



Cloning, sequencing and expression of cDNAs encoding two lysosomal membrane proteins, and generation of monoclonal antibodies against them
by Uthayakumar Selvanayagam

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in
Veterinary Science
Montana State University
© Copyright by Uthayakumar Selvanayagam (1993)

Abstract:

The lgps are heavily glycosylated lysosomal membrane proteins of two similar types, lgp-A and lgp-B, with masses of 110 to 120 kDa. Currently, the lgps are the best characterized lysosomal membrane proteins. cDNAs encoding the lgps have been cloned and sequenced from a number of species, and many polyclonal and monoclonal antibodies are available for studying them. CHO (Chinese hamster ovary) cells have been widely used as a system for studying the targeting and transport of proteins, and many stably-transfected CHO cell lines that express normal or mutant mouse lgp-A are now available. The hamster lgps themselves have not been characterized, however, so there is no way to follow the normal, endogenous lgps or their mRNAs in these transfected cell lines. In an effort to characterize the hamster lgps, we have cloned and sequenced their cDNAs, using mouse lgp-A and rat lgp-B cDNA probes to screen a CHO-K1 cDNA library in bacteriophage lambda gt11. The lgp-B cDNA obtained by this method lacked the 5' 190 bp of the mRNA, but we succeeded in cloning this segment by using the polymerase chain reaction (PCR) to amplify a cDNA preparation from CHO cells. The 2165 bp hamster lgp-A cDNA predicts a 407 amino acid polypeptide with 23 potential N-linked glycosylation sites, a putative signal peptide of 24 amino acids, a transmembrane domain of 25 amino acids and a cytosolic tail of 11 amino acids. The 1422 bp hamster lgp-B cDNA predicts a 410 amino acid polypeptide with 17 potential N-linked glycosylation sites, a putative signal peptide of 28 amino acids, a transmembrane domain of 26 amino acids and a cytosolic tail of 10 amino acids. The amino acid sequences of the hamster lgps are similar to those of other mammalian species, with the same regions of high evolutionary conservation. The hamster lgp-A and lgp-B cDNAs were cloned into eukaryotic expression vectors that confer neomycin resistance, and mouse NIH-3T3 cells were transfected with these plasmids. Monoclonal antibodies against hamster lgp-A and lgp-B were generated by immunizing mice with membrane glycoproteins from CHO cells; after a preliminary screening of the hybridoma products by immunofluorescence of fixed CHO cells, a specific screening for antibodies to hamster lgps was conducted using the transfected NIH-3T3 cells. One hybridoma was found to secrete lgp-A antibodies, and another was found to secrete lgp-B antibodies. The lgp-A antibodies do not recognize hamster lgp-B, and the lgp-B antibodies do not recognize hamster lgp-A; neither recognizes mouse lgps. These antibodies are now useful probes for endogenous lgps in transfected CHO cells expressing foreign lgps. Combined with the cDNA clones that will allow generation of specific nucleic acid probes to distinguish hamster and mouse mRNA levels in transfected cells, they will allow new insight to be gained into the regulation of expression, transport and targeting of lgps. Preliminary experiments to explore the role of the 5' untranslated segment of hamster lgp-A mRNA were also performed, and suggested a possible regulatory function.

The nucleotide sequences reported in this thesis have been submitted to GenBank™ with accession numbers L18986 (hamster lgp-A) and L19357 (hamster lgp-B).

CLONING, SEQUENCING AND EXPRESSION OF cDNAs ENCODING TWO
LYSOSOMAL MEMBRANE PROTEINS, AND GENERATION OF
MONOCLONAL ANTIBODIES AGAINST THEM

by

Uthayakumar Selvanayagam

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Veterinary Science

Montana State University
Bozeman, Montana

July 1993

71378
Se 498

APPROVAL

of a thesis submitted by

Uthayakumar Selvanayagam

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style and consistency and is ready for submission to the College of Graduate Studies.

JULY 27, 1993
Date

Bruce L. Granger
Chairperson, Graduate Committee

Approved for Major Department

AUGUST 2, 1993
Date

C. A. Spivey
Head, Major Department

Approved for College of Graduate Studies

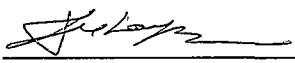
8/14/93
Date

R. L. Brown
Graduate Dean

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Montana State University, I agree that the Library shall make it available to borrowers under the rules of the Library.

If I have indicated my intention to copyright this thesis by indicating a copyright notice page, copying is allowable only for scholarly purpose, consistent with "fair use" as prescribed in the U. S. Copyright Law. Requests for permission for extended quotation from or reproduction of this thesis in whole or in parts may be granted by the copyright holder.

Signature 

date 8. 3. 93

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my sincere appreciation to Dr. Bruce L. Granger for his guidance and support through my studies. I also thank the other members of my thesis committee, Dr. Mitchell Abrahamsen, Dr. Donald Burgess and Dr. Clifford Bond, for their helpful discussions. In addition, I thank Dr. Abrahamsen for his valuable technical advice, and Dr. Burgess for having me as a rotation student in his lab for one quarter.

I am grateful to my father, M. J. Selvanayagam, and mother, U. Chinnathai, for all their love and support throughout my life. My wife, Jagatha, and daughter, Nila, have supported me and made my studies possible. I also thank my brother-in-law, Rangarajan, for helping me to come to the United States of America for my higher studies.

The valuable help rendered by Susan Wimer-Mackin and Sarah Warwood are greatly appreciated. I wish to thank all the people who have helped me in various ways, whose list of names would not fit on this page.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS.....	v
LIST OF TABLES.....	viii
LIST OF FIGURES.....	ix
ABSTRACT.....	x
INTRODUCTION.....	1
The lgp Family of Lysosomal Membrane Glycoproteins.....	1
Gene Structure of lgp-A and lgp-B.....	3
Intracellular Protein Transport.....	3
Transport of Lysosomal Proteins.....	5
Role of Carbohydrate in lgps.....	6
Functions of lgps.....	7
lgps as Tools for Studying Transport and Targeting.....	8
METHODS.....	10
Library Screening.....	10
Preparation of Probes.....	10
Library Screening.....	11
Plaque Purification.....	12
Preparation of Lambda DNA and Identification of Inserts.....	12
Purification of Inserts.....	13
Subcloning of cDNA Inserts into Plasmid Vectors.....	13
Ligation.....	13
Transformation of Bacteria.....	14
Preparation and Analysis of Plasmid DNA.....	14
Sequencing of cDNA Inserts.....	16
Sequencing Reaction.....	16
Electrophoresis and Autoradiography.....	16
Sequence Analysis.....	17
Sequencing Strategy.....	17
Preparation of RNA and DNA.....	18
Genomic DNA Extraction and Southern Blot Analysis.....	18
RNA Extraction.....	18
Single Stranded DNA Production for Sequencing.....	19
RACE (Rapid Amplification of cDNA Ends).....	19
Primer Design.....	19
cDNA Synthesis and Purification.....	20
Anchor Ligation and PCR Amplification.....	22

TABLE OF CONTENTS-CONTINUED

	page
Plasmid Constructs.....	22
Construction of pBSHA and pBGSHA.....	22
Construction of pBSHA5' Δ	24
Construction of pBSHB and pBGSHAB.....	26
Monoclonal Antibodies.....	26
Partial Purification of Antigen.....	26
Hybridomas and Monoclonal Antibodies.....	28
Expression in Heterologous Cells.....	29
Expression of Hamster lgp-A, lgpA5' Δ and lgp-B in Mouse Cells	29
Immunofluorescence.....	29
RESULTS.....	30
Library Screening.....	30
Southern Blot Analysis to Test Probes.....	30
Screening a Hamster cDNA Library for lgp-A.....	30
Screening a Hamster cDNA Library for lgp-B.....	31
Gel Purification and Cloning of cDNA Inserts.....	32
Sequencing.....	32
Sequencing Strategy of Hamster lgp-A and lgp-B.....	32
Obtaining the 5' End of lgp-B cDNA Using RACE.....	35
Hamster lgp-A Sequence.....	35
Hamster lgp-B Sequence.....	38
Monoclonal Antibodies.....	38
Expression of Hamster lgp-A and lgp-B cDNAs.....	41
DISCUSSION.....	44
Structure of lgps and Sequence Comparisons.....	44
Antibodies to lgps.....	51
Heterologous Expression of Hamster lgps.....	52
5' Sequences of lgp-B cDNAs.....	53
Future Directions.....	54
REFERENCES CITED.....	56
APPENDIX.....	62

LIST OF TABLES

Table	Page
1. Percentage of amino acid identity between hamster lgp-A and other known lgp-As.....	48
2. Percentage of amino acid identity between hamster lgp-B and other known lgp-Bs.....	48
3. Percentage of amino acid identity between lgp-A and lgp-B within the same species.....	48

LIST OF FIGURES

Figure	Page
1. Schematic diagram of 500 bp of the 5' end of the hamster lgp-B mRNA and the AmpliFINDER RACE method.....	21
2. Construction of pBGSHA.....	23
3. Construction of pBGSHA5'Δ.....	25
4. Construction of pBGSAHB.....	27
5. Sequencing strategy of lgp-A cDNA.....	33
6. Sequencing strategy of lgp-B cDNA.....	34
7. Nucleotide sequence of hamster lgp-A cDNA.....	36
8. The predicted protein sequence of hamster lgp-A.....	37
9. Nucleotide sequence of hamster lgp-B cDNA.....	39
10. The predicted protein sequence of hamster lgp-B.....	40
11. Immunofluorescence micrographs of endogenous and foreign hamster lgps.....	42
12. Aligned amino acid sequences of lgp-A.....	45
13. Aligned amino acid sequences of lgp-B.....	46
14. Aligned amino acid sequences of hamster lgp-A and lgp-B.....	47

ABSTRACT

The lgps are heavily glycosylated lysosomal membrane proteins of two similar types, lgp-A and lgp-B, with masses of 110 to 120 kDa. Currently, the lgps are the best characterized lysosomal membrane proteins. cDNAs encoding the lgps have been cloned and sequenced from a number of species, and many polyclonal and monoclonal antibodies are available for studying them. CHO (Chinese hamster ovary) cells have been widely used as a system for studying the targeting and transport of proteins, and many stably-transfected CHO cell lines that express normal or mutant mouse lgp-A are now available. The hamster lgps themselves have not been characterized, however, so there is no way to follow the normal, endogenous lgps or their mRNAs in these transfected cell lines. In an effort to characterize the hamster lgps, we have cloned and sequenced their cDNAs, using mouse lgp-A and rat lgp-B cDNA probes to screen a CHO-K1 cDNA library in bacteriophage lambda gt11. The lgp-B cDNA obtained by this method lacked the 5' 190 bp of the mRNA, but we succeeded in cloning this segment by using the polymerase chain reaction (PCR) to amplify a cDNA preparation from CHO cells. The 2165 bp hamster lgp-A cDNA predicts a 407 amino acid polypeptide with 23 potential N-linked glycosylation sites, a putative signal peptide of 24 amino acids, a transmembrane domain of 25 amino acids and a cytosolic tail of 11 amino acids. The 1422 bp hamster lgp-B cDNA predicts a 410 amino acid polypeptide with 17 potential N-linked glycosylation sites, a putative signal peptide of 28 amino acids, a transmembrane domain of 26 amino acids and a cytosolic tail of 10 amino acids. The amino acid sequences of the hamster lgps are similar to those of other mammalian species, with the same regions of high evolutionary conservation. The hamster lgp-A and lgp-B cDNAs were cloned into eukaryotic expression vectors that confer neomycin resistance, and mouse NIH-3T3 cells were transfected with these plasmids. Monoclonal antibodies against hamster lgp-A and lgp-B were generated by immunizing mice with membrane glycoproteins from CHO cells; after a preliminary screening of the hybridoma products by immunofluorescence of fixed CHO cells, a specific screening for antibodies to hamster lgps was conducted using the transfected NIH-3T3 cells. One hybridoma was found to secrete lgp-A antibodies, and another was found to secrete lgp-B antibodies. The lgp-A antibodies do not recognize hamster lgp-B, and the lgp-B antibodies do not recognize hamster lgp-A; neither recognizes mouse lgps. These antibodies are now useful probes for endogenous lgps in transfected CHO cells expressing foreign lgps. Combined with the cDNA clones that will allow generation of specific nucleic acid probes to distinguish hamster and mouse mRNA levels in transfected cells, they will allow new insight to be gained into the regulation of expression, transport and targeting of lgps. Preliminary experiments to explore the role of the 5' untranslated segment of hamster lgp-A mRNA were also performed, and suggested a possible regulatory function.

The nucleotide sequences reported in this thesis have been submitted to GenBank™ with accession numbers L18986 (hamster lgp-A) and L19357 (hamster lgp-B).

INTRODUCTION

Lysosomes may be defined as membrane-delimited terminal degradative compartments of eukaryotic cells. The lysosomal membrane performs several important functions: It provides a stable container for the multitude of acid hydrolases housed inside, it generates and maintains the acidic environment required for the hydrolases, and it also takes part in membrane fusion and fission events, and ion and catabolite transport.

The lgp Family of Lysosomal Membrane Glycoproteins

A family of lysosomal membrane glycoproteins (lgps) has been described in detail in birds and mammals. This family is composed of two highly related proteins, termed lgp-A and lgp-B (lysosomal membrane glycoprotein, types A and B). The cDNAs for these proteins have been cloned and sequenced from several species. The sequences known for lgp-A include rat lgp120 (Howe, *et al.*, 1988), rat LGP 107 (Himeno *et al.*, 1989), mouse LAMP-1 (Chen *et al.*, 1988), mouse lgp120 (Granger *et al.*, 1990), human lamp-1 (Fukuda *et al.*, 1988) and chicken LEP 100 (Fambrough *et al.* 1988), while lgp-B includes rat and mouse lgp110 (Granger *et al.*, 1990), rat LGP 96 (Noguchi *et al.*, 1989), mouse LAMP-2 (Cha *et al.*, 1990) and human lamp-2 (Fukuda *et al.*, 1988; Sawada *et al.*, 1993). Analysis of the deduced amino acid sequences reveals that the lgps are highly conserved type I membrane glycoproteins with 380-396 amino acids. The major portion of these molecules is luminal; the single transmembrane domain is composed of 25 or 26 amino acids, and the cytosolic tail has only 10 or 11 amino acids. In all species, the luminal domain of lgp-A is heavily glycosylated with N-linked glycans, while that of mouse and human lgp-B is glycosylated with N- and O-linked glycans (Fukuda *et al.*, 1988; Granger *et al.*, 1990). The human lgp-A, unlike its counterparts in other species,

also possesses O-linked glycans in addition to N-linked glycans. The major N-linked glycans attached to the lgps are tetraantennary structures (Howe *et al.*, 1988; