A hybrid protein that comprises the β-lactamase signal peptide fused precisely to chicken muscle triosephosphate isomerase is not secreted into the periplasm of Escherichia coli. The protein can be secreted, however, if an arginine residue at position 3 of the isomerase is replaced by either a serine or a proline residue. In contrast, replacement of a neighboring lysine residue has no effect on secretion of the protein. Furthermore, if the arginine is removed from position 3 to generate a secreted protein, but is then reintroduced in place of the neighboring lysine, the blockade to secretion is re-established. The singular effect of the arginine residue on secretion does not result from the role this residue plays in the formation or stabilization of the native isomerase structure: mutational alterations remote from the N terminus of the isomerase that prevent the proper folding of the protein do not relieve the block to secretion. The finding that an arginine residue prevents secretion while a lysine residue does not, suggests that basic residues near the mature N terminus of a secreted protein must be deprotonated if orderly export is to occur. This implies that the signal peptide along with the N-terminal portion of the mature protein partitions directly into the lipid bilayer in the course of the secretory process.

Three types of models have been proposed to explain the almost universal requirement for an N-terminal signal peptide in protein secretion. First, the signal hypothesis envisions that the signal peptide serves primarily as a recognition element. The peptide interacts with a specific receptor such as the signal recognition particle that recruits the protein to a secretory apparatus (1, 2). Once recruited, the protein is extruded linearly across the membrane through a proteinaceous channel (3). In contrast, the membrane trigger hypothesis speculates that the signal peptide allows a secretable protein to undergo a series of conformational changes that propel the protein through the membrane (4). These conformational changes are mediated first by an interaction between the signal peptide and the nascent presecreted protein, then by an interaction of this species with the lipid bilayer, and last, by the proteolytic removal of the signal peptide on the distal side of the membrane. Each conformational change moves the protein along the secretory pathway by virtue of the different partitioning of the various conformations between aqueous and lipid phases. Finally, the direct transfer model (5), the loop model (6), and the helical hairpin hypothesis (7) each suggests that the signal peptide initiates protein secretion by dissolving directly in the lipid bilayer of the membrane. The free energy of insertion of the signal peptide into the membrane pulls the N-terminal residues of the mature protein into the bilayer. Transfer of the remaining residues of the protein into and through the membrane is then energetically more or less neutral, the entry of each residue (or group of residues) into the lipid bilayer being matched by the exit of another residue (or group of residues) on the opposite side of the membrane.

In the preceding paper (8) we demonstrated that the β-lactamase signal peptide alone is not sufficient to drive the normally cytoplasmic enzyme triosephosphate isomerase across the inner membrane of Escherichia coli, although the isomerase can be secreted if a few residues of the mature β-lactamase intervene between the signal peptide and the isomerase. We were thus able to delineate a region of constrained properties that immediately follows the signal peptide, and that encompasses about the first 14 residues of the mature protein. Such a region had been anticipated by those models of protein secretion that postulate the co-ordinate insertion of the signal peptide and the beginning of the mature protein directly into the lipid bilayer (i.e. the direct transfer model, the loop model, and the helical hairpin hypothesis). Consistent with these findings, other studies in both prokaryotes (9–13) and eukaryotes (14, 15) have demonstrated that portions of the mature protein, especially near the N terminus, influence the efficiency of secretion. Here we examine secretion of triosephosphate isomerase in the absence of sequences from the mature β-lactamase in order to define further the characteristics of a mature protein that dictate its secretion competence in E. coli.

