco*RI site of pBR322 (ref. 16 and unpublished data). Plasmid pHC6 was constructed by eliminating about 4.3 kilobases of pHC1 between the *Sac* I site, which occurs just beyond the COOH terminus of *lamB* (see ref. 15), and the *Pvu* II site, which occurs some 400 nucleotides from the replication region of pBR322 (17, 18). Elimination of this DNA, which is neither part of *lamB* nor involved in the maintenance of the plasmid, facilitated isolation of small DNA fragments used in the sequencing procedure (see below). To generate this deletion, we treated the parent plasmid DNA with *Sac* I and *Pvu* II, leaving *lamB*, the replication region, and amp^r (the gene for β-lactamase) on a single DNA fragment. After the single-strand regions of the *Sac* I termini were removed with nuclease S1, the DNA was treated with T4 DNA ligase to circularize the fragment by joining the "blunt ends" (19). (See below for details of enzyme reactions.) Strain pop2044 was transformed with the ligated DNA, and Amp^R transformants were selected on tryptone plates containing ampicillin (50 μg/ml). Plasmid DNA from one clone designated pHC6 was not cleaved by *Sac* I or *Pvu* II and was shown by restriction site mapping to have the structure shown in Fig. 1. Strain pop2044, which is λ^r (resistant to λ phage) due to a deleted *lamB* gene, becomes λ^s when transformed with pHC6. Also, pHC6 DNA programmed a low level of synthesis of pre-λ receptor in an *in vitro* protein-synthesizing system (not shown). Hence, pHC6 contains a functional *lamB* gene. The *malB* DNA fragment in the parental plasmid does not contain the promotor normally responsible for *lamB* expression, and we have not yet identified the promotor responsible for *lamB* expression in pHC6.

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

Abbreviation: bp, base pairs.

* Present address: Howard Hughes Medical Institute, HSE 1504, Dept. of Biochemistry and Biophysics, University of California, San Francisco, CA 94143.

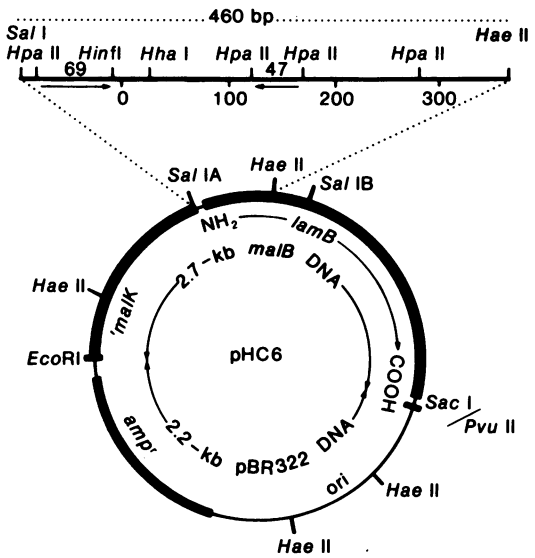

FIG. 1. Structure of pHC6. The notation *'malK* indicates that the *malB* DNA cloned in pHC6 does not contain the region encoding the NH_2-terminal region of the *malK* protein. ori indicates the origin of replication of the cloning vehicle, pBR322 (16). The symbol *Sac* I/*Pvu* II indicates the fusion of the ends generated after treatment of the parental plasmid with these two enzymes. Arrows under fragments *Hpa* II-*Hin*fI 69 and *Hpa* II-*Hpa* 47 indicate the direction in which each of these fragments were used to prime DNA synthesis on λap*malB13 l*-strand and *r*-strand, respectively (see text). The numbers correspond to numbers of base pairs (bp). kb, Kilobases.

To obtain efficient synthesis of the *lamB* protein product *in vitro*, we inserted an *Eco*RI-generated DNA fragment, containing λ promotors P_L and P_R, from plasmid pBH1 (20) at the *Eco*RI site of a plasmid similar to pHC6. In the resultant plasmid (pHC9), *lamB* can be transcribed from λP_R. A detailed description of the construction and properties of the plasmids used here will be published elsewhere.

