Molecular Components of the Signal Sequence that Function in the Initiation of Protein Export

SCOTT D. EMR and THOMAS J. SILHAVY

Cancer Biology Program, National Cancer Institute–Frederick Cancer Research Facility, Frederick, Maryland 21701. Dr. Emr's present address is the Department of Biochemistry, University of California, Berkeley, California 94720.

ABSTRACT We are studying the mechanism by which the LamB protein is exported to the outer membrane of *Escherichia coli*. Using two selection procedures based on gene fusions, we have identified a number of mutations that cause alterations in the LamB signal sequence. Characterization of the mutant strains revealed that although many such mutations block LamB export to >95%, others have essentially no effect. These results allow an analysis of the functions performed by the various molecular components of the signal sequence. Our results suggest that a critical subset of four amino acids is contained within the central hydrophobic core of the LamB signal sequence. If this core can assume an α-helical conformation, these four amino acids comprise a recognition site that interacts with a component of the cellular export machinery. Since mechanisms of protein localization appear to have been conserved during evolution, the principles established by these results should be applicable to similar studies in eukaryotic cells.

"It may be true that in molecular genetics bacteria have had their day in the sun, but in membrane physiology it is not yet noon."

F. M. Harold (17)

All cells synthesize proteins that are exported to various noncytoplasmic locations. In addition, many cells are capable of true protein secretion. These processes of protein localization are selective and efficient in that proteins are strictly compartmentalized to a particular cellular location. During the past decade, considerable effort has been directed towards elucidating the molecular mechanisms by which cells accomplish these processes. These studies suggest at least two pathways of protein localization. One is cotranslational, i.e., export from the cytoplasm is tightly coupled, if not inseparable, from protein synthesis. The other is posttranslational. In eukaryotic cells, proteins that are routed through the rough endoplasmic reticulum fall into the former class. Conversely, proteins destined for certain subcellular organelles, such as the mitochondria, use a posttranslational pathway (4). Although certain steps in the export process may be shared by both pathways, clear differences exist.

Several important principles regarding cotranslational export have emerged: (*a*) The information that determines localization is contained within the structural gene. This information is not read directly from DNA or mRNA but is read from the amino acid sequence of the gene product. (*b*) Most noncytoplasmic proteins are synthesized initially in larger precursor form (*pre*-protein) with a peptide extension at the NH_2-terminal end of the molecule. The peptide extension (signal sequence) contains the information necessary to start the export process (4, 5, 9, 15). (*c*) Export does not occur spontaneously. A cellular export machinery is required. In eukaryotic cells, certain components of this machinery have been purified (22, 38). In prokaryotic cells, certain components have been defined genetically (10, 25). (*d*) The process of cotranslational export appears to have been conserved during evolution. Intragenic information specifying export in a eukaryotic gene can be recognized by a prokaryotic organism and vice versa (13, 30, 37). This conservation of export mechanisms allows rapid and useful exchange of information between scientists working with diverse organisms.

The Gram-negative bacterium *E. coli* contains four cellular compartments: the cytoplasm, an inner or cytoplasmic membrane, an outer membrane, and an aqueous space between the

This work was presented in a symposium on Mechanisms of Protein Sorting at the Twenty-first Annual Meeting of The American Society for Cell Biology (Anaheim, California, November 1981).

two membranes called the periplasm. We are investigating the mechanism(s) by which proteins are exported to the outer membrane. In particular, we are studying the export of the major outer membrane protein LamB.[1] This protein is a component of the maltose transport system, and, as such, its synthesis is induced by the presence of maltose in the growth media (8). The protein also serves as the receptor for certain bacteriophages (29). Mutants lacking this protein or mutants in which the protein is not in the outer membrane are unable to grow on maltodextrins (Dex^-) and are resistant to the bacteriophage λ (λ^r).

Previous studies have shown that the LamB protein is exported in a manner that is strikingly similar to the manner in which proteins are exported to the plasmalemma of eukaryotic cells. LamB is synthesized initially in larger precursor form by ribosomes that are bound to the cytoplasmic membrane (28). (In terms of protein localization, the cytoplasmic membrane of prokaryotic cells functions in a manner analogous to the rough endoplasmic reticular membrane in eukaryotic cells.) The precursor form of the LamB protein contains a signal sequence of 25 amino acids at the NH_2-terminal end of the molecule (19).

An important advantage of studying the process of protein localization in *E. coli* is that sophisticated genetic techniques can be applied. We have developed methods that enable us to isolate a number of mutations that specifically alter the LamB signal sequence. Most of these mutations block LamB export at an early step and result in accumulation of precursor in the cytoplasm (11, 12). However, certain mutations that alter the signal sequence have essentially no effect on LamB export. In this article, we report the isolation, characterization, and analysis of the DNA sequence of these "leaky" signal sequence mutations. In addition, we discuss the implications of our genetic analysis in terms of the functions performed by various molecular components of the LamB signal sequence.

Isolation of Export-defective lamB Mutations

Previously we have described the construction of strains in which the gene coding for LamB (*lamB*) is fused to the gene coding for the cytoplasmic enzyme β-galactosidase (*lacZ*). The resulting hybrid gene specifies a hybrid protein comprised of an NH_2-terminal fragment of LamB and a large COOH-terminal portion of β-galactosidase. This portion of β-galactosidase is functional and enables cells to grow on lactose (Lac^+). Four classes of *lambB-lacZ* fusions have been identified based on the amount of *lamB* DNA contained in the hybrid gene. By determining the cellular location of the hybrid protein specified by these fusions, we demonstrated that export information is contained in the region of the *lamB* gene corresponding to the NH_2-terminal portion of the LamB protein (for review see references 9 and 15).

Our present work shows that one of the *lamB-lacZ* fusions, 42-1, specifies a hybrid protein with a molecular weight of approximately 137,000. It contains ~170 amino acids coded for by *lamB* DNA. (The wild-type LamB protein contains 446 amino acids including the signal sequence [6].) Cellular fractionation of strains containing this gene fusion reveal that ~40% of the hybrid protein is exported to the outer membrane. The remaining 60% is found evenly distributed between the inner membrane and the cytoplasm. We conclude that most of the *lamB* export information is contained within this hybrid gene.

