A Former Amino Terminal Signal Sequence Engineered to an Internal Location Directs Translocation of Both Flanking Protein Domains

EVE PERARA and VISHWANATH R. LINGAPPA
Departments of Physiology and Medicine, and Topogen, University of California, San Francisco, California 94143

ABSTRACT To determine whether a functional amino terminal signal sequence can be active at an internal position, a hybrid gene was constructed in which the entire coding region of bovine preprolactin cDNA was inserted into chimpanzee α-globin cDNA 109 codons downstream from the initiation codon of globin. When RNA synthesized in vitro from this plasmid (pSPGP1) was translated in the rabbit reticulocyte cell-free system, a 32-kD protein was produced that was both prolactin and globin immunoreactive. When microsomal membranes were present during translation (but not when added posttranslationally), a 26-kD and a 14-kD product were also observed. By immunoreactivity and electrophoretic mobility, the 26-kD protein was identical to mature prolactin, and the 14-kD protein appeared to be the globin domain with the prolactin signal sequence attached at its carboxy terminus. From (*a*) posttranslational proteolysis in the presence and absence of detergent, (*b*) sedimentation of vesicles in the presence and absence of sodium carbonate pH 11.5, and (*c*) N-linked glycosylation of the globin-immunoreactive fragment after insertion of an Asn-X-Ser N-linked glycosylation site into the globin coding region of pSPGP1, it appears that all of the 26-kD and some of the 14-kD products, but none of the 32-kD precursor, have been translocated to the lumen of the membrane vesicles. Thus, when engineered to an internal position, the prolactin signal sequence is able to translocate both flanking protein domains.

These data have implications for the understanding of translocation of proteins across the membrane of the endoplasmic reticulum.

The mechanism by which newly synthesized secretory proteins are translocated across the membrane of the endoplasmic reticulum (ER)[1] is an unsolved problem. It is generally accepted that signal sequences play a role in this process, but it is unclear whether they serve as ligands for proteinaceous receptors in the membrane and activate a catalytic translocation mechanism (1), or whether they trigger translocation across the bilayer spontaneously, i.e., without the participation of other membrane proteins (2).

Other fundamental aspects of the translocation process are equally obscure: Is the energy to drive these events derived from the thermodynamics of protein–lipid interactions, from protein synthesis, or does proteinaceous (enzymatic) machinery in the membrane actively move the protein across the bilayer? To what extent is signal sequence function constrained by the nature of the chain being translocated and the location of the signal sequence within that chain? What role, if any, does signal sequence cleavage play in the process of translocation? Why are secretory proteins completely translocated across the bilayer, whereas integral membrane proteins only partially translocated?

Our laboratory is attempting to address such issues in protein translocation by engineering defined coding regions into various and unusual locations or orientations and using these constructions as expression probes in cell-free systems that are amenable to fractionation and reconstitution. As a first step toward understanding signal sequence function we previously demonstrated that a signal sequence coding region

[1] *Abbreviations used in this paper:* Endo H, endoglycosidase H; ER, endoplasmic reticulum.

THE JOURNAL OF CELL BIOLOGY · VOLUME 101 DECEMBER 1985 2292-2301

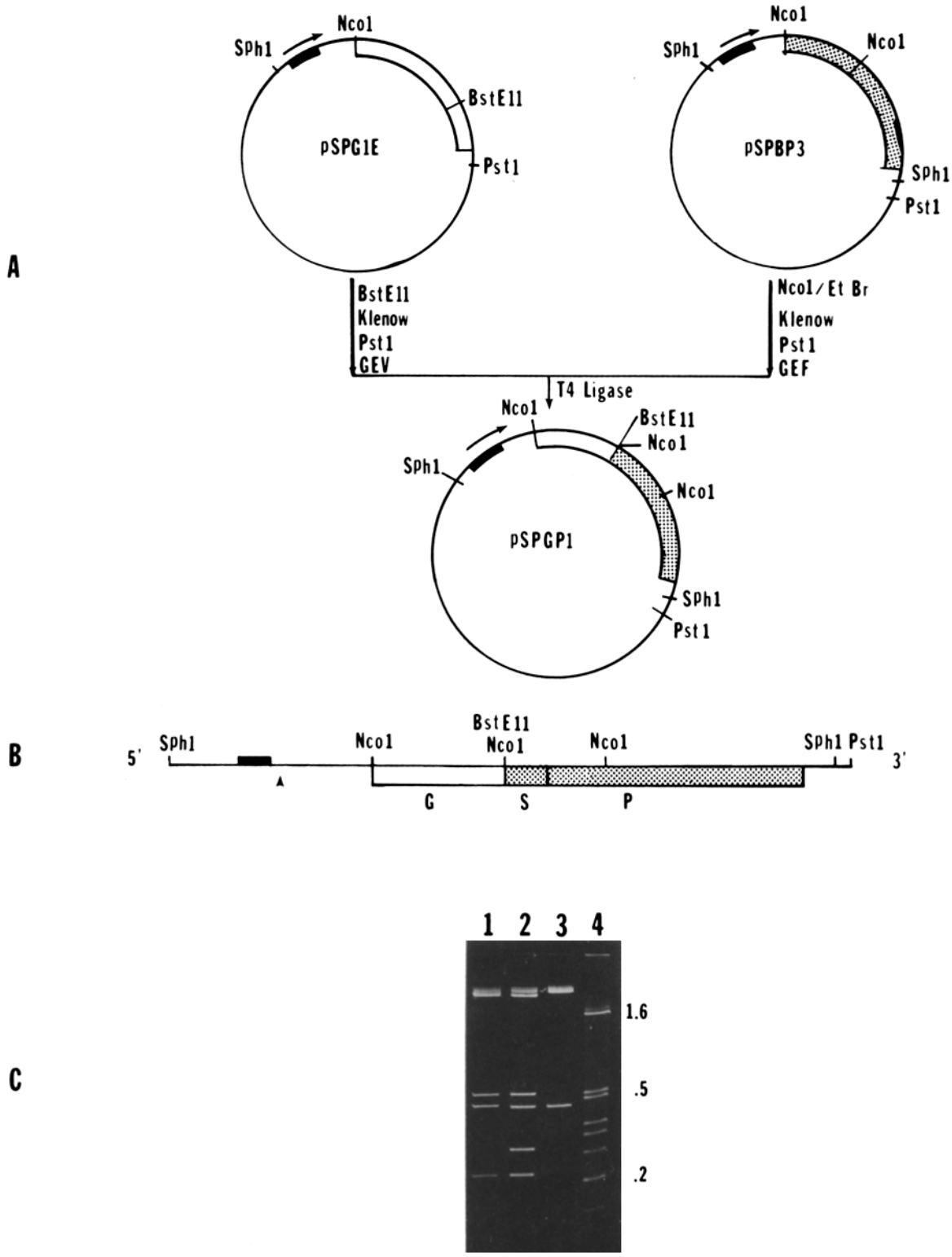

FIGURE 1 Structure of pSPGP1 fusion plasmid. (*A*) Construction scheme for pSPGP1, in which the entire coding region of bovine preprolactin was inserted into an SP6 plasmic containing chimpanzee α-globin cDNA such that the signal sequence of the resulting hybrid fusion protein is located 109 amino acids from the amino terminus (see Materials and Methods for details). Chimpanzee α-globin coding regions are indicated by white bars and preprolactin by stippled bars. The black bars indicate the SP6 promoter, and arrows on the plasmid diagrams indicate the direction of transcription. (*B*) Restriction map of the preGSP coding region of pSPGP1 with corresponding gene product represented underneath. The white bar represents the 109 amino acid globin domain (*G*) and the stippled bar the 229 amino acids of preprolactin (*P*). The 30 amino acid signal sequence of prolactin is also indicated (*S*). The upward pointing arrowhead indicates the transcription start site. (*C*) Restriction endonuclease analysis of pSPBP3, pSPGP1, and pSPG1E. Plasmid DNA (5 μg) was digested in a volume of 10 μl with NcoI and SphI. Samples were prepared and electrophoresed on a 5% polyacrylamide gel in Tris-Borate-EDTA. Lane *1*, pSPBP3; lane *2*, pSPGP1; lane *3*, pSPG1E; lane *4*, pBR322 digested with HinfI as size markers in kilobases.

engineered at the amino terminus was sufficient to translocate the normally cytoplasmic protein globin across microsomal membranes in cell-free systems (3). Similarly, we demonstrated the functional distinction of signal and stop transfer sequences, and that stop transfer sequences require preceding signal sequences for expression of transmembrane integration by the ability or inability of these sequences, engineered into various coding regions, to convert nascent globin domains into secreted versus integral transmembrane proteins (4).