Enzyme Treatment of DNA. Buffers for restriction endonuclease reactions followed the recipes in the New England Biolabs catalog. The S1 nuclease reaction to generate blunt ends for fusion of the *Sac* I and *Pvu* II sites contained 100 mM sodium acetate (pH 4.5), 5 mM $ZnCl_2$, about 1 μg of plasmid DNA, and 2 μl of nuclease S1 (a gift of F. Rougeon) in a total volume of 6 μl. After 30 min at 20°C, 50 μl of Tris/EDTA (10 mM Tris, pH 8/1 mM EDTA) was added and the mixture was extracted three times with phenol equilibrated with Tris/EDTA. The phenol was removed by extracting the aqueous phase with water-saturated ether, and the DNA was precipitated at −20°C after 3 vol of ethanol was added. The DNA was collected by centrifugation. The pellet was washed once with ethanol and resuspended in 19 μl of ligase buffer (10 mM Tris·HCl, pH 7.8/10 mM $MgCl_2$/1 mM dithiothreitol/0.1 mM ATP) and 1 μl of T4 DNA ligase (Miles) was added. After 12 hr at 10°C 5 μl of the solution was used to transform 100 μl of pop2044 prepared as described in ref. 14.

Phage DNA Strand Separation. A previously described method (21) was used with some modification. Phage λap*malB13* (5–10 OD_{260} units) which had been dialyzed extensively against Tris/EDTA was boiled for 5 min in 1.2 ml of a buffer containing 0.125% Na Sarkosyl (Geigy), 6.5 mM Na EDTA, 5 mM Tris·HCl (pH 9.5), and 200 μg of poly(G-U) (Miles). The solution was placed on ice and 5.5 ml of H_2O saturated with CsCl was added. This mixture was centrifuged for 35–40 hr at 15°C in a Beckman 50Ti rotor at 38,000 rpm. The fractions containing separated strands were located by their A_{260} and dialyzed for 18 hr against Tris/EDTA. The DNA was precipitated with ethanol and resuspended to give an A_{260} of 10–30.

Determination of DNA Sequence. The chain-terminating method developed by Sanger *et al.* (22) was used with minor modifications in the experiment presented in Fig. 2. Small double-stranded DNA fragments generated by restriction endonuclease treatment of plasmid pHC6 were separated by electrophoresis in polyacrylamide gels, located by fluorescence in UV light after staining with ethidium bromide, and electroeluted from the acrylamide in 50 mM Tris acetate, pH 8/2 mM sodium EDTA. Particles of acrylamide were removed by passing the eluate through a Millipore filter or centrifuging through a glass-wool plug in the end of a 1-ml plastic Eppendorf pipette tip. The samples were precipitated with ethanol in the presence of 0.3 M sodium acetate and resuspended in water at approximately 0.2–0.5 pmol/μl. In a typical experiment, 0.5 pmol of a DNA fragment in 5 μl of H_2O was placed in a sealed capillary and boiled for 7 min in a water bath. The denatured fragment was mixed immediately with 10 μl of 1.5× *Hin* buffer (21) containing 0.1 pmol of separated λap*malB13* DNA strand, and the mixture was placed at 68°C in a sealed capillary tube for 30–60 min. Three microliters of this mixture was then placed in each of four capillaries containing 1 μCi each of [α-^{32}P]dATP, [α-^{32}P]dTTP, [α-^{32}P]dGTP, and [α-^{32}P]dCTP (300–500 Ci/mmol, Amersham; 1 Ci = 3.7×10^{10} becquerels) in 1 μl of H_2O. Each capillary then received 1 μl of 0.5 mM dideoxyATP, -TTP, -GTP, or -CTP (PL Biochemicals) followed by 0.5 μl of DNA polymerase I Klenow fragment A (Boehringer Mannheim). The solutions were mixed and incubated for 25 min at 20°C. One microliter of a solution containing all four deoxyribotriphosphates at 1 mM was added to each reaction mixture and incubation was continued for 20 min. The mixtures were then added to 6 μl of a solution containing 20 mM sodium EDTA, 95% (vol/vol) formamide, 0.01% xylene cyanol, and 0.01% bromophenol blue. Electrophoresis in 8% acrylamide gels was done as described (23). After electrophoresis, the gels were exposed to x-ray film (Dupont Cronex or Kodak No-screen) at −20°C.

Amino Acid Sequence Analysis. Plasmid pHC9 DNA was used to program protein synthesis *in vitro* as described (9) except that DNA was used at about 100 μg per reaction (300 μl) and, when [^{3}H]leucine (90 Ci/mol) or [^{3}H]isoleucine (50 Ci/mol) was used (about 50 μCi per reaction), the corresponding unlabeled amino acid was omitted from the reaction. Proteins labeled with [^{35}S]methionine or a tritiated amino acid (or both) were separated by electrophoresis (24), located by autoradiography, electroeluted, and precipitated with 20% (wt/vol) trichloroacetic acid, with 200 μg of bovine serum albumin as carrier. The proteins were washed twice by centrifuging and resuspending the pellets in 10% trichloroacetic acid. The washed pellet was resuspended in aqueous solution of 50% (vol/vol) dioxane containing 0.2 M HCl and incubated for 12 hr at 37°C. The mixture was reprecipitated with 10% trichloroacetic acid, washed three times with ether, resuspended in 300 μl of 30% (vol/vol) formic acid, and added to the cup of a Beckman 890C sequenator with 5 mg of whale myoglobin as carrier. Automated degradation followed the Beckman DMBA program with the exception that two "coupling" reactions were done prior to the first cleavage and a second benzene wash was done after the cleavage step in the first two cycles. Proper function of the sequenator was monitored by identifying the expected whale myoglobin amino acid at each cycle by chromatography of an aliquot after converting the thiazolinone cleavage products to phenylthiohydantoin amino acids. The radioactivity in each fraction was determined after resuspending the dried phenylthiohydantoin derivative in 2 ml of Biofluor (New England Nuclear).

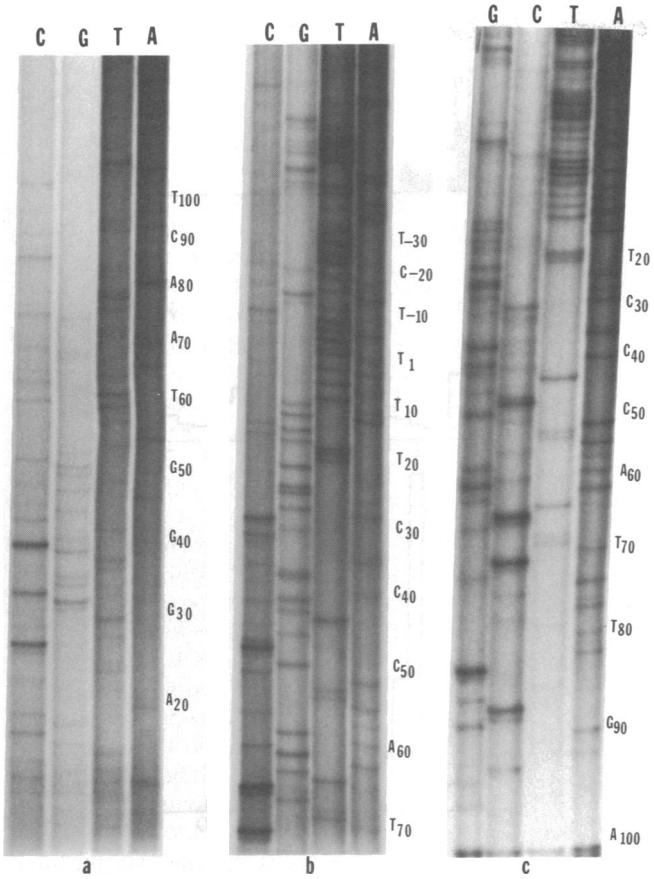

FIG. 2. Gel patterns of DNA fragments generated by extension of DNA primers on λap*malB13* DNA template strands in the presence of chain-terminating analogues. (*a*) Template = *l*-strand, primer = *Hpa* II-*Hin*fI 69; (*b* and *c*) template = *r*-strand, primer = *Hpa* II-*Hpa* II 47. In *a* and *b*, the xylene cyanol was run to the end of the gel (30 cm). In *c*, the xylene cyanol was run about 20 cm. The numbering system corresponds to that in Fig. 3. (Note that the G and C lanes are reversed in *c* compared to *a* and *b*.)