Strains containing the *lamB-lacZ* fusion 42-1 exhibit two novel and characteristic phenotypes. In our present studies, we have been able to exploit these phenotypes to identify mutations in *lamB* that block protein export.

LETHAL EFFECTS OF POORLY LOCALIZED HYBRID PROTEINS: One of the characteristic phenotypes exhibited by strains carrying the 42-1 fusion is related to the inability of the cell to export the LamB-LacZ hybrid protein efficiently. When cells containing this gene fusion are grown in the presence of maltose to induce high-level synthesis of the hybrid protein, they stop dividing, form long filaments, and ultimately lyse. Apparently, synthesis of large amounts of this hybrid protein causes a lethal jamming of the export machinery. This is supported by the observation that, under conditions in which large amounts of the hybrid protein are synthesized, precursors of many other envelope proteins can be detected accumulating in the cytoplasm of the moribund cell (9). If the maltose-sensitive (Mal^s) phenotype is a consequence of the defective export of the hybrid protein, then selecting a maltose-resistant (Mal^r) phenotype should yield mutants in which defective export of the hybrid protein does not occur. To avoid mutants simply defective in the synthesis of the hybrid protein, we require that these mutants retain β-galactosidase activity and therefore a Lac^+ phenotype. We have devised the following procedure to enrich for the desired Mal^r, Lac^+ mutants (12).

Independent colonies of the *lamB-lacZ* fusion strain, pop3186, were inoculated into separate tubes of maltose minimal M63 medium (23), and the tubes were incubated at 37°C for 48 h or until cultures had grown to saturation. Portions (0.05 ml) of each culture were then inoculated into 5 ml of fresh maltose minimal medium, and the cultures were incubated at 37°C for 24 h. This gave rise to almost pure cultures of Mal^r cells. To select for cells present in this population that retained fusion protein (i.e., were still Lac^+), dilutions of the cultures were plated on lactose minimal agar. To ensure that all of the spontaneously occurring mutants analyzed were the result of independent events, only a single Mal^r, Lac^+ colony from each culture was purified and characterized. Nearly 40 mutants have been isolated using this procedure, all of which fail to export the hybrid protein. In the mutant strains, the hybrid protein is found in soluble form in the cytoplasm (12).

All of the mutations that confer Mal^r are linked genetically to the *lamB-lacZ* fusion. This was shown by isolating λ transducing phages that carry the gene fusion (35). When these phages were lysogenized into a wild-type strain (but lacking β-galactosidase), most of the lysogens remained Mal^r even

[1] The following are guidelines for the genetic nomenclature used in this report. Mutant loci are described by a three-letter designation written in lower case, italicized letters. This abbreviation usually is based on a recognizable phenotype. Individual genes that affect the phenotype are described by an italicized capital letter that immediately follows the three-letter designation. Letters or numbers following the capital letter refer to specific mutations (allele numbers). A hyphen between two genetic abbreviations indicates a gene fusion. Components of the fusion are written in order of transcription. The Greek letter Δ indicates a deletion mutation. The following structural genes appear in the manuscript: *malG*, *malE*, *malK* code for inner membrane components of the maltose transport system; *malE* specifies the periplasmic maltose-binding protein; *lamB* specifies the outer membrane protein which is the receptor for bacteriophage λ; *bla* and *phoA* code for the periplasmic enzymes β-lactamase and alkaline phosphatase, respectively. The following phenotypes are used: Dex, maltodextrin metabolism; Mal, maltose metabolism; Lac, lactose metabolism. A plus or minus sign indicates the ability to use the particular sugar as a carbon source. Superscript letters r or s refer to resistance or sensitivity, respectively.

though they became Lac$^+$. Since the phage confers the mutant phenotype, the mutation must be carried by the transducing phage and therefore must be linked to the fusion.

ALTERED ENZYMATIC PROPERTIES OF HYBRID PROTEINS LOCALIZED TO THE MEMBRANE: Many gene fusions that specify a membrane-bound hybrid protein confer another characteristic phenotype. Strains containing such a fusion contain extremely low levels of β-galactosidase activity in the absence of an inducer (i.e., maltose). When a large fraction of the hybrid protein molecules are embedded in a membrane at low concentrations (the uninduced state), they may not effectively tetramerize into active enzyme. Consequently, in the absence of inducer, these strains grow very poorly on lactose. By selecting for Lac$^+$, mutants can be obtained in which the cellular location of the hybrid protein has been altered (25).

The *lamB-lacZ* fusion 42-1 specifies a hybrid protein that is largely membrane-bound (~40% in the outer membrane, ~30% in the inner membrane). Strains containing this fusion grow poorly on lactose. For reasons not presently understood, this defect is temperature dependent. Strains containing this fusion grow slowly on lactose at 30°C; however, at 37°C they do not grow at all on lactose. By selecting for Lac$^+$ at 37°C using the procedure described below, we have been able to isolate export-defective *lamB* mutants.

Independent colonies of strain pop3186 were inoculated into separate tubes of Luria broth (23) and grown overnight. Aliquots (0.2 ml) from each culture were then plated on lactose minimal M63 agar (23). Plates were incubated at 37°C for 2–3 d. Again, to ensure that all of the spontaneously occurring mutants were the result of independent mutational events, only a single Lac$^+$ colony from each plate was purified. The same selection procedure was also used with cultures mutagenized with nitrosoguanidine (23). Six Lac$^+$ mutants obtained by each method were purified and characterized.

All of the Lac$^+$ mutants contain a genetic lesion linked to the *lamB-lacZ* fusion. In addition, all of the mutants fail to export the hybrid protein. In the mutant strains, the hybrid protein is found in soluble form in the cytoplasm. This is evidenced by the fact that >85% of the β-galactosidase activity is found in the supernatant after centrifugation of cell extracts at 100,000 *g* for 1 h. A periplasmic location for the hybrid protein is ruled out by the fact that β-galactosidase activity remains cell associated after cold osmotic shock (24).