We have considered two explanations for the failure of the β-lactamase signal peptide alone to direct the secretion of triosephosphate isomerase. First, when the isomerase is not secreted, it has adopted what appears to be its native tertiary structure in the cytoplasm (8). Folding of the protein thus may preclude its subsequent secretion if translocation across the bilayer requires an unfolded or an alternative conformation. Indeed, Randall and Hardy (16) have correlated the folding of maltose-binding protein into its native structure in the cytoplasm with a loss of secretion competence, and have suggested that the signal peptide acts to delay the folding of the preprotein until interaction with the secretory apparatus has occurred. As a second possibility, we have considered that the positively charged amino acid residues that pepper the N terminus of the isomerase might interfere with secretion of
this enzyme. Proteins that are secreted from bacteria generally have few basic amino acids near the N terminus of the mature protein. The addition of arginine or lysine residues to the N termini of several secreted proteins has been found to inhibit their export in prokaryotes (9, 10, 12, 13).

The results reported here indicate that the secretion incompetence of triosephosphate isomerase is not due to its rapid folding in the cytoplasm. The block to secretion is localized within the first 14 residues of the mature protein, and is relieved when an arginine residue at position 3 (R3) is replaced by an uncharged residue. It is not just the unit positive charge of the arginine residue at neutral pH that inhibits secretion of the isomerase, however, if, having replaced the arginine at position 3 by proline (with the result that the protein is secreted), an arginine residue is reintroduced in place of the proline, secretion is blocked anew. This finding, that an arginine residue inhibits secretion while an identically placed lysine residue does not, suggests that these residues may be required in the nascent chain for secretion.

Experimental Procedures

Bacterial Strains, Bacteriophage, Plasmids, and Media—E. coli strains DH1 (endA1, hsdR17, supE44, thi-1, recA1, gyrA96, relA1 (18), DSF02 (λrha-p42::Tn10), pSF1, pyrD, strA, galB, sup his recA endA2) (19), and JM101 (λарpl, supE44, thi-1, recD, proAB, lacF, lacZΔM15) (20), the bacteriophage M13mp11 (20), and plasmids pBO7 and pB11T (8), pX1 (19), and pAP5 have been described elsewhere. Except when otherwise indicated, bacterial media were prepared as described in Maniatis et al. (22). Dry media were purchased from Difco Laboratories (Detroit, MI). The salts NaCl, NaHPO₄, KH₂PO₄, NaCl, MgCl₂, MgSO₄, and CaCl₂ were obtained from Sigma or Mallinkrodt Inc. α-Lactose was purchased from Fisher Scientific Co. L-Amino acids and antibiotics were purchased from Sigma.

Enzymes and Chemicals—All restriction endonucleases were purchased from New England Biolabs (Beverly, MA), Bethesda Research Laboratories (Gaithersburg, MD), or Boehringer Mannheim. Klenow fragment of DNA polymerase 1 and T4 DNA ligase were purchased from either New England Biolabs or Bethesda Research Laboratories. S1 nuclease was purchased from Boehringer Mannheim. Buffers for these enzymes were either used as provided or prepared as described elsewhere (22). Low-temperature agarose was purchased from FMC BioProducts (Rockland, ME). Nucleoside triphosphates, α-32P and γ-32P methylene diphosphonate, dNTPs, and the codons for the β-lactamase signal peptide and the codons for the triosephosphate isomerase (either the wild-type codons or the R3P altered codons) were purchased from Amersham Corp. Oligonucleotide-directed mutagenesis was performed with the appropriate plasmids were grown overnight in YT medium containing thiamine hydrochloride (2 pg/ml), and appropriate antibiotics.
Amino Acid Substitution Prevents Protein Secretion

Growth was allowed to continue at 37 °C until the optical density (OD), was between 0.5 and 1.0. After the cultures had been chilled on ice for 10 min, the bacteria were isolated by centrifugation at 5000 x g for 10 min at 4 °C. Cell pellets were washed at 4 °C by suspension in chilled M9 salts (5 ml) and were resuspended by centrifugation (5000 x g for 10 min at 4 °C). Cell fractions were then generated as described previously (8). The periplasmic and soluble spheroplast fractions of each strain were assayed for the presence of triosephosphate isomerase enzymatic activity. Triosephosphate isomerase activity was determined at pH 7.6 by the method of Plaut and Knowles (25) using D-gluceraldehyde 3-phosphate as the substrate.