To determine whether a previously amino terminal signal sequence can function from an internal position in a protein, we have engineered the signal sequence of bovine preprolactin between the coding regions for globin and prolactin, and then engineered an N-linked glycosylation site into the globin coding region. Here we present the surprising behavior of these molecules in cell-free systems and their implications for a variety of issues in protein translocation.

MATERIALS AND METHODS

Materials: All restriction endonucleases, nuclease Bal 31, calf intestinal alkaline phosphatase, SP6 RNA polymerase, T4 DNA ligase, and Klenow fragment of *E. coli* DNA polymerase were from Boehringer Mannheim Diagnostics, Inc., Houston, TX or from New England BioLabs, Beverly, MA. RNase inhibitor was from Promega Biotec, Madison, WI; staphylococcal protein A–Sepharose was from Pharmacia, Inc., Piscataway, NJ; rabbit anti–human globin serum was from Cappel Laboratories, Cochranville, PA; rabbit anti–ovine prolactin was from United States Biochemical Corp., Cleveland, OH; proteinase K was obtained from Merck, FRG; endoglycosidase H (Endo H) and [^{35}S]-methionine (translation grade, >800 Ci/mmol) were from New England Nuclear, Boston, MA; Nikkol (octa-ethleneglyco-mono-*n*-dodecyl ether, a nonionic detergent) was from Nikko Chemicals Co., Ltd., Tokyo, Japan. Plasmid pSPBP3 was constructed by William Hansen, Department of Biochemistry and Biophysics, University of California at San Francisco, using bovine preprolactin cDNA (28). All globin-encoding plasmids were derived from pMC18 (4).

Construction of Globin–Prolactin Fusion Plasmid, pSPGP1: As depicted in Fig. 1, plasmid pSPBP3, containing the entire coding region for bovine preprolactin, was linearized with NcoI in the presence of ethidium bromide and the overhang filled in by treatment with *E. coli* DNA polymerase I Klenow fragment in the presence of all four dNTPs. The plasmid was then cut with PstI and the 850-base pair (bp) fragment containing the preprolactin coding region was purified on, and eluted from, a 1% low melting point agarose gel. Plasmid pSPG1E was cut with BstEII, the 5′ overhang filled in as described above, then cut with PstI and the 3-kilobase (kb) vector gel purified. The purified pSPBP3 fragment and pSPG1E vector were treated with T4 DNA ligase. After transformation of *E. coli*, plasmid DNA was prepared from individual ampicillin-resistant colonies and screened by restriction enzyme analysis with NcoI and SphI for appropriate sized fragments (see Fig. 1 *C*).

Insertion of N-Linked Glycosylation Site into pSPGP1: To introduce an N-linked glycosylation site into the globin domain of pSPGP1 we used plasmid pSPSG1 (see Fig. 5),[2] an exact fusion of the β-lactamase signal sequence and chimpanzee α-globin in which a synthetic oligonucleotide encoding Ala-His-*Asn-Gly-Ser*-Gly-Ser-Gly has been inserted into the BssHII site of the globin coding region. The translation product of this plasmid is translocated across the ER membrane and is core glycosylated in vitro.[2] The region encoding the β-lactamase signal sequence was deleted by digestion with NcoI and BglII, treated with Klenow fragment to fill in the 5′ overhangs, and recircularized with T4 ligase. The resulting plasmids were used to transform *E. coli*, and plasmid DNA was prepared from individual ampicillin-resistant colonies and screened by restriction enzyme analysis with HindIII. The 430-bp HindIII fragment of the positive clone, pSPSG3, was inserted into pSPGP1 which had been digested to completion with HindIII and treated with calf intestinal alkaline phosphatase to prevent self-ligation. *E. coli* were transformed and DNA prepared from individual ampicillin-resistant colonies was screened with NcoI to determine the presence and correct orientation of the HindIII insert.

Rabbit Reticulocyte Lysate Transcription-coupled Translation: SP6 plasmids were transcribed in vitro (5) at a concentration of 0.2 mg/ml in a reaction mix containing 40 mM Tris-HCl pH 7.5, 6 mM $MgCl_2$, 2 mM spermidine, 10 mM dithiothreitol, 25 μg/ml calf-liver tRNA, 0.5 mM each ATP, CTP, and UTP, 0.1 mM GTP, 0.5 mM GpppG, 0.9 U/μl RNase inhibitor, and 0.4 U/μl SP6 RNA polymerase. Reactions were carried out at 40°C for 1 h and aliquots used directly in transcription-linked translations in the rabbit reticulocyte lysate cell-free system at a concentration of 20%. Translation reactions were carried out in 20–200 μl vol that contained 43% rabbit reticulocyte lysate (prepared as described, reference 6), 100 mM KCl, 2 mM $MgCl_2$, 0.9 mM GTP, 1 mM ATP, 10 mM creatine phosphate, 0.2 mM each of 19 amino acids minus methionine, 16 mM Tris HCl (pH 7.5), 0.44 mM spermidine, 2 mM dithiothreitol, 0.4 mg/ml creatine phosphokinase, 0.1 mg/ml calf-liver tRNA, and 1 mCi/ml [^{35}S]methionine. Reaction mixtures were incubated at 24°C for 60 min.

Protein Processing and Translocation Assays: In vitro transcripts of SP6 plasmids were translated separately in a rabbit reticulocyte cell-free system in the presence or absence of intact dog pancreas rough microsomes (prepared as described, reference 7). Translation products were immunoprecipitated and separated by SDS PAGE. Bands were localized by autoradiography and quantitated by densitometer scanning of preflashed film, using an LKD

