

A Signal Sequence for the Insertion of a Transmembrane Glycoprotein

SIMILARITIES TO THE SIGNALS OF SECRETORY PROTEINS IN PRIMARY STRUCTURE AND FUNCTION*

(Received for publication, October 3, 1978, and in revised form, October 27, 1978)

Vishwanath R. Lingappa,[‡] Flora N. Katz,[§] Harvey F. Lodish,[§] and Günter Blobel[‡]

From the [‡]Laboratory of Cell Biology, The Rockefeller University, New York, New York 10021, and the [§]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

SUMMARY

The biosynthesis of a secretory protein and a transmembrane viral glycoprotein are compared by two different experimental approaches. (a) NH₂-terminal sequence analysis has been performed on various forms of the transmembrane glycoprotein of vesicular stomatitis virus synthesized in cell-free systems. The sequence data presented demonstrate that the nascent precursor of the glycoprotein contains a "signal sequence" of 16 amino acids at the NH₂ terminus, whose sequence is Met-Lys-Cys-Leu-Leu-Tyr-Leu-Ala-Phe-Leu-Phe-Ile-(His-Val-Asn)-Cys. This signal sequence is proteolytically cleaved during the process of insertion into microsomal membranes prior to chain completion. The new NH₂ terminus of the inserted, cleaved, and glycosylated membrane protein is located within the lumen of the microsomal vesicles and is identical to that of the authentic glycoprotein from virions. (b) Nascent chain competition experiments were performed between this glycoprotein, bovine pituitary prolactin (a secretory protein), and rabbit globin (a cytosolic protein). It was found that the nascent membrane glycoprotein, but not nascent globin, competed with nascent prolactin for membrane sites involved in the early biosynthetic event of transfer across membranes. These data suggest that an initially common pathway is involved in the biogenesis of secretory proteins and at least one class of integral membrane proteins.

Secretory proteins as well as certain integral membrane proteins are synthesized on ribosomes bound to membranes. The question of how these proteins segregate to their respective final locations in the cell—e.g. completely extracellular, plasma membrane, or various intracellular membranes—is of central importance in cell biology. One possibility is that these distinct classes of proteins utilize different receptors and are thus synthesized on different regions of the rough endoplasmic reticulum. A second possibility is that all are biosynthesized by a common route and segregation occurs later.

Previous studies of membrane assembly using glycoprotein of vesicular stomatitis virus as a model transmembrane protein have demonstrated that insertion and subsequent glyco-

* This study was supported by National Institutes of Health Grants CA-12413 (to G.B.) and AI-08814 (to H.L.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

sylation of nascent G proceeded only during translation and that the completed polypeptide was not a substrate for insertion (1, 2). Moreover, the glycoprotein synthesized in heterologous cell-free systems achieved an asymmetric orientation identical to that of pulse-labeled G found in rough microsomes of infected cells (1). This was consistent with the hypothesis (3-5) that insertion of membrane proteins into the bilayer proceeds by a mechanism, highly conserved through evolution, analogous to that described in the signal hypothesis for the transfer of secretory proteins across membranes (4, 5).

In this paper, NH₂-terminal sequence analysis of various forms of *in vitro*-synthesized G protein, as well as nascent chain competition experiments, are presented which allow us to extent the comparison of these two mechanisms. We demonstrate that this integral transmembrane glycoprotein, like most secretory proteins, is biosynthesized with a transient "signal sequence"; that this signal sequence is proteolytically removed during insertion into microsomal membranes; and that the resulting new NH₂-terminal residues exclusively within the lumen of the membrane vesicles. We also present evidence that nascent G and nascent bovine pituitary prolactin compete for membrane sites involved in chain segregation. These results are interpreted to provide further support for a conception of membrane protein biogenesis based on the signal hypothesis (5).

MATERIALS AND METHODS¹

RESULTS

Previously, translation of VS virus² mRNA in the wheat germ cell-free protein-synthesizing system in the absence of a microsomal membrane supplement has yielded as major products the virus-coded proteins N, NS, M, and G. Except for the G protein, all of the proteins synthesized *in vitro* co-migrated with their authentic viral counterparts (1, 11, 12). The identical pattern of four major polypeptides was observed upon translation of VS virus mRNA in the reticulocyte lysate cell-free system (Fig. 1, Lane 3). Moreover, in either cell-free system, translation in the absence or in the presence of microsomal membranes during translation resulted in the synthesis of two distinct intermediates of G (1). Pre-G₀ (Fig. 1, Lane 3, upward pointing arrow) synthesized in the absence of microsomal membranes is a 60,000-dalton, nonglycosylated, putative precursor of the 67,000-dalton authentic glycoprotein (G₂) isolated from virions. G₁ (Fig. 1, Lane 5, downward pointing arrow) synthesized in the presence of membranes has been previously characterized (1) to represent a partially glycosylated form of G which is inserted into the microsomal membrane with an asymmetric, transmembrane orientation identical to that of the pulse-labeled intermediate of G observed in rough microsomes of infected cells. When membranes were added after translation, the conversion of pre-G₀ to G₁ did not take place, i.e. the completed pre-G₀ was not a substrate for insertion nor, hence, for glycosylation, an exclusively luminal event.

In the case of secretory proteins, the information for translation-coupled segregation (7, 9, 13-15) usually resides in an

¹ Portions of this paper (including "Materials and Methods, References, and Figs. 2, 3, and 6) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full-size photocopies are available from the Journal of Biological Chemistry, 9650 Rockville Pike, Bethesda, Md. 20014. Request Document No. 78C-335, cite author(s), and include a check or money order for \$1.00 per set of photocopies.

² The abbreviation used is: VS virus (VSV in miniprint), vesicular stomatitis virus; G, glycoprotein.

From the data presented here and previously (1, 16), it seems clear that the information which initiates the process of translation-coupled insertion of the G protein resides in the NH₂-terminal signal sequence, and moreover, that proteolytic removal of this information takes place before chain completion. It seemed possible however that separate sets of membranous receptors were involved in segregation and insertion, the latter set allowing only partial transfer of proteins across membranes and, hence, resulting in a transmembrane deposition. To investigate this possibility, competition experiments were designed involving nascent chains of G, or bovine pituitary prolactin (a nonglycosylated secretory protein), or rabbit globin (a cytosolic protein lacking a signal sequence). In these experiments, the behavior of nascent prolactin in the presence of varying amounts of nascent G protein is evaluated. Except where indicated, the amount of membranes was kept constant and subsaturating with respect to segregation activity. Competition was monitored by assaying for processing of prolactin which is an accurate and readily quantifiable indicator of the extent of segregation (7).

It can be seen that the presence of low (Fig. 5, Lane 2) or high (Fig. 5, Lane 3) concentrations of VS virus mRNA resulted in decreased levels of processed prolactin relative to preprolactin, when compared to translation of prolactin mRNA in the absence of added VS virus mRNA (Fig. 5, Lane 1). Similar experiments with increasing amounts of globin mRNA had no such effect at either low (Fig. 5, Lane 4) or high (Fig. 5, Lane 5) globin mRNA concentration. That competition between nascent pre-G₀ and nascent preprolactin

was most likely for membrane sites and not for soluble factors was supported by the demonstration (Fig. 5, Lane 6) that increasing membrane concentrations abolished the competition (Fig. 5, Lane 3). In order to quantitate the extent of inhibition of segregation of nascent prolactin by nascent pre-G₀, the radioactivity in preprolactin and in processed prolactin was determined separately and segregation was expressed as a percentage of total radioactivity in completed prolactin chains (processed prolactin/preprolactin + processed prolactin). It can be seen in Fig. 6 that increasing the amount of VS virus mRNA changed the ratio of processed prolactin to preprolactin from approximately 3:1 to 1:1 (*i.e.* reduced the percentage of processed prolactin chains from 77% to 52%). In contrast, globin mRNA had no effect on the percentage of prolactin segregation (Fig. 6). In either case, increasing the concentration of the competing mRNA resulted in an increase in both total protein synthesis (data not shown) and in specific products coded for by the competing species (*arrowheads*, Fig. 5). For example, at 16 μg/ml of globin mRNA, approximately 182,000 cpm of globin were synthesized while at the same mRNA concentration 122,000 cpm of VS virus products were made of which G comprised less than 5%. Thus, even a vast excess of nascent globin did not affect the percentage processing of prolactin.

DISCUSSION

The demonstration (1) that cell-free systems are capable of membrane assembly and that asymmetric integration into the bilayer proceeds only during translation has permitted detailed studies into this fundamental and ubiquitous intracellular process. Kinetic studies (16) have indicated a requirement for the NH₂-terminal region in the insertion process. The structural studies presented in this communication demonstrate conclusively that nascent G protein of VS virus contains a hydrophobic sequence of 16 amino acids at the NH₂ terminus which is proteolytically removed during insertion prior to chain completion.

A comparison of the VS virus signal with sequence data on signals of secretory proteins reveals several common features including a typical region of leucine-rich hydrophobic amino acids in positions 4 through 12 and the presence at either end of the signal, *e.g.* at residues 2, 3, 13, 15, and 16, of hydrophilic or polar amino acids. Similarly, the presence of a tyrosine in position 6 is consistent with the presence of at least one free hydroxyl-containing side group in most signal sequences to date. Apart from these general features, few specific homologies with other signal sequences can be discerned, which is not surprising in view of the fact that secretory signals which have been demonstrated to recognize similar receptors (19) often have little apparent homology with each other (20). For this reason, certain distinctive features of the VS virus pre-G₀ signal may not be a significant indication of functional distinction. For example, the presence of Cys in positions 3 and 16 while unusual, is not unique, in that prealbumin has been demonstrated to contain Cys in positions 6 and 17 of its 19-residue-long signal peptide (22) and pre-MOPC 41 a and b light chains both have a cysteine, as does pre-G₀ at the penultimate site before cleavage (23).

In general, then, the primary structural studies suggest similarities between the structural prerequisites for the early events in protein secretion and those of membrane assembly, two processes which pose conceptually related biosynthetic problems. In both processes, correct topological biogenesis involves exclusively the nascent form of the polypeptide and, in both processes, the information for transfer usually resides in a cleavable hydrophobic NH₂-terminal sequence. However, that a signal sequence need be neither cleavable nor highly hydrophobic (20) in order to function in the transfer of pro-

	1	5	10	15	20	25	30	35
a) RNA sequence	M	K	C	L	L	Y	L	
b) pre-G ₀	M	K	C	L	L	Y	L	A
c) G ₁								
d) G ₁								

FIG. 4. Summary of sequence data on VS virus G. *a*, from partial RNA sequence (Ref. 21); *b*, VS virus pre-G₀ (see Fig. 2); *c*, VS virus G₁ (see Fig. 3a); *d*, VS virus G₁' (see Fig. 3b). Arrow indicates cleavage site for signal peptidase. Underlined residues indicate basis for alignment of pre-G₀ with G₁. M = Met, K = Lys, C = Cys, L = Leu, Y = Tyr, A = Ala, F = Phe, I = Ile, H = His, V = Val, N = Asn, P = Pro, X = unknown.

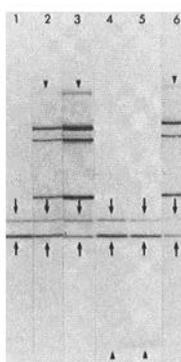


FIG. 5. Competition for membranous receptors by nascent VS virus G and nascent prolactin. 0.5 A₂₆₀ unit/ml of prolactin mRNA was translated in the reticulocyte lysate system (6) containing 5 A₂₆₀ units/ml (Lanes 1 to 5) or 10 A₂₆₀ units/ml (Lane 6) of EDTA-stripped nuclease-digested dog pancreas microsomal membranes in the presence of 0.1 (Lane 2) or 0.4 (Lanes 3 and 6) A₂₆₀ unit/ml of VS virus mRNA; or 0.1 (Lane 4) or 0.4 (Lane 5) A₂₆₀ unit/ml of globin mRNA. Downward pointing arrows refer to preprolactin and upward pointing arrows refer to processed prolactin. Downward pointing arrowheads refer to G₁, upward pointing arrowheads refer to globin. Lanes 1 to 5 were derived from one slab gel while Lane 6 was derived from a different slab gel and realigned according to the migration of processed prolactin.

teins across membranes in cell-free systems (19) has been demonstrated. It seems possible that hydrophobicity plays a role in the recognition by signal peptidase, a membrane-bound enzyme, and that recognition by receptors for transfer across the membrane is mediated by as yet poorly understood conformational features common to all signal sequences.

From our data, it appears that nascent G and nascent prolactin recognize at least one common membranous component. This component seems unlikely to be involved in processes of protein synthesis other than transfer across (into) membranes, since nascent globin, translated by the same pool of (initially) free ribosomes as G and prolactin, did not interfere with segregation of prolactin. Rather, it seems likely that competition was for a receptor on the cytoplasmic face of the membrane which recognizes at least in part, nascent, signal-containing chains and participates in their transfer across (into) the membrane. Although we cannot rule out competition for other components, *e.g.* signal peptidase on the luminal aspect of the membrane, in view of previous studies which show that nascent ovalbumin (glycosylated but not processed by signal peptidase) and nascent prolactin (processed by signal peptidase but not glycosylated) compete for common receptors, these explanations seem unlikely. Moreover, no new intermediates of G were seen (*e.g.* unprocessed but segregated and glycosylated precursors).

While detailed mechanistic studies await characterization of the membranous apparatus involved in segregation, the competition experiments presented here afford a preliminary statement on the mechanistic similarity of insertion and of segregation. These results support a biosynthetic model involving a common set of early events in membrane assembly and protein secretion, wherein similar signal sequences recognize a common set of receptors to initiate the process of

ribosome attachment, tunnel formation, and initiation of chain transfer across the membrane. Subsequent to the onset of chain transfer, the events diverge; secretory chains such as prolactin are entirely transferred within the vesicle lumen while the transfer of transmembrane proteins such as G are arrested prior to completion to transfer, resulting in a polypeptide spanning the bilayer. In both cases, co-translational modification, *e.g.* cleavage of signals and glycosylation of suitable acceptor sites, proceeds during translation and, probably, prior to cessation of transfer. In such a model, the information which arrests transfer—termed a “stop transfer” sequence (24)—would be expected to reside elsewhere in nascent G other than in the NH₂-terminal sequence, which is cleaved. Such a stop transfer sequence might function by abrogating or altering the ribosome membrane junction or self-associated tunnel components, and results in retention of the COOH-terminal polypeptide region as a cytoplasmic domain. Earlier studies have firmly documented vast evolutionary conservation of the role of signal sequences in the transfer of chains across membranes (9). By corollary, it seems likely that the stop transfer role is a conserved function of a set of defined polypeptide sequences, adjacent to, or itself constituting, the region of the inserted polypeptide which actually spans the bilayer.

One implication of this model of early biosynthetic events is that information other than that involved in the initial biosynthetic disposition (*i.e.* on free *versus* bound ribosomes, and hence, cytosolic *versus* extracellular consignment of polypeptides) is required to consummate the later stages of subcellular sorting.

Acknowledgment—We thank H. Desruisseaux for help with the Beckman Sequencer.

SUPPLEMENT
to
A Signal Sequence for the Insertion of a Transmembrane Glycoprotein:
Similarities to the Signals of Secretory Proteins in Primary Structure and Function.

by
Vishwanath R. Lingappa, Flora M. Katz, Harvey F. Lodish, and Günter Blobel

MATERIALS AND METHODS

Most of the procedures used in the present study have been detailed previously, including the preparation of mRNA from rabbit reticulocytes (4), bovine sterility glands (7), and 30S cells infected with VSV Indiana (8); the preparation of EDTA-stripped, nuclease-digested dog pancreas microsomal membranes (6,14); assays such as polyacrylamide slab gel electrophoresis in sodium dodecyl sulfate and subsequent autoradiography (4); and proteolysis of membranes with trypsin and chymotrypsin (4). The translation of mRNAs in a staphylococcal nuclease-treated rabbit reticulocyte lysate was as before (6,8) except that rabbit liver tRNA was added to a final concentration of 100 μ M. Other post-translational procedures such as immunoprecipitation are detailed in the figure legends. Simple preparation for sequence analysis: Following immunoprecipitation of G from total translation products of VSV mRNA as described in Figure 1, and subsequent polyacrylamide slab gel electrophoresis, gels were dried without fixing or staining and exposed to autoradiography. G protein forms were excised from gels and electroblotted in the presence of 1 mg ovalbumin as a carrier as previously described (9). To the eluted samples in a volume of 3 ml were added serially 0.2 ml of 5 mg/ml sperm whale apomyoglobin, 0.1 ml of 10 mg/ml ovalbumin, 0.4 ml of 4.0 M KCl and 3.5 ml of 50% trichloroacetic acid. After vortexing, incubation on ice for 15 minutes, centrifugation at 600 g for 10 minutes, and aspiration of supernatants, the resulting pellet was resuspended in 1 ml of water then precipitated with 9 volumes of acidified acetone at -20°C for 15 minutes. Resuspension, acetone precipitation, and centrifugation were repeated and the remaining pellet dissolved in 50% heptafluorobutyric acid prior to loading into the Beckman 890C sequencer running on a DMAA program. After each cycle of Edman degradation, thiazolidine derivatives were collected in ninhydrin, air-dried and counted by standard double-label procedures in 3 ml of toluene liquid fluor. Repetitive yields were calculated (10) normalized to the first cycle.

Tritiated amino acids and [³⁵S]cysteine were purchased from New England Nuclear, Boston, MA 02118, at the highest available specific activities. [³⁵S]methionine was purchased from Amersham Corp., Arlington Heights, IL 60005. Trypsin and chymotrypsin were from Boehringer, Mannheim, Germany; Protein A Sepharose was from Pharmacia Fine Chemicals, Piscataway, NJ 08854. Triethylamine was from FBA Pharmaceuticals, New York, NY; rabbit liver tRNA was from Grand Island Biological Co., Grand Island, NY 14077.

REFERENCES

- Katz, F.N., Rothman, J.E., Lingappa, V.R., Blobel, G., and Lodish, H.F. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 3278-3282.
- Tonegato, F., and Blobel, G. (1974) *Proc. Natl. Acad. Sci. U.S.A.* 71, 715-719.
- Rothman, J.E., and Leonard, J. (1977) *Science* 195, 743-753.
- Blobel, G., and Dobberstein, B. (1975) *J. Cell Biol.* 67, 853-851.
- Blobel, G. (1977) *3rd International Cell Biology* (P. D. Boylston, R. R., and Porter, K.R., ed.) pp. 318-325. The Rockefeller University Press, New York.
- Shields, D., and Blobel, G. (1978) *J. Biol. Chem.* 253, 3753-3756.
- Lingappa, V.R., Dewillier-Thierry, A., and Blobel, G. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 2432-2436.
- Pelham, H.R.B., and Jackson, R.J. (1976) *Eur. J. Biochem.* 67, 247-256.
- Shields, D., and Blobel, G. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 2059-2063.
- Waterfield, W.D., and Bridges, J. (1975) in *Instrumentation in Amino Acid Sequence Analysis* (F. H. R. N. ed) pp. 41-71. Academic Press, New York.
- Merrison, T.C., and Lodish, H.F. (1975) *J. Biol. Chem.* 250, 4955-4962.
- Both, G.W., Moyer, S.A., and Bamerjee, A.K. (1975) *Proc. Natl. Acad. Sci. U.S.A.* 72, 274-278.
- Sacchena, E., and Noize, I. (1976) *Proc. Natl. Acad. Sci. U.S.A.* 73, 1179-1183.
- Blobel, G., and Dobberstein, B. (1975) *J. Cell Biol.* 67, 852-862.
- Lingappa, V.R., Lingappa, J.R., Prasad, R., Ebner, K.E., and Blobel, G. (1978) *Proc. Natl. Acad. Sci. U.S.A.* 75, 2338-2342.
- Rothman, J.E., and Lodish, H.F. (1977) *Nature*, 269, 775-780.
- Tonegato, F., and Chosh, H.P. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 1516-1520.
- Hagen, D.A., Arnes, L., Yasumura, K.T., and Coon, M.J. (1977) *Biochem. Biophys. Res. Commun.* 72, 967-971.
- Lingappa, V.R., Shields, D., Woo, S.L.C., and Blobel, G. (1978) *J. Cell Biol.* In press.
- Palmiter, R.D., Gagnon, J., and Walsh, K.A. (1978) *Proc. Natl. Acad. Sci. U.S.A.* 75, 94-98.
- Ross, J.K. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 3677-3678.
- Palmiter, R.D., Thibodeau, S.W., Gagnon, J., and Walsh, K.A. (1977) in *Proceedings of the 11th FEBS Meeting*, Copenhagen, 1977, Regulatory Proteolytic Enzymes and their Inhibitors (Magnusson, S., Othman, M., Folmann, B., Dano, K., and Neurath, H., eds), 32, pp. 10-13. Pergamon Press, Oxford.
- Scheel, K., and Bursstein, Y. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 716-720.
- Blobel, G. (1978) in *Proceedings of the 11th FEBS Meeting*, Copenhagen, 1977, Gene Expression, (Clark, B.F.C., ed.) 43, pp. 99-108. Pergamon Press, Oxford.

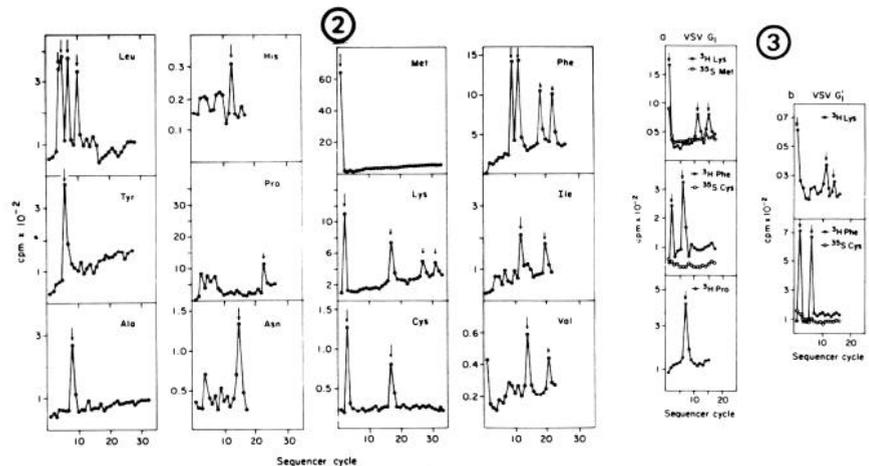


Fig. 2. Amino terminal sequence analysis of VSV preG₁. 1 ml translation reactions containing 250 μ Ci of a single tritiated amino acid, 250 μ Ci of [³⁵S] Met or Cys and 1.0 A₂₆₀ u/ml of VSV mRNA were carried out as previously described (6). Sample preparation was described in methods. The preG₁ was taken through 20-35 cycles of Edman degradation and residues analysed as described in methods. Sequence assignments in the signal region of preG₁ are indicated by long arrows; those in the authentic region of G₁ by short arrows. Input radioactivity was: [³⁵S] Met, 120,000 cpm; [³H] Lys, [³H] Pro, and [³H] Phe, 40,000 cpm; [³H] Leu, [³H] Ile, [³H] Ala, [³H] Tyr, 10-15,000 cpm; [³H] Val, [³H] Asn, [³⁵S] Cys, and [³H] His, 5,000 cpm. Repetitive yields of 96% were routinely obtained.

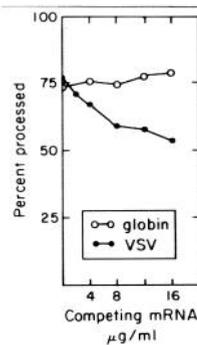


Fig. 3. Quantitation of competition for membranous receptors by nascent G of VSV and nascent prolactin. A constant amount of prolactin mRNA (0.5 A₂₆₀ u/ml) was translated in the presence of 5 A₂₆₀ u/ml of microsomal membranes in the absence or presence of increasing amounts of an additional mRNA (closed circles, VSV mRNA; open circles, globin mRNA). Translation products were analysed (see for example Fig. 5) and the radioactivity of gel slices containing preprolactin and prolactin was determined as previously (6). The extent of processing is expressed on the ordinate as a percentage of processed prolactin over that of total prolactin (which is preprolactin + processed prolactin) and is plotted as a function of competing mRNA concentration. As an example, 28,150 cpm were found in total prolactin and 21,400 cpm in processed prolactin when no competing mRNA was added.