RESULTS

DNA Sequence. Previous experiments suggested that the beginning of *lamB* occurred between the *Sal* I site labeled A in Fig. 1 and the *Hae* II site immediately to the right (25). To carry out DNA sequence analysis in this region, we first determined the location of cleavage sites for several restriction enzymes within the *Sal* I-*Hae* II 460-base-pair (bp) fragment. The *Sal* IA-*Sal* IB 720-bp fragment was labeled at its 5′ extremities with ^{32}P and treated with *Hae* II, and the isolated *Sal* I-*Hae* II fragment (labeled only at the *Sal* I extremity) was used to map the sites for *Hpa* II, *Hha* I, and *Hin*fI by the method of Smith and Birnstiel (26). Next, using this restriction map (Fig. 1) as a guide, small fragments of the *Sal* I-*Hae* II 460 fragment were isolated as described in *Materials and Methods* and used to prime DNA synthesis in the presence of dideoxyribonucleoside triphosphates on the isolated DNA strands of phage λap*malB13*, which contains this region of *malB* (15). From the orientation of the *mal*B region in phage λap*malB13* we could predict that fragment *Hpa* II-*Hin*fI 69 would prime DNA synthesis on the *l*-strand in the same direction as the *malK-lamB* mRNA transcript and that fragment *Hpa* II-*Hpa* II 47 would prime synthesis across the complementary region on the *r*-strand.

From the gel pattern shown in Fig. 2 one can deduce nearly all the DNA sequence in the region between the *Hpa* II sites at −80 and +110 (Fig. 3). The DNA sequence corresponding to the *malK-lamB* mRNA transcript can be read from position

```
                                SD          1   2   3   4   5
                                         f-MET MET ILE THR LEU
AAAAAAGAAAAGCAATGACTCAGGAGATAGA ATG ATG ATT ACT CTG
TTTTTTCTTTTCGTTACTGAGTCCTCTATCT TAC TAC TAA TGA GAC
-30          -20     HinfI -10        -1 +1            10

 6   7   8   9  10  11  12  13  14  15  16  17  18
ARG LYS LEU PRO LEU ALA VAL ALA VAL ALA ALA GLY VAL
CGC AAA CTT CCT CTG GCG GTT GCC GTC GCA GCG GGC GTA
GCG TTT GAA GGA GAC CGC CAA CGG CAG CGT CGC CCG CAT
HhaI  20              30              40              50

 19  20  21  22  23  24  25
MET SER ALA GLN ALA MET ALA val asp phe his gly tyr
ATG TCT GCT CAG GCA ATG GCT GTT GAT TTC CAC GGC TAT
TAC AGA CGA GTC CGT TAC CGA CAA CTA AAG GTG CCG ATA
      60              70              80              90

ala arg ser gly
GCA CGT TCC GGT
CGT GCA AGG CCA
        100  HpaII
```

FIG. 3. DNA and amino acid sequence in the region of the NH_2 terminus of *lamB*. The notation SD (Shine–Dalgarno sequence) indicates a possible ribosome-binding site for initiation of translation of *lamB* mRNA (27). Amino acids found in the extra NH_2-terminal peptide are given in capital letters; those occurring in the mature protein are given in lowercase letters.

+7 extending past position +100 (Fig. 2*a*). The complementary strand sequence can be read from position +99 to −48 (Fig. 2 *a* and *b*). Ambiguities in the sequence of one strand could be resolved by reading the complementary sequence. For example, in Fig. 2*a*, it is not absolutely apparent that two C residues occur at positions 38 and 39 or that three G residues occur at positions 48–50. These uncertainties are resolved by reading the pattern given by the other strand (Fig. 2 *b* and *c*). During many such sequence runs, we found that such ambiguities almost always involved sequences rich in G and C residues and also that *r*-strand templates gave clearer results with a lower background than *l*-strand templates. This latter could be because *l*-strand is contaminated with small amounts of double-stranded DNA since *l*-strand density in CsCl is nearer that of double-stranded DNA than is *r*-strand density.