Effect of the Mutations on the Export of the Wild-Type LamB Protein

To determine the effect of the mutations on an otherwise wild-type LamB protein, they were recombined from the *lamB-lacZ* hybrid gene into a wild-type *lamB* gene (Fig. 1). We found that most of the mutations that were selected as Mal^r and one of the mutations that was selected as Lac$^+$ in the parent fusion strain confer a typical LamB$^-$ phenotype to wild-type strains, i.e., the inability to grow on maltodextrins (Dex$^-$) and resistance to phage λ. These phenotypes permit a fine structure genetic analysis by deletion mapping (12) (Fig. 1). Results demonstrate that all of these mutations lie in the region of the *lamB* gene that codes for the signal sequence.

The effect of the mutations that confer a LamB$^-$ phenotype on the localization of the LamB protein was determined by fractionating the mutant cells into the four cellular compartments (cytoplasm, periplasm, inner membrane, and outer membrane). Immune precipitation of each of these fractions with

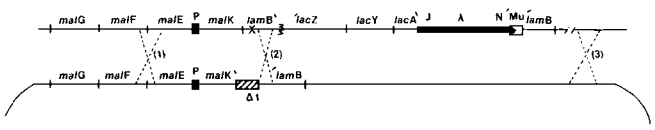

FIGURE 1 Recombination of the export-defective mutations from the hybrid *lamB-lacZ* gene into an otherwise wild-type *lamB* gene. Shown are the two divergent operons that comprise the *malB* locus in *E. coli*. These two operons specify five proteins that, together, make up the active transport system for maltose and maltodextrins. Both operons, and thus the synthesis of all five proteins, are induced by maltose in the growth media. The *malE* gene specifies the periplasmic maltose-binding protein; *malG*, *malF*, and *malK* specify proteins associated with the inner membrane. Transcription of these two operons is initiated from a central region designated *P*. Strains containing the deletion Δ(*malB*)1 are unable to grow on maltose (Mal$^-$) because the deletion removes part of the *malK* gene. This deletion also removes the portion of the *lamB* gene that codes for the signal sequence. If this strain is transduced to a Mal$^+$ phenotype, the transductants must inherit from the donor that portion of the *lamB* gene coding for the signal sequence, since the recombinational event that incorporated *malK* DNA must also incorporate *lamB* DNA (12).

The upper line represents the transducing phage fragment of DNA, derived from a Mal^r, Lac$^+$ *lamB-lacZ* fusion strain, brought in during the transduction. The lower line represents the homologous region of the chromosomal DNA in a strain containing the Δ(*malB*)1 deletion with which the transducing phage fragment can recombine. Shown are the possible sites of recombination that will give rise to Mal$^+$ transductants. Recombination at positions *1* and *3* will lead to incorporation of the *lamB-lacZ* fusion and adjacent bacteriophage λ DNA (present because of the method of construction of the gene fusion) into the recipient chromosome. Recombination at positions *1* and *2* will lead to the replacement of DNA covered by the Δ(*malB*)1 deletion. These two classes of recombinants can be distinguished because recombinants of the first type will be Lac$^+$ (they contain the *lamB-lacZ* fusion) and recombinants of the second type will be Lac$^-$.

This experiment allowed us to determine the location of the mutation by deletion mapping. If all Mal$^+$, Lac$^-$ transductants contain the export-defective mutation then the mutation must lie in that region defined by Δ(*malB*)1, i.e., the region of *lamB* coding for the signal sequence.

anti-LamB serum showed that the mutant protein precursor is in the cytoplasmic fraction of these cells. Presumably, in this location the protein is sequestered from the signal peptidase, which is necessary for cleavage of the signal sequence from the precursor.

Ten of the mutants isolated using the Mal^r, Lac$^+$ selection and nearly all of the mutants isolated using the Lac$^+$ selection present something of an enigma. All contain a genetic lesion in the *lamB* portion of the hybrid gene. All of them prevent export of the hybrid protein. As stated above, the hybrid protein in these strains is found in the cytoplasm. However, when these mutations are recombined into an otherwise wild-type *lamB* gene, the resulting strains all exhibit a normal LamB$^+$ phenotype (Dex$^+$, λ^s). Cell fractionation and immune precipitation revealed an apparently normal LamB protein in the outer membrane in wild-type amounts (Fig. 2). By performing immune precipitation on radioactively labeled whole-cell extracts, we were able to demonstrate the presence of a small amount (<2%) of LamB protein precursor (Fig. 3). Since we have never detected precursor in wild-type cells, we conclude that these mutations do effect LamB export; however, the effect is quite small.

Since these mutations are phenotypically silent when present

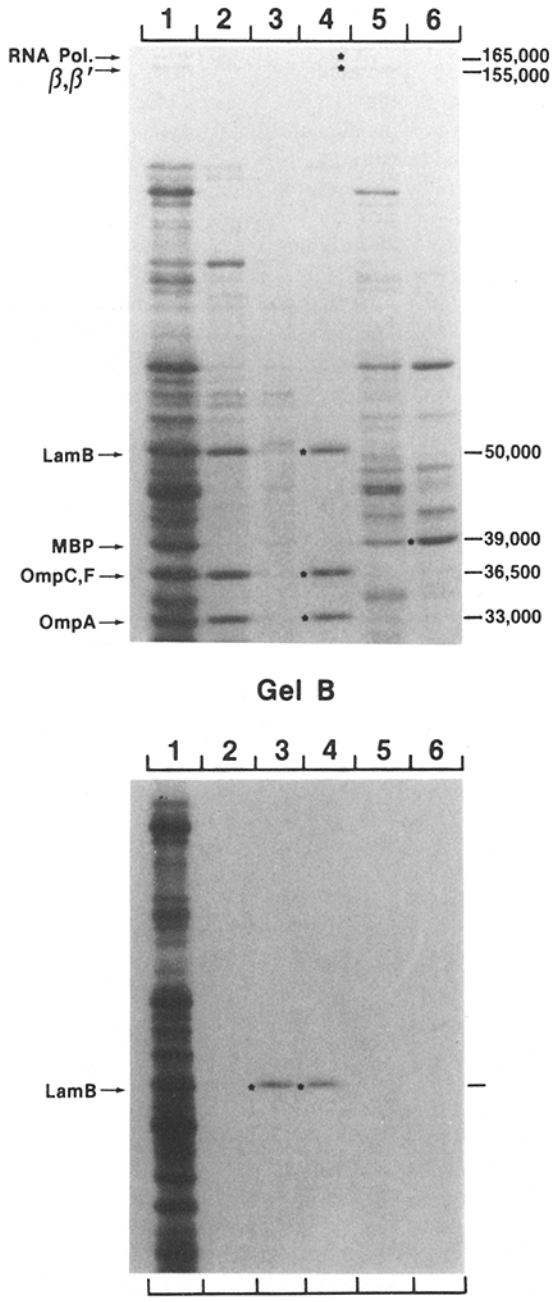

FIGURE 2 Location of LamB protein in the mutant cell. Strains containing the phenotypically silent *lamB* mutations were grown at 28°C in maltose minimal medium (250 ml) to mid log phase. The cultures were then labeled with ^{14}C-uniformly labeled amino acids (1 μCi/ml) for 7 min (12). Labeling was stopped by diluting cultures 1:5 with ice cold Luria broth. Cells were then pelleted and fractionated.

Inner and outer membranes were separated by the selective solubilization technique described by Schnaitman (32) or by isopycnic sucrose density-gradient centrifugation as described by Osborn et al. (26). The bacterial periplasmic fraction was obtained by cold osmotic shock (24). In gel *A*, samples from each of the cellular fractions from a representative mutant (SE2073) were subjected to electrophoresis in a 9% SDS polyacrylamide gel (21). The gel was then stained with Coomassie Brilliant Blue. Samples from each of the cellular fractions were also subjected to immune precipitation with rabbit anti-LamB serum and to gel electrophoresis as described (12, 34). Gel *B* is the autoradiogram of such a gel. Gel *A*: lane *1*, whole cell extract; lane *2*, total membrane fraction (inner and outer); lane *3*, inner membrane; lane *4*, outer membrane; lane *5*, total soluble fraction (cytoplasm and periplasm); lane *6*, periplasm. Marker proteins known to be localized to specific cellular compartments are indicated. These include: the outer membrane proteins OmpC, OmpF, and OmpA; the cytoplasmic RNA polymerase subunits (β and β'); and the periplasmic maltose-binding protein. Gel *B*: lane *1*, whole cell extract; lane *2*, anti-LamB precipitation from the inner membrane fraction; lane *3*, anti-LamB precipitation from the outer membrane fraction; lane *4*, marker wild-type λ receptor protein; lane *5*, anti-La precipitation from the soluble fraction; lane *6*, anti-LamB precipitation from the periplasmic fraction.

in a wild-type gene, genetic analysis is difficult. Their presence in an otherwise wild-type gene can only be detected genetically using techniques of marker rescue. This is done by lysogenizing strains that are thought to carry the mutation with λ transducing phages that carry the parent *lamB-lacZ* fusion 42-1. These lysogens are Lac$^-$ at 37°C. However, when these lysogens are plated on minimal lactose agar, Lac$^+$ recombinants appear at high frequency, i.e., the mutant phenotype of the fusion can be rescued through genetic recombination by the mutation present in the otherwise wild-type *lamB* gene in the lysogen.

Deletion mapping experiments were done with fusion strains containing these phenotypically silent mutations as described in Fig. 1. Twenty of the resulting transductants were scored for the presence of the mutation by marker rescue. All were found to carry the silent mutation. From this we conclude that these mutations must lie within or very close to the region of the *lamB* gene that codes for the signal sequence.

DNA Sequence Analysis

Previously we have reported the DNA sequence analysis of the mutations that block LamB export and confer a LamB$^-$ phenotype. All of these mutations cause alterations in the LamB signal sequence (11). These results are summarized in Fig. 4.

Three of the mutations that are phenotypically silent when present in an otherwise wild-type gene were chosen for DNA sequence analysis. Two of those, *lamBS96* and *lamBS73*, occurred spontaneously. The other, *lamBS110* was isolated after nitrosoguanidine mutagenesis. As predicted by genetic analysis, all of these mutations alter the region of the *lamB* gene that codes for the signal sequence. The spontaneous mutation *lamBS96* is a transversion that changes the glycine codon at position 17 to an arginine codon. The remaining mutations both lead to a base substitution that changes the same glycine codon to an aspartic acid codon. Consistent with the types of mutations known to be caused by nitrosoguanidine, this event corresponds to an A:T to G:C transition mutation (23).

DISCUSSION

The *lamB-lacZ* fusion strain, pop3186, exhibits two characteristic phenotypes that we have been able to exploit to isolate export-defective *lamB* mutants. One of these phenotypes, maltose sensitivity (Mal^s), relates to the inability of the cell to export large amounts of the LamB-LacZ hybrid protein efficiently. The other phenotype relates to the low β-galactosidase activity of the membrane-associated hybrid protein. By selecting for relief of the Mal^s phenotype (Mal^r) or for increased β-galactosidase activity (Lac$^+$), we have been able to isolate mutant strains in which export of the hybrid protein from the cytoplasm is blocked. In all of the mutant strains, the hybrid protein is found in soluble form in the cytoplasm.

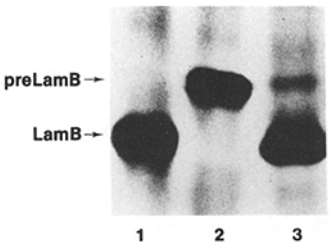

FIGURE 3 Immune precipitation of LamB and preLamB protein from wild-type and mutant strains. Cultures (1 ml) were grown at 28°C in maltose minimal medium to mid log phase (OD_{600} = 0.5) at which time 10 μCi of [^{35}S]methionine was added per milliliter of culture. Cells were labeled for 4 min. Labeling was stopped by placing cultures in an ice bath. Immune precipitation with anti-LamB was then carried out on whole cell extracts, and the precipitation was run on 9% SDS polyacrylamide gels as previously described (10). Lane *1*, precipitation from the wild-type parent strain; lane *2*, precipitation from a representative export-defective mutant; lane *3*, precipitation from a representative leaky mutant. Only the relevant portion of the autoradiogram is shown. The positions of LamB and preLamB protein are indicated. The gel was purposely overexposed to clearly show the presence of the preLamB protein present in the leaky mutant strain.