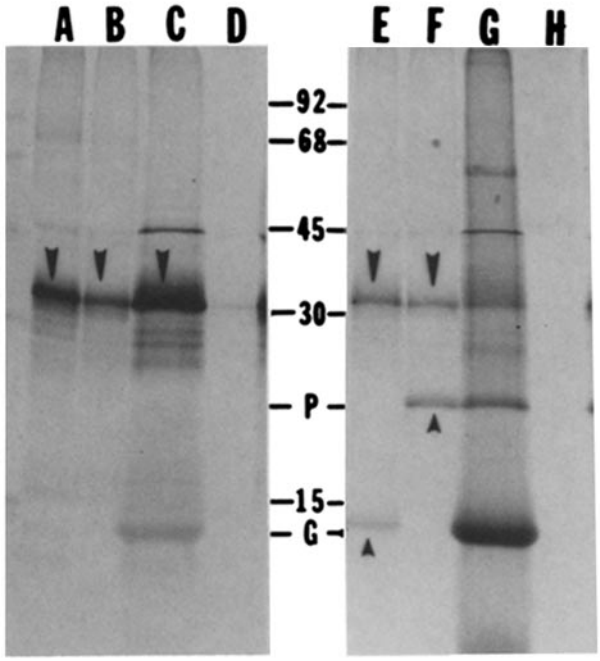

FIGURE 2 In vitro translation products of pSPGP1. Plasmid DNA was transcribed by SP6 RNA polymerase in a volume of 10 μl and a 2-μl sample translated in the rabbit reticulocyte cell-free system as described in Materials and Methods. After 1 h at 23°C, aliquots (0.5 μl) were immunoprecipitated with either rabbit anti–human globin (lane *A*), rabbit anti–ovine prolactin (lane *B*), or normal rabbit serum (lane *D*). Samples were prepared and separated by SDS PAGE and bands viewed by autoradiography. One aliquot was not immunoprecipitated but applied directly to SDS PAGE (lane *C*). Translations were simultaneously performed as above except that dog pancreas rough microsomes were present at a concentration of 5 A_{280} U/ml. Samples were prepared as before. Lane *E*, anti-globin immunoprecipitate; lane *F*, anti-prolactin immunoprecipitate; lane *G*, total translation products; lane *H*, nonimmune serum immunoprecipitate. Large arrowheads pointing downward indicate pre GSP and the arrowheads pointing upward show the cleavage products, GS1 and P1 in lanes *E* and *F*, respectively. Positions of authentic mature prolactin and globin (143 amino acids) are indicated as P and G, respectively, in the center lane where molecular weight markers are also indicated in kilodaltons. Note: the low molecular weight band in lanes *C* and *G* (total translation products) that runs at the same apparent molecular weight as the globin marker, G, is rabbit globin synthesized in the in vitro system from mRNA in the rabbit reticulocyte lysate. The different relative intensities of this band in lanes *C* and *G* are due to different exposure times for lanes *A–D* and lanes *E–H*. Lanes *E–H* required a longer exposure time because of the inhibition of protein synthesis by the addition of dog pancreas rough microsomes.

[2] Perara, E., D. Cashman, Y. W. Kan, and V. R. Lingappa. Manuscript in preparation.

2202 Ultroscan Laser Densitometer from LKB Instruments, Inc., Gaithersburg, MD.

To determine percentage processing of pSPGP1 translation products in vitro, translation reactions were carried out in the presence of varying membrane concentrations (0, 1.25, 2.5 or 5.0 A_{280} U/ml). Reactions were stopped after 1 h at 24°C by chilling on ice. Samples were split in half (5 μl each) and immunoprecipitated with either globin or prolactin antiserum. Samples were prepared and subjected to SDS PAGE (7). The intensities of preGSP, GS1, and P1 bands were quantitated by densitometry. Percentage processing of preGSP to GS1 was determined by [(GS1 × 11/4)/preGSP + (GS1 × 11/4)] × 100 and processing to P1 by [(P1 × 11/7)/preGSP + (P1 × 11/7)] × 100 to compensate for the differential methionine contents of the three proteins (preGSP contains eleven, P1 seven, and GS1 four).

Translocation of preGSP and its cleavage products, GS1 and P1, across microsomal membranes was determined by two independent methods: (*a*) sensitivity of translation products to protease digestion and (*b*) addition of carbohydrate to the Asn-X-Ser N-linked glycosylation site engineered into the globin coding region.

Protease protection experiments were done as follows: After 1 h at 24°C, translation reaction mixtures were chilled on ice, adjusted to 10 mM $CaCl_2$, and divided into equal aliquots of 5 or 10 μl. Some were treated with proteinase K (dissolved in 10 mM $CaCl_2$, 50 mM Tris pH 7.5 and preincubated at 37°C for 15 min) at a final concentration of 0.1–0.4 mg/ml either in the presence or absence of 1% Nikkol (a nonionic detergent used to disrupt the lipid bilayer). All samples were incubated at 0°C for 1 h. Proteinase digestion was stopped by the addition of 2 mM phenylmethylsulfonyl fluoride and immediately transferred to 4–5 vol 1% SDS 0.1 M Tris-HCl (pH 8.9), preheated to 100°C, then incubated at 100°C for 10–15 min. Samples were diluted 20-fold in a solution of 1% Triton, 0.1 M Tris pH 8.0, 10 mM EDTA, 100 mM NaCl, subjected to immunoprecipitation with either 0.5 μl anti-prolactin antiserum or 4 μl anti-globin antiserum, and protein A-Sepharose-CL4B, followed by SDS PAGE. Bands were viewed by autoradiography and quantitated by densitometric scanning. Percentage protection was determined for all bands by [band (+ protease)/band (− protease)] × 100.

Endo H digestion was used to determine core glycosylation of translation products. Endo H removes simple N-linked carbohydrates, causing a shift to a lower apparent molecular weight upon SDS PAGE. Translation products of pSPgGP1 obtained in the presence or absence of membranes were immunoprecipitated and eluted in 100 μl of 0.1 M sodium citrate pH 5.5, 0.1% SDS. Samples were heated to 100°C for 2 min, supernates removed and divided into two aliquots. Endo H added to one aliquot to a final concentration of 1 μg/ml, and both aliquots incubated at 37°C for 12 h. After digestion, 10 μg of carrier bovine serum albumin (BSA) was added, and samples were chilled and adjusted to 15% ice-cold trichloroacetic acid, precipitates collected by centrifugation, and samples prepared for SDS PAGE as usual.

Membrane Sedimentation and Carbonate Extraction: Translation products (10 μl) obtained in the presence of 4 A_{280} U/ml dog pancreas rough microsomes were diluted 250-fold (to 2.5 ml) with either ice-cold 0.1 M sodium carbonate pH 11.5 (7) or ice-cold 0.25 M sucrose, 5 mM $MgCl_2$, 1 mM dithiothreitol, 50 mM triethanolamine pH 7.4, 0.1 M KCl and incubated at 0°C for 30 min. The samples were centrifuged at 0°C for 1 h at 50,000 rpm in polycarbonate tubes in a Beckman 70.1 Ti rotor (Beckman Instruments, Inc., Palo Alto, CA). The supernatants were removed, the sides of the tube carefully dried with a kimwipe, and the membrane pellets dissolved in a 1% Triton buffer. The pH of the carbonate samples was adjusted to 7–7.5 with acetic acid, and all samples were immunoprecipitated with either anti-prolactin or anti-globin serum, prepared and subjected to SDS PAGE and autoradiography as usual.

RESULTS

Construction of Globin-Prolactin Fusion Plasmid

Fig. 1, *A* and *B* shows the scheme by which the entire coding region for bovine preprolactin was inserted, in frame, 109 codons downstream from the initiation codon of a chimpanzee α-globin cDNA clone previously engineered behind the SP6 promoter (pSPG1E). Putative positive colonies were selected for ampicillin resistance and screened for the presence of restriction fragments of predicted sizes of the globin, signal sequence, and prolactin coding domains, as well as the SP6 promoter. These characteristic restriction fragments from both parent plasmids, pSPG1E and pSPBP3, and from the new construction, pSPGP1, are displayed in Fig. 1 *C*.

Membrane-dependent Processing of Nascent Pre GSP to GS1 and P1

Upon expression in a transcription-coupled rabbit reticulocyte lysate cell-free translation system, pSPGP1 encoded a fusion protein of ~32 kD with both globin and prolactin immunoreactivity (Fig. 2, lanes *A–D*) called preGSP. When translation reactions were supplemented with microsomal membranes two additional translation products, not present in the absence of membranes, were seen after electrophoresis on polyacrylamide gels in SDS with subsequent autoradiography (Fig. 2, lane *G*). One of these bands, termed P1, was found to be anti-prolactin but not anti-globin immunoreactive and to co-migrate with authentic mature bovine prolactin (Fig. 2, lane *F*). The other product, termed GS1, was anti-globin but not anti-prolactin immunoreactive, and migrated with an apparent molecular weight slightly greater than that of authentic full-length globin (14 kD, Fig. 2 lane *E*). When membranes were added after completion of protein synthesis with further incubation, neither of these bands were generated (data not shown). The difference in the relative intensities of the bands in the autoradiographs is due to the methionine distribution in the [^{35}S]methionine-labeled, newly synthesized proteins. Upon processing, of the eleven methionines in pre GSP, P1 contains seven and GS1 only four.

Fig. 3 demonstrates that when membranes were present during translation, GS1 and P1 were generated in a 1:1 ratio and that the percentage processing of both products increased correspondingly with membrane concentration. These data indicate that GS1 and P1 are generated from a common

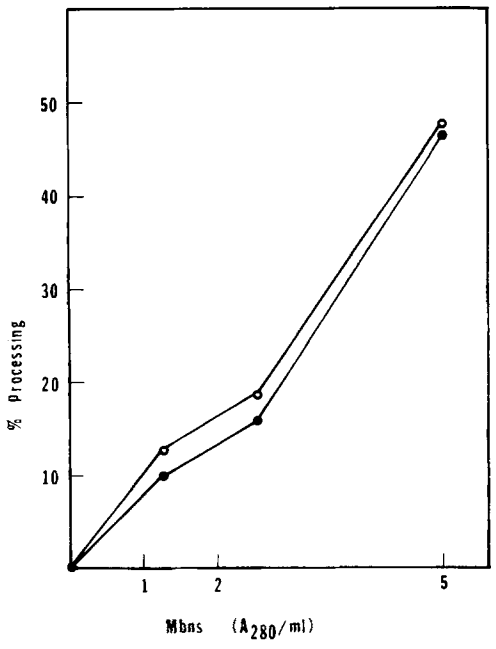

FIGURE 3 Co-translational conversion of nascent pre GSP to GS1 and P1 as a function of membrane concentration. Transcription-linked translations of pSPGP1 were performed in the rabbit reticulocyte lysate cell-free system (as described in Materials and Methods) in the presence of varying membrane concentrations (0, 1.25, 2.5, or 5 A_{280} U/ml). Aliquots (5 μl) were immunoprecipitated with either rabbit anti-human globin serum (●) or rabbit anti-ovine prolactin (○). After SDS PAGE, bands were viewed by autoradiography (using preflashed film) and quantitated by scanning densitometry as described in Materials and Methods.

(nascent) precursor and that the processing activity is associated with the microsomes.

Our interpretation of these results is that the prolactin signal sequence, now localized internally, is still functional as evidenced by accessibility of the signal sequence cleavage site to signal peptidase, a lumenally disposed enzyme of the endoplasmic reticulum (9). The products of this cleavage are authentic prolactin (P1) and globin with the prolactin signal sequence attached at its carboxy terminus (GS1).

Both GS1 and P1 are Localized inside Microsomal Vesicles

To determine which of these products are completely translocated across the membranes, we used two different experimental approaches. First, posttranslational proteolysis with proteinase K was used to localize the precursor and the two cleavage products. Any polypeptide that is completely translocated across the bilayer will be resistant to proteolysis unless the integrity of the membrane is abolished by the addition of nonionic detergents (7). If a protein spans the membrane, its molecular weight will be reduced by digestion of the cytoplasmically disposed domains (10, 11). A protein localized completely outside of the membrane vesicles will be totally digested by the added protease (3, 4). As can be seen in Fig. 4, the precursor, preGSP, is completely degraded (lanes *H–J*). However, some chains of GS1 (23%) and almost all chains of P1 (89%) were protected from proteolysis (bands were quantitated by scanning densitometry as described in Materials and Methods; data not shown). When nonionic detergent was included with the protease to disrupt the protecting lipid bilayer, all protection from proteolysis was abolished (lanes *I* and *J*). To rule out the possibility that the relatively poor protection of GS1 was due to a slightly increased intrinsic protease resistance of that molecule relative to preGSP, rather than to translocation across a protecting lipid bilayer, we investigated the kinetics of proteolysis. We varied both the time of protease digestion (from 15 to 90 min) and the protease concentration (from 0.1 to 0.4 mg/ml final concentration of proteinase K). At all time points and at all protease concentrations virtually none of the preGSP chains but the same percentage of GS1 chains were protected (<1% and ~20%, respectively, as determined by scanning densitometry; data not shown). Our interpretation of these data is that the now internal prolactin signal sequence is able to direct not only the subsequently synthesized prolactin domain across the microsomal membrane, but also the already completed globin domain, albeit with lower efficiency.

Engineering an N-Linked Glycosylation Site into the Globin-coding Region of pSPGP1

To further test our interpretation we engineered a glycosylation site into the globin-coding region of pSPGP1. Previously

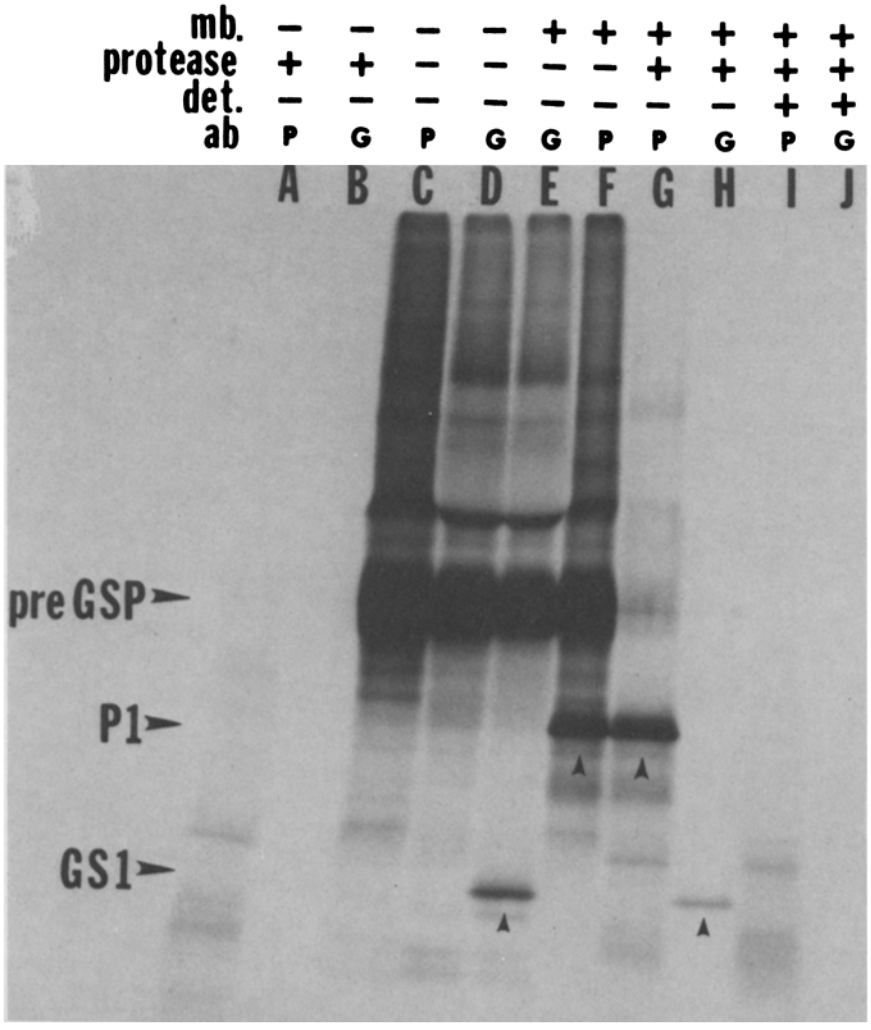

FIGURE 4 Transmembrane translocation of globin–prolactin fusion protein. Plasmid DNA from pSPGP1 was transcribed and translated in vitro. Some translation reaction mixtures included dog pancreas membranes at 2.5 A_{280} U/ml (+ *mb*, lanes *E–J*). After translation some aliquots (5 μl) were treated for 1 h at 0°C with proteinase K at 0.1 mg/ml (+ *protease*, lanes *A*, *B*, and *C–J*), in some cases in the presence of 1% Nikkol to disrupt the membranes (+ *det.*, lanes *I* and *J*). See Materials and Methods for details. Samples were all diluted and immunoprecipitated with either anti-globin (*G*, lanes *B*, *D*, *E*, *H*, and *J*) or anti-prolactin serum (*P*, lanes *A*, *C*, *F*, *G*, and *I*), separated by SDS PAGE and viewed by autoradiography.