RESULTS

Mutations That Alter Tertiary Structural Stability Do Not Affect Secretion of Triosephosphate Isomerase—Plasmid pB0T encodes a hybrid protein that comprises the β-lactamase signal peptide fused precisely to chicken muscle triosephosphate isomerase. This hybrid protein is not secreted into the periplasm of E. coli, and the preprotein accumulates in the cytoplasm (8). To examine the possibility that the isomerase is not secreted because it folds rapidly into a conformation that is not competent for export, we have introduced structure-altering mutations into the isomerase and have determined the cellular location of the resulting constructs (Table I).

In mutant D226N, the aspartate residue at position 226 of the native isomerase has been replaced by asparagine. The rationale for this change derives from consideration of the primary and tertiary structure of triosephosphate isomerase. Aspartate 226 is a conserved residue in eight of the nine sequenced triosephosphate isomerases (26, 27), and the crystal structures of the chicken and yeast enzymes suggest that this residue forms a structurally important ion pair with a similarly conserved arginine residue at position 3 (R3) (28, 29).

Aspartate 226 is a conserved residue in eight of the nine sequenced triosephosphate isomerases (26,27), and the crystal structures of the chicken and yeast enzymes suggest that this residue forms a structurally important ion pair with a similarly conserved arginine residue at position 3 (R3) (28, 29). (In the one exceptional case, that of the triosephosphate isomerase from trypanosomes, both R3 and D226 are replaced by uncharged residues (P and N, respectively: see Ref. 27).)

The coordinate replacement of these two conserved residues (Table I).

Although the D226N mutation disrupts the normal structure of the isomerase, this change could be too subtle to influence secretion of the protein. Consequently, a more drastic alteration was made that removes residues 22-69 of the isomerase, and replaces them with a single alanine residue (226-69A). Since triosephosphate isomerase monomers normally fold into reasonably self-contained domains, this brutal deletion of 47 residues should abolish all elements of the native tertiary structure of the isomerase. Moreover, the deletion occurs early in the translated sequence and is expected to interfere from the outset with the normal pathway of events that occur in the folding of this protein. Yet, as is shown in Fig. 3A, even this truncated protein is not secreted into the periplasm. A single species, corresponding in size to the precursor, is found in the spheroplast fraction. As expected, this polypeptide is digested by protease in lysed spheroplasts, indicating that the native (protease-resistant) struc-

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Description</th>
<th>Effect on secretion</th>
<th>Protease sensitivity</th>
<th>Enzymatic activity</th>
<th>Amino acid sequence of altered isomerase</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>Wild-type</td>
<td>Not secreted</td>
<td>Resistant</td>
<td>Active</td>
<td>Ala-Pro-Arg-Lys...Asp...</td>
</tr>
<tr>
<td>D226N</td>
<td>Aspartate 226 to asparagine</td>
<td>Not secreted</td>
<td>Resistant</td>
<td>Active</td>
<td>Ala-Pro-Arg-Lys...Asn...</td>
</tr>
<tr>
<td>Δ22-69A</td>
<td>47 amino acid deletion</td>
<td>Not secreted</td>
<td>Resistant</td>
<td>Active</td>
<td>Ala-Pro-Arg-Lys...Asp...</td>
</tr>
<tr>
<td>R3S</td>
<td>Arginine 3 to serine</td>
<td>Not secreted</td>
<td>Resistant</td>
<td>Active</td>
<td>Ala-Pro-Arg-Lys...Asp...</td>
</tr>
<tr>
<td>R3P</td>
<td>Arginine 3 to proline</td>
<td>Secreted</td>
<td>Active</td>
<td>Ala-Pro-Arg-Lys...Asp...</td>
<td></td>
</tr>
<tr>
<td>R3P,D226N</td>
<td>Arginine 3 to proline and aspartate 226 to asparagine</td>
<td>Secreted</td>
<td>Active</td>
<td>Ala-Pro-Arg-Lys...Asp...</td>
<td></td>
</tr>
<tr>
<td>K4Q</td>
<td>Lysine 4 to glutamine</td>
<td>Not secreted</td>
<td>Resistant</td>
<td>ND</td>
<td>Ala-Pro-Arg-Gln...Asp...</td>
</tr>
<tr>
<td>K4Q,D266N</td>
<td>Lysine 4 to glutamine, aspartate 226 to asparagine</td>
<td>Not secreted</td>
<td>Resistant</td>
<td>ND</td>
<td>Ala-Pro-Arg-Gln...Asp...</td>
</tr>
<tr>
<td>R3P</td>
<td>Arginine 3 to proline and lysine 4 to arginine</td>
<td>Not secreted</td>
<td>Resistant</td>
<td>Active</td>
<td>Ala-Pro-Arg...Asn...</td>
</tr>
<tr>
<td>R3K</td>
<td>Arginine 3 to lysine and aspartate 226 to asparagine</td>
<td>Not secreted</td>
<td>Resistant</td>
<td>Active</td>
<td>Ala-Pro-Lys-Lys...Asp...</td>
</tr>
<tr>
<td>R5K,D226N</td>
<td>Arginine 3 to lysine and aspartate 226 to asparagine</td>
<td>Not secreted</td>
<td>Resistant</td>
<td>Active</td>
<td>Ala-Pro-Lys-Lys...Asn...</td>
</tr>
</tbody>
</table>