Identification of NH_2-Terminal Region of *lamB* by Correlation of DNA and Partial Amino Acid Sequences. Starting at position −10, the nucleotide sequence corresponding to the mRNA contains the sequence A-G-G-A-G followed five bases later by A-T-G-A-T-G (Fig. 3). The former is complementary to five bases at the 3′ terminus of 16S rRNA and, in the corresponding mRNA, could comprise the ribosome attachment site for initiation of protein synthesis at either of the two ATG (AUG in the mRNA) triplets (27). Proceeding in the direction of translation, there is no chain-terminating triplet in phase with the ATG triplets for 35 triplets in the sequence shown and actually for at least 150 triplets from other sequence data not shown. To test whether this coding sequence was that of *lamB*, we synthesized the *lamB* product *in vitro* with different labeled amino acids, isolated the labeled protein by polyacrylamide gel electrophoresis, then subjected the purified proteins to automated sequential Edman degradation to see if a given labeled amino acid occurred at the position predicted by the DNA sequence. An example of gel purification of the products synthesized *in vitro* is shown in Fig. 4. The protein labeled pre-λ receptor in Fig. 4 was previously identified as the primary translation product of the *lamB* gene because: (*i*) it was precipitated with anti-λ receptor antiserum, (*ii*) it yielded peptides in common with λ receptor when digested with protease V8, and (*iii*) it contained formyl[^{35}S]methionine at its NH_2 terminus when synthesized *in vitro* in the presence of f[^{35}S]Met-tRNA$_i$. However, as shown in Fig. 4 and elsewhere (9), its molecular

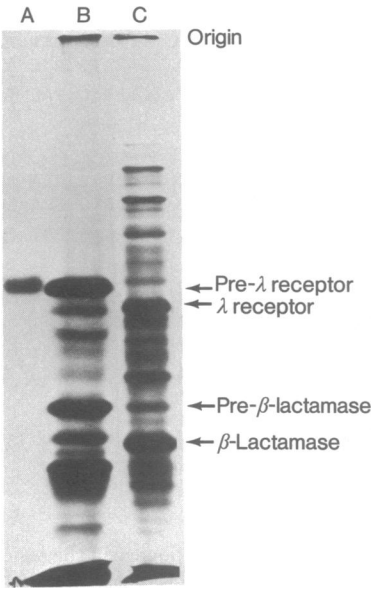

FIG. 4. Electrophoretic isolation of [^{35}S]methionine-labeled pre-λ receptor from other labeled proteins made in a cell-free system programmed by pHC9 DNA (lane B). A sample of previously isolated [^{35}S]methionine-labeled pre-λ receptor was rerun after preparation for sequential Edman degradation (lane A). (Pre-λ receptor synthesized *in vitro* in the presence of other labeled amino acids was isolated in the same way.) For comparison, mature λ receptor synthesized in minicells containing plasmid pHC9 is shown (lane C). Minicell-producing strain JH103, a *malT*$^-$ derivative of DS410 (28), was transformed with pHC9, and minicells were isolated as described (28) except that only one sucrose gradient step was used and the minicells were somewhat contaminated with whole cells. Labeling of the plasmid-containing minicells with [^{35}S]methionine will be described elsewhere. Pre-β-lactamase (lane B) and β-lactamase (lane C) were identified by peptide fingerprinting and partial amino acid sequence analysis and are indicated here only for comparison. Comparison of the DNA sequence of *amp*^{r} and the NH_2 terminus of mature β-lactamase predicts that pre-β-lactamase should be synthesized as a precursor with an extra 23 amino acids at its NH_2 terminus (see ref. 18).