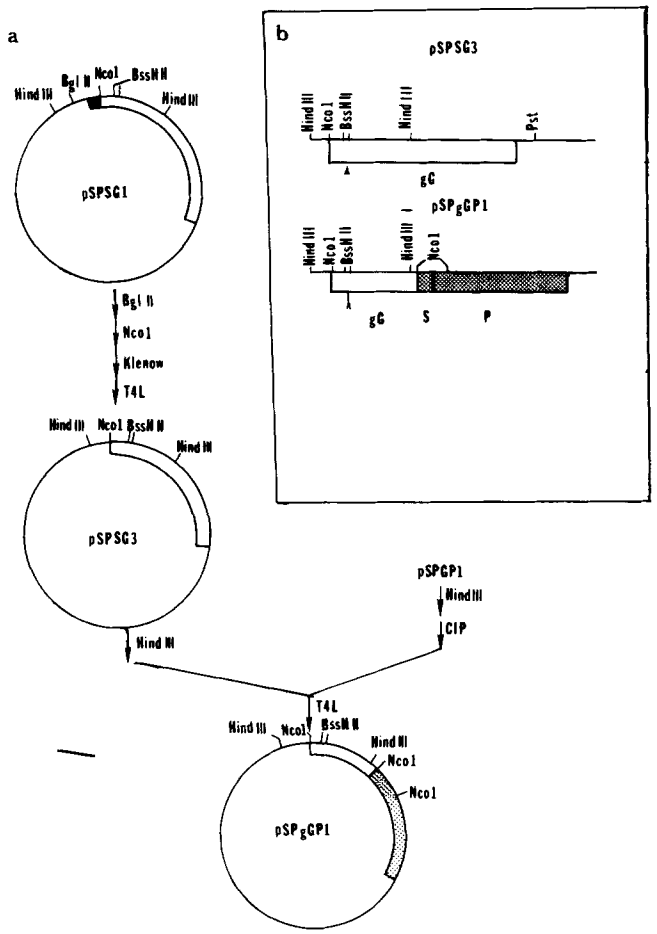

FIGURE 5 Construction of plasmid, pSPgGP1. (a) Construction scheme for pSPgGP1. The signal sequence coding region was deleted from pSPSG1 which encodes the β-lactamase signal sequence fused precisely to the amino terminus of chimpanzee globin into which a synthetic N-linked glycosylation site had been inserted at the BssHII site. The 430-bp HindIII fragment of the resulting plasmid, pSPSG3, was then inserted into the HindIII site of pSPGP1 to create pSPgGP1. Details are described in Materials and Methods. Stippled areas indicate prolactin-coding regions, black the lactamase signal sequence, and white indicates the globin coding regions. (b) Restriction maps and corresponding gene products of coding regions of pSPSG3 (upper drawing) and pSPgGP1 (lower drawing). The stippled bar represents prolactin, and white bars represent globin domains. The arrowhead pointing upward indicates the position of the Asn-X-Ser glycosylation site in the globin molecule encoded by pSPSG3; the branched symbol represents the carbohydrate that is added to the Asn of the Asn-X-Ser inserted into pSPgGP1 (but is not added to the translation product of pSPSG3; see Fig. 7).

we have shown that artificial glycosylation sites engineered into either the BssHII site (~20 amino acids from the amino terminus) or the BstE II site (~110 amino acids from the amino terminus) of globin result in translation-coupled core glycosylation of the globin domain when an amino terminal signal sequence is present to direct the nascent globin chains into the ER lumen.[2] Fig. 5 illustrates the scheme by which the coding region for one of these glycosylated globin constructs (pSPSG1) was first modified by deletion of the signal sequence coding region to generate pSPSG3. The HindIII fragment of this plasmid (lacking a signal sequence but containing an N-linked glycosylation site engineered into the BssH II site) was excised and ligated in place of the corresponding HindIII fragment of pSPGP1 (which lacked the glycosylation addition site). The resulting construction, pSPgGP1, differed from pSPGP1 only in the presence of the 24-bp oligonucleotide encoding a glycosylation addition site, and the encoded proteins are identical except for the insertion of eight amino acids, Ala-His-Asn-Gly-Ser-Gly-Ser-Gly, between amino acids 20 (Gly) and 21 (Ala) in the globin domain of the pSPgGP1 gene product. Since glycosylation is restricted to the lumen of the ER (12, 13), addition of N-linked sugars represents a definitive assay for translocation that is independent of other criteria such as protection from proteolysis or signal peptide cleavage. Thus, addition of carbohydrate to the globin domain encoded by pSPgGP1 upon transcription-linked translation in the presence of microsomal membranes would constitute direct and conclusive evidence for translocation of the globin domain of the globin–prolactin fusion protein.

Glycosylation and Protection of gGS1

Fig. 6 demonstrates that when pSPgGP1 is transcribed and translated as described previously for pSPGP1, a ~32-kD globin and prolactin immunoreactive protein was synthesized (lanes *D* and *I*, arrows pointing downward). The co-translational (but not posttranslational, see lane *C*) addition of microsomal membranes resulted in appearance not only of the 26-kD P1 band (lane *J*) and the 14-kD gGS1 band (lane *E*, small arrowhead pointing upward), analogous to GS1 in the presence of pSPGP1, but also in the appearance of a 16-kD globin but not prolactin immunoreactive band (lane *E*, large arrow pointing upward). This new 16-kD band, termed gGS1′, was believed to be the glycosylated derivative of gGS1. Consistent with this interpretation, gGS1′ was well protected from proteases (lane *F*), as was P1 (lane *K*), while gGS1 was relatively poorly protected (lane *F*, small arrow pointing upward). Protection of gGS1 approximated that of GS1 (20%), presumably representing those chains of gGS1 that were translocated but not glycosylated, an intermediate often observed in glycoprotein biosynthesis both in vivo and in vitro (Lingappa, V., unpublished observation). Also, as was seen with preGSP, the precursor, pre gGSP, was not protected.

Our interpretation of these data was confirmed by treatment with Endo H which shifted the position of gGS1′ on SDS PAGE to that of gGS1 (Fig. 7, lanes *D–F*), thereby demonstrating the presence of carbohydrate on gGS1′. Similar treatment of the protein encoded by pSPSG3, which contains the glycosylation site but lacks a signal sequence to allow its utilization, demonstrates neither shift up on SDS PAGE with co-translational membranes (Fig. 7, lane *B*) nor protection from proteolysis (data not shown) nor shift down with Endo H digestion (Fig. 7, lane *C*).

Carbonate Extraction of gGS1′

Having demonstrated unequivocally that both P1 and gGS1′ were faithfully translocated across microsomal membranes using the internal signal sequence of prolactin encoded in pSPgGP1, we proceeded to study the disposition of these molecules in the lumen. In particular, we wanted to know if gGS1′ was integrated into the bilayer on the cisternal side or whether it was soluble in the lumen, i.e., did gGS1 display properties of a secretory or of an integral membrane protein?

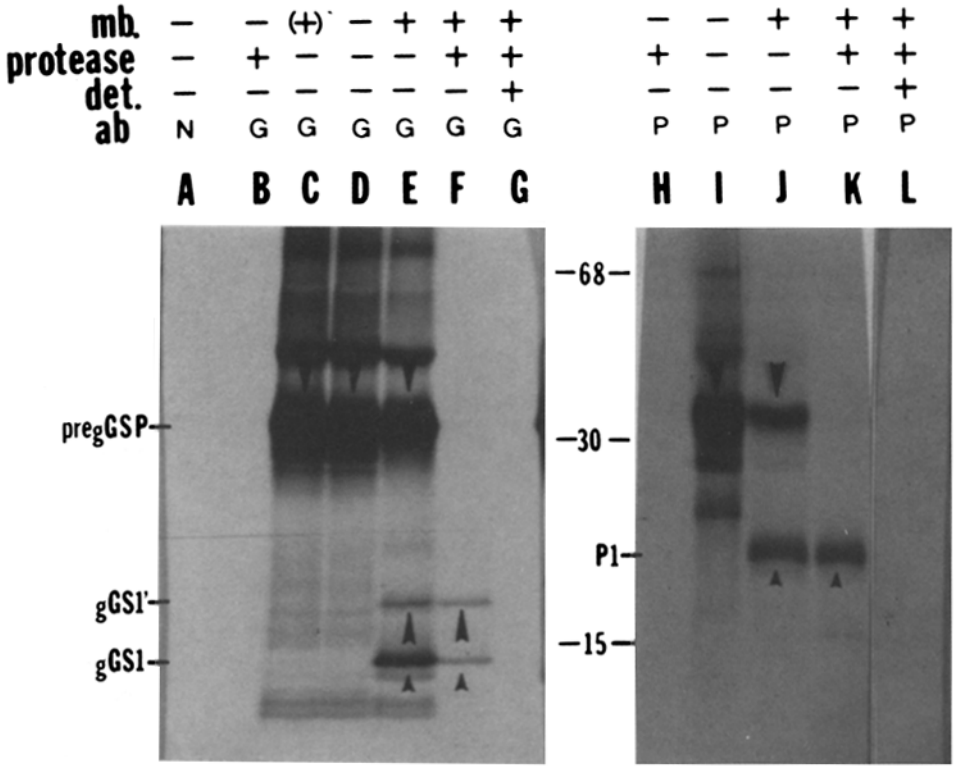

FIGURE 6 In vitro translation products encoded by pSPgGP1 and localization by proteolysis. Plasmid pSPgGP1 was transcribed and translated as described in Materials and Methods with membranes either absent (– *mb.*, lanes *A, B, D, H,* and *I*) or present at a concentration of 4 A_{280} U/ml (+ *mb.*, lanes *E–G* and *J–L*) during the translation reaction, or added posttranslationally to the same concentration with additional incubation (60 min at 24°C; (+) *mb.*, lane *C*). Some aliquots were then treated with proteinase K (+ *protease*) in the presence (+ *det.*, lanes *G* and *L*) or absence of detergent (– *det.*, lanes *B, F, H,* and *K*) as described and immunoprecipitated with globin antiserum (*G*, lanes *B–G*), prolactin antiserum (*P*, lanes *H–L*), or nonimmune rabbit serum (*N*, lane *A*). Samples were prepared and subjected to SDS PAGE and autoradiography.

Similarly we investigated whether pre gGSP and those gGS1 chains that are not protected from proteases (Fig. 6, lane *F*) are peripherally bound to membranes, integrated into membranes, or free in the cytosol. Fig. 8 shows the results of sedimentation of membranes in either isotonic sucrose buffer or after extraction with sodium carbonate pH 11.5, a procedure designed to strip off nonintegral proteins and content proteins from microsomal membranes (8). It can be seen that all forms (pre gGSP, gGS1, gGS1′, and P1) are extracted by carbonate (lanes *E–H*) and that both P1 and gGS1′ sedimented quantitatively with membranes in sucrose buffer (lanes *A–D*). Control extractions with a translation product known to integrate into membranes co-translationally demonstrated the fidelity of carbonate extraction in our hands (data not shown). Both gGS1 and pre gGSP were found in both the supernate and the pellet after membrane sedimentation in sucrose. The majority of gGS1 sedimented with microsomes, while pre gGSP was split approximately evenly between the membrane pellet and the supernate.

These data indicate that gGS1′ is not integrated into the vesicle membrane but exists either free in the microsomal lumen or peripherally associated with membrane proteins of the vesicle lumen. Also, those chains of gGS1 that are not protected from proteolysis are apparently in large measure bound to the membranes. The significance of this finding is currently under investigation.

DISCUSSION

Our original purpose in constructing pSPGP1 was to determine if amino terminal signal sequences were functionally related to internal signal sequences by converting a normally amino terminal signal into an internal one.

The existence of internal signal sequences has been proposed (1, 14) for integral membrane proteins as a means of accounting for multiple transmembrane loops, since alternating signal and stop transfer sequences in register could serve to stitch the nascent polypeptide into the bilayer in a programed fashion. Recent evidence demonstrates the existence of internal signal sequences directly in the case of vertebrate rhodopsin (Friedlander, M., and G. Blobel, personal communication; Lingappa, V. R., manuscript in preparation) and hepatitis B virus surface antigen (Eble, B., V. R. Lingappa, and D. Ganem, manuscript submitted for publication) by the ability of internal coding regions to translocate domains that would otherwise remain in the cytoplasm. The existence of internal signal sequences in secretory proteins is more controversial. Ovalbumin lacks a cleaved signal sequence (15), yet its nascent chain competes with nascent preprolactin for membrane receptors for translocation (16), suggesting the evidence of an uncleaved signal sequence somewhere in the ovalbumin chain. An internal tryptic fragment from ovalbumin displayed rough ER-specific competition for membrane receptors involved in prolactin translocation. Because this internal region, but not the amino terminus, displayed strong homology to amino terminal signal sequences of other proteins of the chicken oviduct, it was proposed that the uncleaved signal sequence of ovalbumin resided in this internal position rather than at the amino terminus and therefore, that amino terminal and internal signal sequences were structurally and functionally similar (14). Subsequent work, both indirect (17, 18) and direct (19), has demonstrated the existence of a signal sequence for ovalbumin that is closer to, but not necessarily at, the amino terminus, but has not ruled out the existence of a signal sequence at the more internal location.

In principle, internal and amino terminal signal sequences could represent two classes of ligands. For example, amino terminal signal sequences might be capable of directing the translocation only of polypeptide synthesized subsequently while internal signal sequences could translocate both preceeding and subsequently synthesized domains. Our demon-