* Differences from the wild-type sequence are in bold type.
* ND, not determined.
Amino Acid Substitution Prevents Protein Secretion

20085

The Role of Basic Residues in the Inhibition of Secretion—Among the first 14 amino acids of triosephosphate isomerase, positively charged residues fall at positions 3 (arginine), 4 (lysine), and 12 (lysine). There are no acidic residues in this region, so the N terminus of the isomerase carries a charge of +3 at physiological pH (Fig. 5). The role of these charged residues in the secretion of the isomerase has been assessed from the behavior of the mutants described below (summarized in Table I).

When arginine 3 is changed to either a serine or a proline, a fraction of each of the resulting mutant isomerases R3S and R3P is secreted into the periplasm (Fig. 6, A and C). Aside from the processed periplasmic species, a substantial portion of each mutant is processed but retained with the spheroplast fraction, and some preprotein is also found in the spheroplast (Fig. 6, A and C). The processed species in the spheroplast fraction is a form that has been more or less completely translocated, but not released from the inner membrane. This location is demonstrated by protease treatment of whole spheroplasts and detergent-solubilized spheroplasts, the results of which are illustrated in Fig. 6, B and D. Behavior of this type (that is, translocation but not release) has been observed for other β-lactamase/triosephosphate isomerase hybrid proteins (8), for the wild-type β-lactamase secreted from cells grown at lower temperatures (30), and for several normally secreted proteins that have been altered by mutation (31–33). We presume that this results from improper folding of the protein that precludes release from the membrane. The protease treatment also reveals that the isomerase moiety of the R3S and R3P preproteins has not adopted its native structure in the cytoplasm: both preproteins are completely digested by the protease when the spheroplast membrane is

When the spheroplast membrane is intact the molecule is protected from proteolysis (Fig. 3B, lane S + K): the preprotein is thus completely intracellular. We must conclude from these results that the secretion incompetence of these hybrid proteins is not due to their rapid folding into a stable native structure.

The Unsecretability of Triosephosphate Isomerase Derives from Residues Near the N Terminus—Since the failure of the isomerase to be secreted is not explained by its rapid folding,
disrupted by detergent (Fig. 6, B and D, lanes LS + K). It is likely, therefore, that the R3 alterations allow secretion of the isomerase by eliminating the primary arginine-based block to secretion and, possibly, by preventing formation of the enzyme's native structure.