Most of the mutations isolated by selecting Mal^r and one of the mutations isolated by selecting Lac$^+$ confer a typical LamB$^-$ phenotype when recombined into an otherwise wild-type *lamB* gene. In these recombinants, the mutant LamB is found in soluble form in the cytoplasm with the signal sequence still attached. Genetic and DNA sequence analysis revealed that all of these mutations alter the LamB signal sequence. These results demonstrate that a functional signal sequence is required for export. In addition, they indicate that the signal sequence functions at a very early stage in the export process. If the step mediated by the signal sequence is blocked, protein export does not initiate.

The remaining mutations that were isolated using these selections also block export of the LamB-LacZ hybrid protein. However, they do not confer a LamB$^-$ phenotype when recombined into an otherwise wild-type *lamB* gene. They have essentially no affect on either the export or processing of the LamB protein. In this case, the mutations are very "leaky." Evidence obtained in eukaryotic systems indicates that the

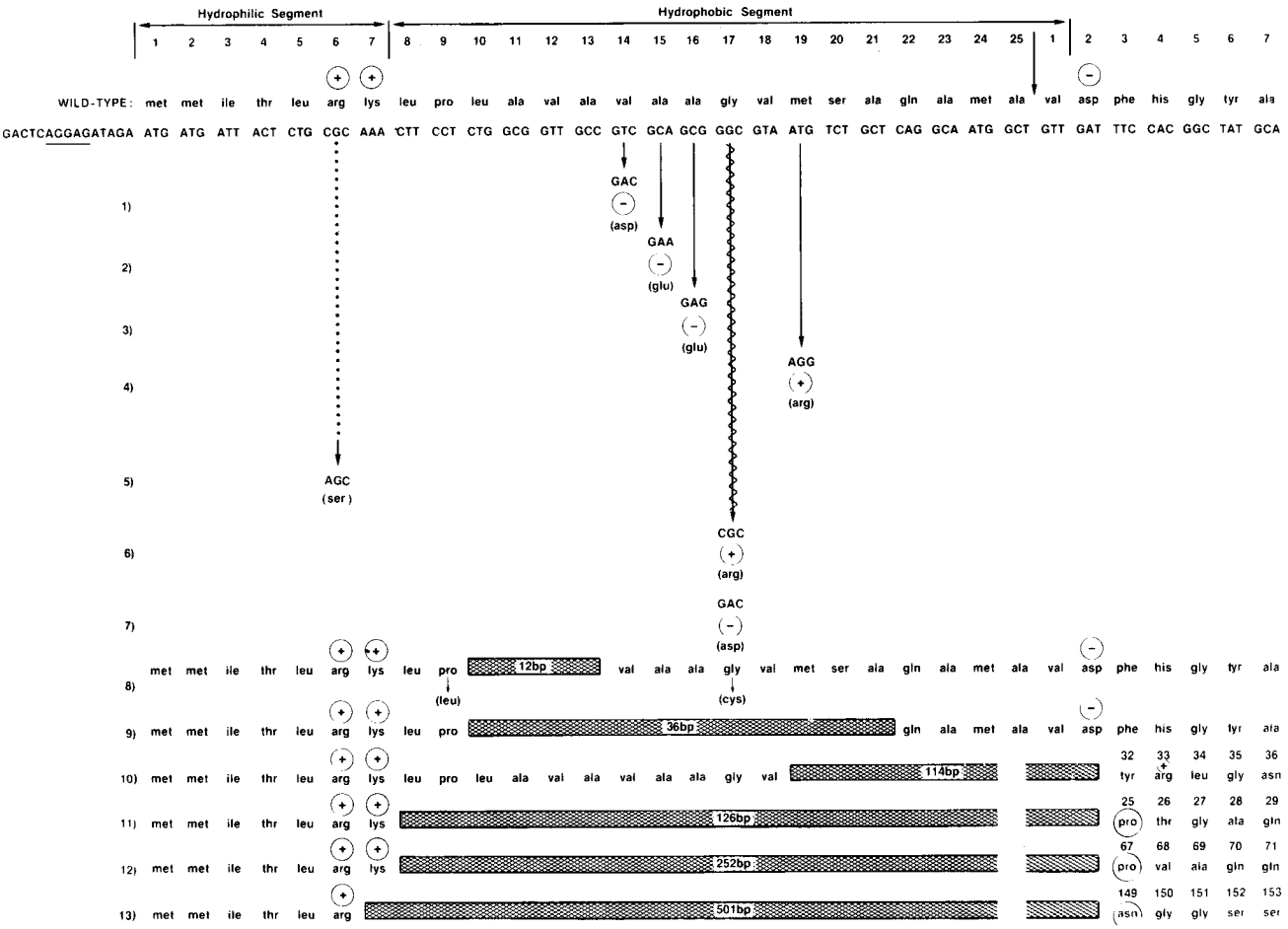

FIGURE 4 Mutations presently known that lead to alterations in the LamB signal sequence. The mutations are divided into three classes and indicated by straight lines, a dotted line, or a squiggly line, all with arrowheads. The amino acid alterations caused by point mutations *1-4* and deletion mutations *8-13* prevent export of the LamB protein (9). The amino acid substitution in point mutant *5* seems to interfere with a cellular mechanism that couples the export and translation of LamB (16, 17). The point mutations *6* and *7* are the mutations described in this report. They do not lead to any significant block in LamB export, i.e., they exhibit a very "leaky" phenotype. Each of the amino acid residues in parentheses below line *8* represent substitutions that restore function to this mutant signal sequence (see text). Numbers above the amino acid residues indicate position in either the precursor or mature LamB protein sequence. The amino acid directly following deletions *11-13* (shown in circles) are not normally present in the wild-type LamB protein sequence at the positions indicated (6). They are coded for by fused codons comprised of nucleotides located directly before and after each deletion. The site of LamB signal sequence processing is indicated above the wild-type sequence by a vertical arrow. The extent of each deletion is indicated as number of base pairs deleted in the shaded bars. The charge exhibited by certain of the amino acids in either the wild-type or mutant signal sequences is indicated in circles above each of the charged residues.