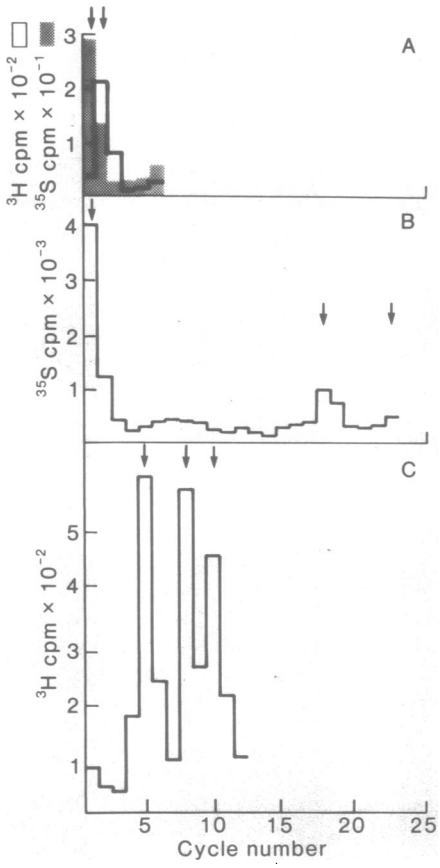

FIG. 5. Sequential Edman degradation of radioactive pre-λ receptor. Pre-λ receptor, synthesized *in vitro* in the presence of [^{3}H]-isoleucine and [^{35}S]methionine (*A*), [^{35}S]methionine (*B*), or [^{3}H]-leucine (*C*), was purified (Fig. 4) and subjected to automated sequential degradation. The radioactivity released at each cycle was determined. The ^{35}S released in *A* is represented by the shaded area.

weight is about 2000 greater than that of λ receptor isolated from cells.

Pre-λ receptor synthesized in the presence of [^{35}S]methionine and [^{3}H]isoleucine yielded ^{35}S at cycle 1 and ^{3}H at cycle 2 with a small amount of ^{35}S at position 2 and a small amount of ^{3}H at position 3 (Fig. 5*A*). In another experiment, pre-λ receptor synthesized in the presence of [^{35}S]methionine yielded ^{35}S at positions 1, 2, 18, 19, and 23, where the sequence run was stopped (Fig. 5*B*). It seemed to us that the pre-λ receptor used in these two experiments consisted of two populations of molecules, a major fraction contained Met-Ile at its NH_2 terminus and a minor fraction contained Met-Met-Ile at the NH_2 terminus. The precautions taken in the sequenator program (see *Materials and Methods*) ensured that the occurrence of ^{35}S at cycle 2 and ^{3}H at cycle 3 in Fig. 5*A* was not due to an artifact of the automated sequence procedure. Both partial sequences are compatible with the DNA sequence shown, the major fraction having the methionine labeled 2 in Fig. 3 at its NH_2 terminus, the minor fraction beginning with the methionine labeled 1. The *N*-formyl group and the NH_2-terminal methionine are removed from some proteins during *in vitro* synthesis; this may be prevented by incubating the S30 extract at 37°C immediately prior to carrying out the synthesis reaction (29). Pre-λ receptor was synthesized in the presence of [^{3}H]leucine by using an S30 extract that had been incubated for 30 min at 37°C just before addition of DNA, amino acids, and cofactors. In this case, ^{3}H occurred at positions 5, 8, and 10, as predicted by the DNA sequence if the protein began with the methionine corresponding to the codon labeled 1 in Fig. 3. We believe, therefore, that initiation of pre-λ receptor synthesis *in vitro* occurs at the methionine codon labeled 1 in Fig. 3 and that the fMet is removed by a protease(s) that is inactivated at 37°C. Support for this notion comes from the fact that pre-β-lactamase (see Fig. 4), synthesized in the same reaction as the pre-λ receptor whose sequence is given in Fig. 5*A*, contained no methionine at its NH_2 terminus and [^{3}H]isoleucine occurred at positions 2 and 10 instead of 3 and 11 as predicted by the DNA sequence (17), indicating that the fMet had been removed (unpublished data). We cannot, however, exclude the possibility that pre-λ receptor may, under certain conditions, initiate at methionine codon 2; also we do not know which of the two methionine codons is used *in vivo*. Nevertheless, the partial amino acid sequence we have determined at the NH_2 terminus of pre-λ receptor is that predicted by the DNA sequence in Fig. 3, indicating that this is indeed the beginning of *lamB*.

Correlation of DNA Sequence with NH_2 Terminus of λ Receptor: Site of Proteolytic Cleavage of Pre-λ Receptor. λ receptor purified from outer membrane of *E. coli* had the NH_2-terminal sequence Val-Asp-Phe-X-Gly-Tyr-Ala (13). The seven triplets beginning at codon 26 (Fig. 3) code for the sequence Val-Asp-Phe-His-Gly-Tyr-Ala. The correlation between DNA and protein sequence confirms that the DNA sequence presented is that of *lamB* and that pre-λ receptor contains a signal peptide of 25 amino acids that is cleaved, apparently, during the export of the protein.

DISCUSSION

The DNA sequence strategy used here to determine the sequence at the beginning of *lamB* should prove useful in other systems where transducing phage carrying the DNA region of interest may be constructed. It would seem especially useful where the sequence analysis of series mutants is to be done because mutants can be generated and characterized genetically while the gene is carried in a lysogen of the transducing phage. For sequence analysis, the phage is produced, the strands are separated, and DNA fragments, isolated from a plasmid (or phage) carrying the wild-type gene, can be used as primers in the chain-terminating sequence method (22). When both strands may be used as templates, the method is reasonably reliable, as shown here by the agreement between DNA sequence and partial amino acid sequence.

We can identify the nucleotide sequence starting with +1 in Fig. 3 as the beginning of the *lamB* coding sequence because (*i*) it occurs within the expected region of the *malB* restriction map, (*ii*) three different radioactive amino acids were incorporated into the primary translation product of *lamB* (pre-λ receptor) at the positions predicted by the DNA sequence, and (*iii*) the DNA sequence between positions +76 and +96 (Fig. 3) codes for a heptapeptide found at the NH_2 terminus of λ receptor (13). Thus, as expected from previous work (8, 9), pre-λ receptor contains an extrapeptide at its NH_2 terminus. The high number of hydrophobic amino acids (19 of 25) in this peptide is its most striking feature and suggests that it is indeed a signal peptide (2, 3). Furthermore, the long stretch of hydrophobic or noncharged amino acids (beginning with leucine 8) following a basic peptide (ending with lysine 7) is a common feature of all known bacterial (4, 18, 30, 31) and many eukaryotic (3) signal sequences. According to the "signal hypothesis" (2), the signal peptide initiates protein transport by inserting into the cytoplasmic membrane, and it is tempting to suggest that the basic residues preceding the hydrophobic stretch may be necessary to interact with the phosphates of the inner phospholipid layer of the membrane (30). However, direct evidence that the function of the signal peptide is to interact with the membrane is lacking. In fact, one current model, stemming primarily from experiments with integral membrane proteins with signal peptides, proposes that signal peptides are necessary to maintain the protein in a soluble state in the cytoplasm prior to insertion of the protein into the membrane (see ref. 4). *E. coli* systems such as *lamB* are amenable to genetic analysis, and several recently isolated mutants of *lamB* have been shown to produce a pre-λ receptor that is not transported (ref. 12; Emr and Silhavy, personal communication). In all these mutants the signal peptide appears to be altered. Correlation of the alteration from the wild-type signal peptide structure of these and other mutants with the change in physiological behavior of their protein products may suggest how the signal peptide functions in pre-λ receptor transport.

Finally, one may ask why signal sequences are cleaved during transport. It is known that a mutant *E. coli* lipoprotein is transported to the outer membrane, its normal location, with the signal peptide uncleaved (32). Also, it has been shown that bovine pancreas preribonuclease may have the same specific nuclease activity as processed ribonuclease (33). Hence, although these may be special instances, they suggest that cleavage of signal sequence may not be required for translocation or for activation. One might speculate that in the case of these "secreted" proteins cleavage is necessary to make translocation a unidirectional process and prevents the protein from returning into the cytoplasmic membrane once it has been translocated.

We thank Genevieve Garreau for excellent technical assistance. We thank M. de Wolff for operating the Beckman sequenator and B. Keil for use of the machine. We are grateful to Pat Clausen for preparing this manuscript. This work was supported by grants from the Centre National de la Recherche Scientifique (LA No. 271), Institut National de la Santé et de la Recherche Médicale (Unitè U163), and the North Atlantic Treaty Organization (Grant 1297). J.H. was a visiting scientist of the Centre National de la Recherche Scientifique and benefited from a special grant from the Mutatest Fund.

1. Palmiter, R. W., Gagnon, J. & Walsh, K. A. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 94–98.
2. Blobel, G. & Dobberstein, B. (1975) *J. Cell Biol.* **67**, 835–851.
3. Habener, J. F., Rosenblatt, M., Kemper, B., Rosenberg, H. M., Rich, A. & Potts, J. I., Jr. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 2616–2619.
4. Wickner, W. (1979) *Annu. Rev. Biochem.* **48**, 23–45.
5. Braun, V. & Krieger-Braun, H. (1977) *Biochem. Biophys. Acta* **469**, 89–98.
6. Szmeleman, S. & Hofnung, M. (1975) *J. Bacteriol.* **124**, 112–118.
7. Randall-Hazelbauer, L. L. & Schwartz, M. (1973) *J. Bacteriol.* **116**, 1436–1446.
8. Randall, L. L., Hardy, S. J. S. & Josefson, L. G. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 1209–1212.
9. Marchal, C., Perrin, D., Hedgpeth, J. & Hofnung, M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 1491–1495.
10. Silhavy, T. S., Shuman, H. A., Beckwith, J. & Schwartz, M. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5411–5415.
11. Emr, S. D., Schwartz, M. & Silhavy, T. S. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 5802–5806.
12. Endermann, R., Hindennach, I. & Henning, U. (1978) *FEBS Lett.* **88**, 71–74.
13. Braun-Breton, C. & Hofnung, M. (1978) *Mol. Gen. Genet.* **159**, 143–149.
14. Cohen, S. M., Chang, A. C. Y. & Hsu, L. (1971) *Proc. Natl. Acad. Sci. USA* **69**, 2110–2114.
15. Marchal, C., Greenblatt, J. & Hofnung, M. (1978) *J. Bacteriol.* **136**, 1109–1119.
16. Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M., Heyneker, H. L., Boyer, H. W., Crosa, J. H. & Falkow, S. (1977) *Gene* **2**, 95–113.
17. Sutcliff, J. G. (1979) *Cold Spring Harbor Symp. Quant. Biol.* **43**, 77–90.
18. Backman, K., Betlach, M., Boyer, H. W. & Yanofsky, S. (1979) *Cold Spring Harbor Symp. Quant. Biol.* **43**, 69–76.
19. Sgaramella, V. & Khorana, H. G. (1972) *J. Mol. Biol.* **72**, 427–444.
20. Hedgpeth, J., Ballivet, M. & Eisen, H. (1978) *Mol. Gen. Genet.* **163**, 197–203.
21. Szybalski, W., Kubinshi, H., Hradenca, Z. & Summers, W. C. (1971) *Methods Enzymol.* **21D**, 383–413.
22. Sanger, F., Nicklen, S. & Coulsen, R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
23. Sanger, F. & Coulson, R. (1978) *FEBS Lett.* **87**, 107–110.
24. Laemmli, U. (1970) *Nature (London)* **227**, 680–685.
25. Raibaud, O., Clement, J. M. & Hofnung, M. (1979) *Mol. Gen. Genet.* **174**, 261–267.
26. Smith, H. O. & Birnstiel, M. (1976) *Nucleic Acids Res.* **3**, 2387–2398.
27. Shine, J. & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 1342–1346.
28. Reeve, J. N. (1977) *Mol. Gen. Genet.* **158**, 73–79.
29. Belin, D. (1979) *Mol. Gen. Genet.* **171**, 35–42.
30. DiRienzo, J. M., Nakamura, K. & Inouye, M. (1978) *Annu. Rev. Biochem.* **47**, 481–532.
31. Sugimoto, K., Sugisaki, H., Okamoto, T. & Takanami, M. (1977) *J. Mol. Biol.* **111**, 487–507.
32. Lin, J. J. C., Kanazawa, H., Ozols, J. & Wu, H. C. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4891–4895.
33. Haugen, T. H. & Heath, E. C. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 2689–2693.