To explore the generality of the beneficial R3P change, this mutation was engineered into the β-lactamase/triosephosphate isomerase hybrid protein described in the previous section. Consonant with the results for the complete isomerase, the R3P mutation dramatically relieves the blockade to secretion in this chimera (BTB, R3P) (Fig. 7A). Proteinase K sensitivity of the proteins encoded by pBOT(K4Q), pBOT(K4Q,D226N), and pBOT(R3P,K4R). A, location of the proteins encoded by pBOT(K4Q). Cells were pulse-labeled for 30 s with [35S]methionine and chased with an excess of unlabeled methionine for 30 s, 4 min, and 12 min. The samples were separated into spheroplast-associated (S) and periplasmic (P) proteins. Triosephosphate isomerase species were immune precipitated and analyzed by denaturing polyacrylamide gel electrophoresis and fluorography. w-t, native triosephosphate isomerase. B, proteinase K sensitivity. Cells were pulse-labeled for 90 s with [35S]methionine and chased with an excess of unlabeled methionine for 30 s, 4 min, and 12 min. The samples were then separated into spheroplast-associated (S) and periplasmic (P) proteins. Triosephosphate isomerase species were immune precipitated and analyzed by denaturing polyacrylamide gel electrophoresis and fluorography.

In contrast to the replacement of R3 by uncharged residues, a similar change in the lysine residue at position 4, as in K4Q, does not result in protein secretion (Fig. 7A). Proteolysis experiments suggest that the K4Q mutation does not fully disrupt the normal structure of the isomerase: the isomerase domain of the prepolymer does show some resistance to proteolysis (Fig. 7B). To eliminate any secondary effects due to folding that may mask the full effect of substitution for the lysine, the D226N change described earlier that blocks the folding of the isomerase was combined with the K4Q change. The double mutant (K4Q,D226N) is, however, still not secreted (Fig. 7C). Complete protease digestion after detergent solubilization of the spheroplasts demonstrates that the tertiary structure of the K4Q,D226N protein has been disrupted (Fig. 7D, lane LS + K). (The D226N mutation itself does not prevent protein secretion: the R3P,D226N double mutant is secreted slightly more rapidly than that of the R3P protein.) We must conclude, therefore, that R3 plays a different role from K4 in blocking the secretion process.

Since the R3S and R3P mutant isomerases are secreted, whereas the K4Q mutant isomerase is not, it seemed that either the position or the higher pK of the arginine could be responsible for the inhibition of protein export. To examine this question, an arginine residue was reintroduced to the secretable R3P protein in place of its lysine residue at position 4. In the resulting double mutant R3P,K4R, the N terminal residue was indicated.
Amino Acid Substitution Prevents Protein Secretion

charge is the same as that of the secreted R3P protein, and this charge is located at the same position in the primary structure of the isomerase. The only difference is that this charge is now carried by an arginine residue rather than a lysine residue. As Fig. 7E shows, the arginine for lysine substitution completely abolishes secretion of the hybrid protein, which now remains as the preprotein in the cytoplasm.

The preprotein is protected from proteolysis in intact spheroplasts, and is digested when the spheroplast membrane is disrupted by detergent (Fig. 7F). The molecule is therefore intracellular and has not adopted the native structure of the isomerase. As expected from its sensitivity to proteolysis, this polypeptide is not enzymatically active.

The finding that a lysine residue at position 4 does not block secretion of triosephosphate isomerase whereas an arginine residue does, led us to ask whether the isomerase might be rendered secretable by simply replacing the arginine residue at position 3 with a lysine. The resulting R3K mutant is, however, not secreted (data not shown). The secretory mechanism can therefore tolerate a single lysine residue at the N terminus of the isomerase, but either an arginine or 2 lysine residues in this region is unacceptable.

DISCUSSION

To examine the features of a protein that influence its secretion into the periplasm of E. coli we have made a series of changes in a nonsecreted hybrid protein that comprises the \( \beta \)-lactamase signal peptide fused precisely to the complete triosephosphate isomerase from chicken muscle. Some of the alterations have been designed to disrupt the tertiary structure of the isomerase while leaving the N-terminal sequence of the protein unchanged, while other alterations have changed the basic amino acid residues that lie near the N terminus of the mature isomerase (Table I).

Our results show that formation of the native structure of the isomerase is not responsible for the secretion incompetence of the protein. Rather, the nature of the N-terminal residues of the isomerase is the key factor in determining whether the \( \beta \)-lactamase signal peptide can direct the isomerase into the periplasm. Observations consistent with this conclusion have been reported for the secretion of other proteins in prokaryotes (9-13) and eukaryotes (14, 15). It has been suggested by von Heijne (17) that positive charge may not be tolerated near the N terminus of proteins that are secreted in bacteria, and it has been shown that the introduction of basic residues into this region can inhibit secretion in lower organisms (9, 10, 12). We have found here, however, that the basic side chains of arginine and lysine are not equivalent in terms of their effect on secretion. Thus, whereas replacement of arginine 3 with serine or proline allows secretion of the isomerase (in mutants R3S or R3P), replacement of the vicinal lysine at position 4 with a glutamine residue (in K4Q) has no beneficial effect. Moreover, if, after replacing the arginine residue at position 3 (thus generating the secreted mutant R3P), an arginine is reintroduced in place of the lysine residue at position 4 (to create R3P,K4R), the secretion blockade is re-established. The arginine residue evidently prevents secretion of the isomerase while an identically positioned lysine residue does not.

The beneficial effect of removing the arginine residue at position 3 does not derive from the role of this residue in the formation or stabilization of the native structure of the isomerase, since structure-altering mutations remote from the N terminus do not allow the enzyme to be secreted. Neither a point mutation that breaks a structurally important ion pair between arginine 3 and aspartate 226 at the distal residue nor a brutal deletion that removes from the isomerase the 47 residues that follow glycine 21 enables the \( \beta \)-lactamase signal peptide to guide secretion of the isomerase. Furthermore, the introduction of just the first 14 residues of the isomerase between the \( \beta \)-lactamase signal peptide and the mature \( \beta \)-lactamase blocks secretion of the \( \beta \)-lactamase, and this blockade is dramatically eliminated when the offending arginine is replaced by proline (see Fig. 4).

At physiological pH, both lysine and arginine side chains are essentially fully protonated (the \( pK_a \) of lysine's \( \epsilon \)-ammonium group is about 10.5, and that of arginine's guanidinium group is about 12.5), and each carries a charge of +1. Yet an arginine residue near the N terminus of the mature protein prevents export, while a similarly placed lysine residue does not. This finding indicates that the block to secretion does not derive from the interaction of the cationic group with the membrane potential (34), nor from a requirement that the N terminus of a secretable protein be bipolar (17), but rather from the primary difference between the two side chains, namely their \( pK_a \) values. The singularly deleterious effect of an arginine residue near the mature N terminus of a protein that is to be secreted therefore suggests that the residues immediately following the signal peptide must be deprotonated at some point in the secretory process. Such a requirement implies that these residues leave the aqueous phase and enter a non-polar environment such as the membrane bilayer. Indeed, the fact that monocarboxylic acids and monoamines (each having a \( pK_a \) value about 3 units away from the pH of most physiological systems) can passively cross biological membranes, whereas phosphates and guanidinium compounds cannot (35), is consistent with the difference between Lys and Arg noted here.

The length of the mature N terminus that must be buried in the bilayer is suggested by the results described in the previous paper (8). There it was shown that efficient secretion of the \( \beta \)-lactamase/triosephosphate isomerase hybrid proteins requires intervention of the first 12 amino acids of mature \( \beta \)-lactamase between the end of the signal peptide and the start of the triosephosphate isomerase. This places arginine 3 of the isomerase at position 10 of the mature hybrid protein. Thus, at least the first 14 N-terminal residues of the mature portion of a bacterial secreted protein may have to become buried in the membrane in order that protein secretion be properly initiated. (It is arguable whether this step is an early step, and evidence has been presented that the hydrophobic core of the signal sequence has its effect not during the attachment of the precursor protein to the membrane but rather during the subsequent translocation [36].) Outside this critical region, charged residues seem not to interfere with translocation across the membrane (8). Within this critical region, however, there is a limit to the hydrophilicity that is compatible with entry into the secretion pathway. Thus we find that 1 lysine residue near the N terminus of triosephosphate isomerase is tolerable, but that 1 arginine or 2 lysine residues, are not. Unsurprisingly in this context, it has been observed that acidic residues can counterbalance basic residues within the mature N terminus (8, 12, 37, 38), and we may presume that the beneficial effect of counterbalancing charges results from the formation of ion pairs that enables these charged residues to enter the membrane without the unfavorable energetic consequences of (de)protonation. Such ion pairs will be of particular benefit when they involve an arginine residue, since the energetic cost to deprotonate this residue is so high. Indeed arginine residues are never observed near the mature N terminus of secreted bacterial proteins in the absence of nearby acidic residues (for a compilation of
secreted protein sequences, see Ref. 39).

The results of the present work are difficult to reconcile with the initial events postulated either by the signal hypothesis (1, 2) or by the membrane trigger hypothesis (4). Neither model envisions the obligate insertion of the early part of the mature protein into a nonpolar environment, nor is there another basis in these theories for any discrimination between arginine and lysine residues near the N terminus of the mature protein. On the other hand, the direct transfer model (5), the loop model (6), and the helical hairpin hypothesis (7) are each based on the central tenet that the N terminus of a secreted protein partitions directly into the lipid bilayer. These models each propose that the principal function of a signal peptide is to provide the free energy necessary to pull the first residues of the mature protein into the membrane. Indeed, we should predict that proteins having a marked imbalance of charged residues at their N termini will be unsecretable, because the free energy gained upon insertion of the hydrophobic core of the signal peptide into the membrane will not be sufficient to drive the coordinate insertion of the attached N terminus of the mature protein. We therefore believe that the early events in the mechanism of protein secretion in bacteria occur much as has been postulated by these direct insertion models. Our data do not, however, address the mechanism of subsequent translocation of the protein across the bilayer. Genetic evidence suggests that a proteinaceous apparatus is ultimately involved in this process (40–42). Nevertheless, it seems that one barrier to protein secretion in bacteria derives from the direct partitioning of the N terminus of the preprotein into a nonpolar phase such as the lipid bilayer, and that one role of the signal peptide is to provide the driving force for this event.

Interestingly, in contrast to secreted proteins of prokaryotes, proteins secreted from eukaryotic cells are not necessarily characterized by having few (or balanced) charges near the N terminus of the mature protein (17, 39). Indeed, several secreted eukaryotic proteins such as rat, bovine, and human secretory proteins (43, 44), bovine and human parathyroid hormone (45, 46), human apolipoprotein A-II (14), and human β2-macroglobulin (21) have extremely basic stretches that immediately follow the signal peptide. On the basis of our present results, we predict that these proteins will be unsecretable in E. coli. In fact, human β2-macroglobulin is not secreted in E. coli: only if an acidic linker is placed at the beginning of the mature protein is secretion observed (37). This emphasizes a difference between early events in the secretory process of prokaryotes and eukaryotes.

In summary, we have shown that the residues near the N terminus of a mature protein influence its secretability, and that this influence most likely derives from the requirement that these residues partition into a hydrophobic phase. Charged residues appear particularly harmful to secretion because of the unfavorable energetic consequences of neutralization, although the formation of ion pairs between side chains of opposite charge may abate this effect (8, 12, 37, 38).

These findings are consistent with those models of protein secretion that postulate the direct insertion of the signal peptide and a portion of the mature protein into the membrane such as the direct transfer model (5), the loop model (6), and the helical hairpin hypothesis (7).

Acknowledgments—We thank Drs. A. Minsky, J. Hermes, D. Wiley, and G. Guidotti for useful discussions.

REFERENCES