signal sequence must initiate protein export before synthesis of most of the protein has occurred (31). If we assume that this is true of our system, then we would predict that export initiates before the *lacZ* portion of the hybrid gene is translated. The fact that certain signal sequence mutations block export of the LamB-LacZ hybrid protein but not of LamB is not consistent with this prediction. At present, we do not understand this anomaly. It may be that the export is not completely cotranslational and that information downstream affects export initiation. Alternatively, the presence of the mutation together with β-galactosidase sequences may cause the export process to initiate but then abort at some later stage.

Most of the mutations that confer Mal^r also confer $LamB^-$ (Dex^-, λ^r) when recombined from the hybrid gene to the wild-type gene, whereas most of the Lac^+ mutations did not. This phenomenon did not seem to be caused by a difference in stringency between the two selections because the Lac^+ mutants are all Mal^r and vice versa.

The "leaky" mutations described here bring the total number of known LamB signal sequence alterations to 13. The effects of all of these mutations on LamB export have been determined. Taken together, these results permit an analysis of the functions performed by the various molecular components of the signal sequence in the initiation of protein export.

Like all signal sequences, the LamB sequence can be divided into two distinct domains: an NH_2-terminal hydrophilic segment and a central hydrophobic core that extends near to the site of processing. These two domains are generally separated by one or two basic amino acids, especially in prokaryotic sequences (Fig. 4). Sequence comparison of all known signal sequences reveals no other striking homologies except for the presence of an amino acid with a small side chain (ala, ser, gly, cys) at the processing site.

The NH_2-terminal domain, excluding the basic amino acid residues, does not appear to play a critical role in export. Several lines of evidence support this claim. First, no export-defective mutations are known which lie in this region. Second, this sequence varies in composition and varies in size from as small as one (methionine, specified by the initiation codon) to as large as seven. Indeed, Talmadge et al. (36) have placed 18 extra amino acids in this region of the signal sequence of preproinsulin. Export from the cytoplasm of *E. coli* appears to occur normally.

The central hydrophobic core clearly plays an important role in the initiation of protein export. All of the export-defective *lamB* mutations alter this region of the signal sequence. Similarly, all export-defective *malE* (1) (codes for the periplasmic maltose-binding protein), *bla* (codes for β-lactamase; D. Koshland and D. Botstein, personal communication) and *phoA* (codes for the periplasmic enzyme alkaline phosphatase; S. Michaelis and J. Beckwith, personal communication) mutations alter this region as well. Since no sequence homology can be recognized between various signal sequences and since all export-defective *lamB* mutations, and nearly all others as well, result in the presence of a charged amino acid in this region, it could be argued that this sequence functions simply because it is hydrophobic. Although we do not doubt the obvious importance of hydrophobicity, we believe that certain amino acid residues in this region play a more critical role in export initiation. In the LamB signal sequence, the presence of a charged residue at positions 14, 15, 16, or 19 blocks export to >95%. However, as we have demonstrated here, a charged residue, either acidic or basic, at position 17 has essentially no effect on the export of LamB to the outer membrane. Thus, it is the position of the alteration, not the presence of a charge, which determines the effect of the mutation.

We believe that the residues at positions 14, 15, 16, and 19 define an important recognition site. These four residues probably interact directly with a cellular component of the protein export machinery. If one of these residues is altered by mutation, this critical recognition cannot occur and the export process does not initiate. The result is the accumulation of precursor in the cytoplasm. The data we have obtained by genetic analysis of the LamB signal sequence are consistent with this proposal. All of the export-defective *lamB* mutations (14 base substitutions and 13 deletion mutations) alter one or remove one or more of these critical four residues. The two point mutations described here that do not alter one of these residues do not block export.

An apparent exception to the statement that all export-defective *lamB* mutations alter at least one of these critical four residues is the small deletion mutation *lamBS78*. This twelve-base-pair deletion removes amino acids 10, 11, 12, and 13 from the LamB signal sequence. It blocks export to >95%. Although this deletion certainly does not alter one of the four critical amino acids directly, evidence that we have obtained recently (13) indicates that the mutation alters the recognition site indirectly by altering the secondary conformation of residues 14, 15, and 16.

Using rules to predict peptide secondary structure (7), it has been determined that the hydrophobic core of the LamB signal sequence most probably exists in an α-helical conformation (2, 3). Since two amino acids in this core region, proline at position 9 and glycine at position 17, destabilize helical structures, it is predicted that the helix terminates in the region of these two residues. According to these rules, none of the point mutations that alter the LamB signal sequence would alter this secondary structure. However, the small deletion mutation *lamBS78*, which removes residues 10, 11, 12, and 13, would alter the secondary structure because in the mutant signal sequence the helix-destabilizing residues, proline and glycine, are too close to each other (three residues apart instead of seven as in the wild-type sequence) to permit a helix to form between them. Consequently, the critical residues 14,15, and 16 cannot form the helical conformation required for recognition.

The contention that the *lamBS78* deletion alters the secondary structure as described above has been tested genetically. Since the critical recognition site is still intact in the mutant signal sequence, we predicted that function would be restored by a second mutation that permits the critical region to assume an α-helical conformation. That is what we observed. Secondary mutations that change the proline at position 9 to leucine or that change the glycine at position 17 to cysteine restore function to the mutant signal sequence. Both of these changes permit the recognition site to assume an α-helical conformation (Fig. 5).

The various molecular components of the LamB signal sequence and the function each appears to perform in initiating protein export can be summarized as follows: (*a*) The NH_2-terminal domain does not appear to be required. (*b*) The central hydrophobic core is essential. Furthermore, this core must be able to assume an α-helical conformation to allow recognition to occur. (*c*) A critical subset of four amino acids contained within the hydrophobic core comprises a recognition site that interacts directly with a component of the cellular export machinery.

The nature of the cellular component that interacts with the recognition site in the hydrophobic core is not known. We

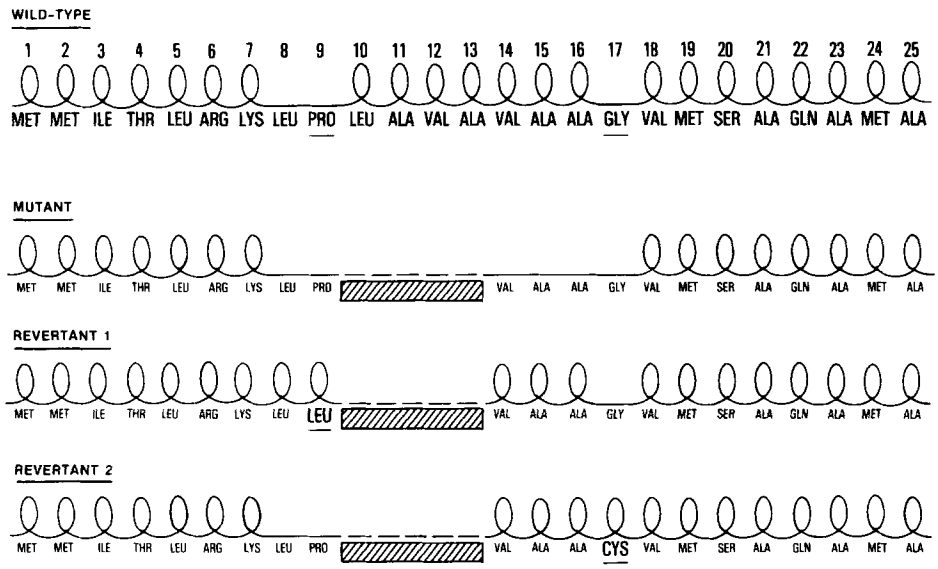

FIGURE 5 Secondary structure analysis of the LamB signal sequence from wild-type, deletion mutant (SE2078), and double mutant (revertants 1 and 2) cells. The *lamBS78* deletion present in strain SE2078 and in the revertants is indicated by a shaded bar. Regions of predicted α-helical (loops) or random coil (straight line) secondary structure are indicated above each sequence. Amino acid substitutions in the revertants are underlined.

presume, however, that the components will be defined genetically by mutations like *prlA*. Such mutations alter a cellular component and restore recognition of mutationally altered signal sequences (10). We do not mean to imply that the only function performed by the signal sequence is in the initiation of export. Protein localization is likely to be a multistep process. Conceivably, the signal sequence could function in several of these steps.

The function of the basic amino acid residues that separate the two signal sequence domains remains unclear. None of the export-defective mutations that we have isolated alters one of these residues. This would suggest that these residues do not function in export initiation. Recently, Schwartz et al. (33) isolated a mutant in which the arginine at position 6 of the LamB signal sequence is changed to a serine. Results that were obtained with this mutant suggest that the mutation may interfere with the cellular mechanism that couples export and translation (16, 17). Analogous mutations have been constructed in vitro in the gene coding for lipoprotein, a major outer membrane protein of *E. coli*, and similar results were obtained (20). Recently Walter and Blobel (39) have isolated a protein factor from eukaryotic cells that appears to mediate the coupling of export and translation. Although more work is required, it seems likely that the similarlities between prokaryotic and eukaryotic export processes extend to the details of export initiation.

This research was sponsored by the National Cancer Institute, Department of Health and Human Services, under Contract N01-CO-75380 with Litton Bionetics, Inc., Kensington, MD. The contents of this publication do not necessarily reflect the views of policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U. S. Government.

Received for publication 11 June 1982, and in revised form 18 September 1982.

REFERENCES

1. Bedouelle, H., P. J. Bassford, A. V. Fowler, I. Zabin, J. Beckwith, and M. Hofnung. 1980. Mutations which alter the function of the signal sequence of the maltose-binding protein of *Escherichia coli*. *Nature* (*Lond.*) 285:78–81.
2. Bedouelle, H., and M. Hofnung. 1981. Functional implications of secondary structure analysis of wild-type and mutant bacterial signal peptides. *Prog. Clin. Biol. Res.* 63:309–403.
3. Bedouelle, H., and M. Hofnung. 1981. On the role of the signal peptide in the initiation of protein exportation. *In* Intermolecular Forces. B. Pullman, editor. D. Rudel Publishing Co., Hingham, ME. 361–372.
4. Blobel, G. 1980. Intracellular protein topogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 77:1496–1500.
5. Blobel, G., and B. Dobberstein. 1975. Transfer of proteins across membranes. I. Presence of proteolytically processed and nonprocessed nascent immunoglobulin chains on membrane-bound ribosomes of murine melanoma. *J. Cell Biol.* 67:835–851.
6. Clement, J. M., and M. Hofnung. 1981. Gene sequence of the λ receptor, an outer membrane protein of *E. coli* K12. *Cell.* 27:507–514.
7. Chou, P. Y., and G. D. Fasman. 1978. Empirical predictions of protein conformation. *Annu. Rev. Biochem.* 47:251–276.
8. Debarbouille, M., H. A. Shuman, T. J. Silhavy, and M. Schwartz. 1978. Dominant constitutive mutations in *malT*, positive regulator gene of maltose regulon in *Escherichia coli*. *J. Mol. Biol.* 124:359–371.
9. Emr, S. D., M. N. Hall, and T. J. Silhavy. 1980. A mechanism of protein localization–the signal hypothesis and bacteria. *J. Cell Biol.* 86:701–711.
10. Emr, S. D., S. Hanley-Way, and T. J. Silhavy. 1981. Suppressor mutations that restore export of a protein with a defective signal sequence. *Cell.* 23:79–88.
11. Emr, S. D., J. Hedgpeth, J.-M. Clement, T. J. Silhavy, and M. Hofnung. 1980. Sequence analysis of mutations that prevent export of phage lambda receptor, an *Escherichia coli* outer membrane protein. *Nature* (*Lond.*). 285:82–85.
12. Emr, S. D., and T. J. Silhavy. 1980. Mutations affecting localizations of an *Escherichia coli* outer membrane protein, the bacteria phage lambda receptor *J. Mol. Biol.* 141:63–90.
13. Emr, S. D., and T. J. Silhavy. Genetic evidence for the role of signal sequence secondary structure in protein secretion. *Proc. Natl. Acad. Sci. U. S. A.* in press.
14. Fraser, T. H., and B. J. Bruce. 1978. Chicken ovalbumin is synthesized and secreted by *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 75:5936–5940.
15. Hall, M. N., S. D. Emr, and T. J. Silhavy. 1981. Genetic studies on mechanisms of protein localization in *Escherichia coli* K-12. *J. Supramol. Struct.* 13:147–164.
16. Hall, M. N., J. Gabay, M. Debarbouille, and M. Schwartz. 1982. A role for mRNA secondary structure in the control of translation initiation. *Nature* (*Lond.*). In press.
17. Hall, M. N., and M. Schwartz. 1982. Reconsidering the early steps of protein secretion. *Annales de l'Institut Pasteur.* In press.
18. Harold, F. M. 1972. Conservation and transformation of energy by bacterial membranes. *Bacteriol. Rev.* 36:172–230.
19. Hedgpeth, J., J. M. Clement, C. Marchal, D. Perrin, and M. Hofnung. 1980. DNA-sequence encoding the NH_2-terminal peptide involved in transport of lambda receptor, an *Escherichia coli* secretory protein. *Proc. Natl. Acad. Sci. U. S. A.* 77:2621–2625.
20. Inouye, S., T. Franceschini, K. Nakamura, X. Soberon, K. Itakura, and M. Inouye. 1982. Requirement of positive charge at the amino terminal region of the signal peptide: an evidence for the loop model for protein secretion across the membrane. *Proc. Natl. Acad. Sci. U. S. A.* In press.
21. Laemmli, U. K. 1970. Cleavage of structure proteins during assembly of the head of bacteriophage T4. *Nature* (*Lond.*). 227:680–685.
22. Meyer, D. I., and B. Dobberstein. 1980. Identification and characterization of a membrane component essential for the translocation of nascent proteins across the membrane of the endoplasmic reticulum. *J. Cell Biol.* 87:503–508.
23. Miller, J. H. 1972. Experiments in Molecular Genetics, New York: Cold Spring Harbor Laboratory.
24. Neu, H. C., and L. A. Heppel. 1965. The release of enzymes from *Escherichia coli* by osmotic shock and during the formation of spheroplasts. *J. Biol. Chem.* 240:3685–3692.
25. Oliver, D. G., and J. Beckwith. 1981. *Escherichia coli* mutant pleiotropically defective in the export of secreted proteins. *Cell.* 25:765–772.
26. Osborn, M. J., J. E. Gander, E. Parisi, and J. Carson. 1972. Mechanism of assembly of the outer membrane of *Salmonella typhimurium*: isolation and characterization of cytoplasmic and outer membrane. *J. Biol. Chem.* 247:3962–3972.
27. Raibaud, O., M. Roa, C. Braunbreton, and M. Schwartz. 1979. Structure of the *malB* region in *Escherichia coli* K12. I. Genetic map of the *malK-lamB* operon. *Mol. Gen. Genet.* 174:241–248.
28. Randall, L. L., S. J. S. Hardy, and L.-G. Josefsson. 1978. Precursors of three exported proteins in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 75:1209–1212.
29. Randall-Hazelbaur, L. L., and M. Schwartz. 1973. Isolation of the bacteriophage lambda receptor from *Escherichia coli*. *J. Bacteriol.* 116:1436–1446.
30. Roggenkamp, R., B. Kustermann-Kuhn, and C. Hollenberg. 1981. Expression and proc-

essing of bacterial β-lactamase in the yeast *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* 78:4466–4470.
31. Rothman, J. E., and H. F. Lodish. 1977. Synchronized transmembrane insertion and glycosylation of a nascent membrane protein. *Nature (Lond.).* 269:775–780.
32. Schnaitman, C. A. 1971. Solubilization of the cytoplasmic membrane of *Escherichia coli* by triton X-100. *J. Bacteriol.* 108:545–552.
33. Schwartz, M., M. Roa, and M. Debarbouille. 1981. Mutations that affect *lamB* gene expression at a posttranscriptional level. *Proc. Natl. Acad. Sci. U. S. A.* 78:2937–2941.
34. Shuman, H. A., T. J. Silhavy, and J. R. Beckwith. 1980. Labeling of proteins with beta-galactosidase by gene fusion—identification of a cytoplasmic membrane component of the *Escherichia coli* maltose transport system. *J. Biol. Chem.* 255:168–174.
35. Silhavy, T. J., E. Brickman, P. J. Bassford, M. J. Casadaban, H. A. Shuman, V. Schwartz, L. Guarente, M. Schwartz, and J. R. Beckwith. 1979. Structure of the *malB* region in *Escherichia coli* K12. 2. Genetic Map of the *malE, F, G* operon. *Mol. Gen. Genet.* 174:249–259.
36. Talmadge, K., J. Brosius, and W. Gilbert. An internal signal sequence directs secretion and processing of proinsulin in bacteria. *Nature (Lond.).* 294:176–178.
37. Talmadge, K., J. Kaufman, and W. Gilbert. 1980. Bacteria mature preproinsulin to proinsulin. *Proc. Natl. Acad. Sci. U. S. A.* 77:3988–3992.
38. Walter, P., and G. Blobel. 1980. Purification of a membrane-associated protein complex required for protein translocation across the endoplasmic reticulum. *Proc. Natl. Acad. Sci. U. S. A.* 77:7112–7116.
39. Walter, P., and G. Blobel. 1981. Translocation of proteins across the endoplasmic reticulum. III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes. *J. Cell Biol.* 91:557–561.