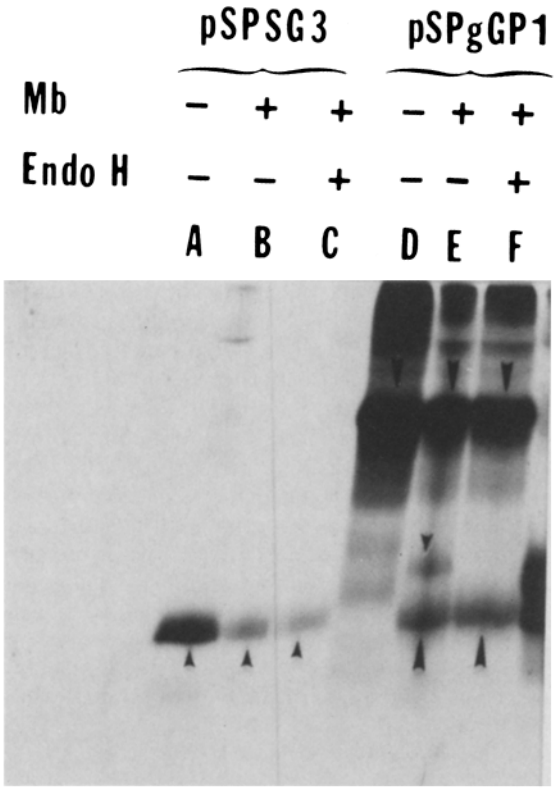

FIGURE 7 Endo H digestion of products encoded by pSPSG3 and pSPgGP1. Plasmid pSPSG3 (lanes *A–C*) and pSPgGP1 (lanes *D–F*) were transcribed and translated in separate reactions in the presence (+ *mb.*, lanes *B*, *C*, *E*, and *F*) or absence (− *mb.*, lanes *A* and *D*) of dog pancreas membranes (4 A_{280} U/ml). Samples were immunoprecipitated with globin antiserum and digested with endoglycosidase H (*Endo H* +, lanes *C* and *F*) or mock digested (*Endo H* −, lanes *B* and *E*) as described in Materials and Methods. The small arrow pointing downward in lane *E* indicates gGS1′, apparently the glycosylated form of gGS1 (large arrows pointing upward, lanes *E* and *F*). Large arrows pointing downward indicate pre gGSP1 (lanes *D–F*), and small arrows pointing upward indicate translation products of pSPSG3.

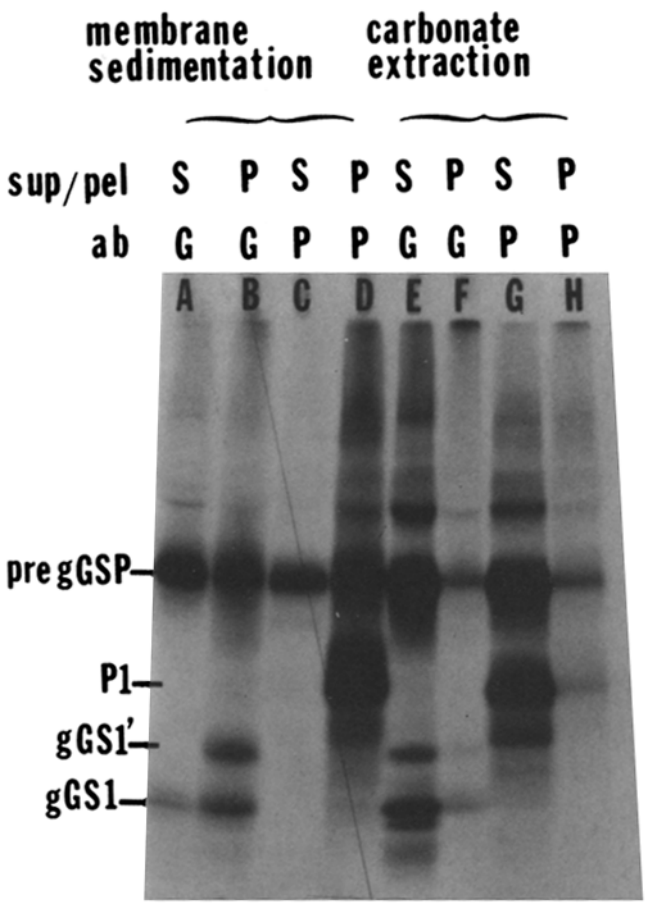

FIGURE 8 Carbonate extraction and vesicle sedimentation of translation products encoded by pSPgGP1. Plasmid pSPgGP1 was transcribed in vitro and products translated for 1 h at 24°C in the presence of 4 A_{280} U/ml of dog pancreas membranes. The translation reaction was then split into two equal aliquots, one of which was extracted with carbonate and the other sedimented in isotonic sucrose buffer (see Materials and Methods for details). Lanes *A–D* display supernatant (*S*) and pellet (*P*) after membrane sedimentation, and lanes *E–H* are the supernatant (*S*) and pellet (*P*) after carbonate extraction. Samples were immunoprecipitated with anti-globin (*G*, lanes *A*, *B*, *E*, and *F*) or anti-prolactin (*P*, lanes *C*, *D*, *G*, and *H*) serum, electrophoresed on a polyacrylamide gel, and viewed by autoradiography.

stration that the prolactin signal sequence functions when relocated internal to a globin domain, and generates soluble rather than integral membrane products in the lumen, suggests that internal signal sequences can exist in the case of secretory proteins. The observation that the internally relocated prolactin signal will translocate both amino and carboxy flanking domains (albeit with signficantly different efficiencies) indicates that internal signal sequences may be structurally and functionally similar to amino terminal signal sequences.

The finding that the two cleavage products are translocated with different efficiencies while using a common signal sequence is surprising. It suggests that GS1 and P1 domains of nascent preGSP are translocated independently. Perhaps this reflects a difference in translocation between domains that are nascent (such as P1 which is synthesized subsequent to the signal sequence) versus those that have already folded (such as the globin domain that is synthesized before the signal sequence). The location of the signal sequence relative to the translocated domain (i.e., at the carboxy terminus of globin versus at the amino terminus of prolactin) may also effect translocation efficiency of one versus the other domain, as might retention or removal of the signal sequence. Finally, we cannot rule out the possibility that the signal sequence retained by the globin domain can direct the reverse reaction of transport of GS1 from lumen to cytosol, to some extent, hence giving the appearance of inefficient translocation into the lumen. It should be possible to determine the kinetics of GS1 and P1 translocation as well as to evaluate the various explanations for the different efficiencies of translocation, using this cell-free expression system.

Translocation of the globin domain of 109 amino acids that is synthesized before the signal sequence is a remarkable observation with surprising implications for the mechanism of chain translocation. Two models for translocation of nascent polypeptides across the ER membrane have been put forth (1, 2). One of these postulates a proteinaceous transport tunnel whose assembly in the plane of the lipid bilayer is directed by the signal sequence (1). In this view, both targeting and translocation are the result of ligand–receptor-like inter-

actions between domains of nascent chains (topogenic sequences) and membrane proteins serving as receptors and transporters; the other model predicts that the free energy gain from insertion of the hydrophobic signal sequence directly into the lipid bilayer as a helical hairpin more than offsets the unfavorable energetics of pulling the contiguous hydrophilic chain into the bilayer, and thereby provides a thermodynamic basis for spontaneous insertion and translocation across the membrane, with the signal sequence retained in the bilayer after cleavage. This model requires no proteins in the membrane to facilitate translocation, and can be modified to account for signal recognition particle and its receptor as a requirement for targeting rather than translocation.

Since protein folding is likely begun during the window of time between initiation of protein synthesis and extrusion of the signal sequence from the large subunit of the ribosome (20, 21), it appears that a (partially) folded globin domain is presented to the membrane for translocation during the synthesis of nascent preGSP. Any hypothesis for translocation must account not only for this observation but also for the apparently soluble, nonintegrated form in which the gGS1′ chains are found. The topogenic sequence model can comfortably account for both of these considerations by postulating either an enzyme capable of denaturing the folded globin domain before transport as a component of the translocation machinery, or that the proteinaceous tunnel is of a dimension large enough to accomodate a partially folded domain. While the spontaneous insertion model might accomodate the idea of translocation of a folded globin domain providing it is denatured before the process of translocation, the observation that the gGS1′ chains are not integrated into the microsomal membrane argues that the signal sequence was never inserted directly into the bilayer.

Thus, regardless of whether a partially folded or completely denatured domain is translocated, our results suggest that not only targeting (i.e., signal recognition particle–receptor interaction [22–24]) but also translocation itself, involves proteinaceous machinery in the membrane.

Another implication of the globin domain being synthesized before the signal sequence, is that the energy expended in synthesis of a length of polypeptide cannot be the sole driving force for its translocation. We cannot at this time rule out a role for continued synthesis of the prolactin domain in translocation of the already completed globin domain, although we doubt this possibility. It seems more likely to us that the driving force for translocation resides in the membrane. If this hypothesis is correct, it should be possible to dissociate translocation of arrested nascent domains from continued protein synthesis, although the inability of completed and released chains to be translocated argues for some additional requirement such as for the ribosome or other protein cofactors (such as SRP) in the cytoplasm.

The placement of a globin domain at the amino terminus of a signal sequence provides the first marker for localization of a signal sequence after its cleavage. While we cannot rule out that such a bulky group has altered the fate of the signal, these data suggest that the signal is itself translocated, although it may be pulled back into the cytosol for degradation subsequently. In any case it appears not to be retained in the lipid bilayer, in view of our results with carbonate extraction.

We have also demonstrated a novel and direct approach to assaying for transfer of proteins across the ER membrane, that is the engineering of an N-linked glycosylation site (Asn-X-Ser/Thr) into the protein domain of interest. While failure to glycosylate does not rule out translocation, the addition of carbohydrate can be a powerful independent line of evidence in studying these events.

These results change our understanding of the vectorial discharge of proteins synthesized on membrane-bound ribosomes (25). Since in this case, translocation of the globin domain must take place after completion of its synthesis, it would appear that only a subset of the early events culminating in segregation in the ER cisternae need be vectorial.

Finally, the translocation of a globin domain when placed amino terminal to a signal sequence should not be taken to imply that any protein domain regardless of size or of secondary structure would be so translocated. The limits on translocation of domains both amino terminal as well as carboxy terminal to a signal sequence remain to be determined. There may well exist domains or conformations incompatible with translocation. At least two cases of cytoplasmic proteins unable to be translocated by amino terminal signal sequences have been proposed (26, 27). Further studies in cell-free systems should resolve these issues.

This work was supported by National Institutes of Health grant GM31626.

Received for publication 26 August 1985.

REFERENCES

1. Blobel, G. 1980. Intracellular protein topogenesis. *Proc. Natl. Acad. Sci. USA.* 77:1496–1500.
2. Engleman, D. M., and T. A. Steitz. 1981. The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis. *Cell.* 23:411–422.
3. Lingappa, V. R., J. Chaidez, C. S. Yost, and J. Hedgpeth. 1984. Determinants for protein localization: lactamase signal sequence directs globin across microsomal membranes. *Proc. Natl. Acad. Sci. USA.* 81:456–460.
4. Yost, C. S., J. Hedgepeth, and V. R. Lingappa. 1983. A stop transfer sequence confers predictable transmembrane orientation to a previously secretory protein in cell-free systems. *Cell.* 34:759–766.
5. Kreig, P. A., and D. A. Melton. 1983. Functional mRNAs are produced by SP6 in vitro transcription of cloned cDNA's. *Nucleic Acids Res.* 12:7057–7070.
6. Merrick, W. C. 1983. Translation of exogenous mRNAs in reticulocyte lysates. *Methods Enzymol.* 101:606.
7. Blobel, G., and B. Dobberstein. 1975. Transfer of proteins across membranes. II. *J. Cell Biol.* 67:852–862.
8. Fujiki, Y., A. L. Hubbard, S. Fowler, and P. B. Lazarow. 1982. Isolation of intracellular membranes by means of sodium carbonate treatment: application to endoplasmic reticulum. *J. Cell Biol.* 93:97–102.
9. Jackson, R. C., and G. Blobel. 1977. Post-translational cleavage of presecretory proteins with a detergent extract of rough microsomes. *Proc. Natl. Acad. Sci. USA.* 74:5598–5602.
10. Katz, F. N., J. E. Rothman, V. R. Lingappa, G. Blobel, and H. F. Lodish. 1977. Membrane assembly in vitro: synthesis, glycosylation and asymmetric insertion of a transmembrane protein. *Proc. Natl. Acad. Sci. USA.* 74:3278–3282.
11. Lingappa, V. R., F. N. Katz, H. F. Lodish, and G. Blobel. 1978. A signal sequence for the insertion of a transmembrane glycoprotein. *J. Biol. Chem.* 253:8667–8670.
12. Rothman, J., F. N. Katz, and H. F. Lodish. 1978. Glycosylation of a membrane protein is restricted to the growing polypeptide chain but is not necessary for insertion as a transmembrane protein. *Cell.* 15:1447–1454.
13. Lingappa, V. R., J. R. Lingappa, R. Prasad, K. E. Ebner, and G. Blobel. 1978. Coupled cell-free synthesis, segregation, and core glycosylation of a secretory protein. *Proc. Natl. Acad. Sci. USA.* 75:2338–2342.
14. Lingappa, V. R., J. R. Lingappa, and G. Blobel. 1979. Chicken ovalbumin contains an internal signal sequence. *Nature (Lond.).* 281:117–121.
15. Palmiter, R. D., J. Gagnon, and K. A. Walsh. 1978. Ovalbumin: a secreted protein without a transient hydrophobic leader sequence. *Proc. Natl. Acad. Sci. USA.* 75:94–98.
16. Lingappa, V. R., D. Shields, S. Woo, and G. Blobel. 1978. Nascent chicken ovalbumin contains the functional equivalent of a signal sequence. *J. Cell Biol.* 79:567–572.
17. Meek, R. L., K. A. Walsh, and R. D. Palmiter. 1982. The signal sequence of ovalbumin is located near the NH_2 terminus. *J. Biol. Chem.* 257:12245–12251.
18. Braell, W. A., and H. F. Lodish. 1982. Ovalbumin utilizes an NH2 terminal signal sequence. *J. Biol. Chem.* 257:4578–4582.
19. Tabe, L., P. Krieg, R. Strachan, D. Jackson, E. Wallis, and A. Colman. 1984. Segregation of mutant ovalbumins and ovalbumin-globin fusion proteins in Xenopus oocytes. *J. Mol. Biol.* 180:645–666.
20. Harrison, S. C., and R. Durbin. 1985. Is there a single pathway for the folding of a polypeptide chain? *Proc. Natl. Acad. Sci. USA.* 82:4028–4030.
21. Karplus, M., and D. L. Weaver. 1976. Protein-folding dynamics. *Nature (Lond.).* 260:404–406.
22. Walter, P., and G. Blobel. 1981. Translocation of proteins across the endoplasmic

reticulum. *J. Cell Biol.* 91:557–561.
23. Gilmore, R., P. Walter, and G. Blobel. 1982. Protein translocation across the endoplasmic reticulum: isolation and characterization of the SRP receptor. *J. Cell Biol.* 95:470–477.
24. Meyer, D. I., E. Krause, and B. Dobberstein. 1982. Secretory protein translocation across membranes: the role of the 'docking protein'. *Nature (Lond.).* 297:647–650.
25. Redman, C. M., and D. D. Sabatini. 1966. Vectorial discharge of peptides released by puromycin from attached ribosomes. *Proc. Natl. Acad. Sci. USA.* 56:608–615.
26. Moreno, F., A. V. Fowler, M. Hall, T. Silhavy, I. Zabin, and M. Schwartz. 1980. A signal sequence is not sufficient to lead β-galactosidase out of the cytoplasm. *Nature (Lond.).* 286:356–359.
27. Kadonaga, J. T., A. E. Gautier, D. R. Straus, A. D. Charles, M. D. Edge, and J. R. Knowles. 1984. The role of the β-lactamase signal sequence in the secretion of proteins by *Escherichia coli*. *J. Biol. Chem.* 259:2149–2154.
28. Sasavage, N. L., J. H. Nilson, S. Horowitz, and F. M. Rottman. 1982. Nucleotide sequence of bovine prolactin messenger RNA. *J. Biol. Chem.* 257:678–681.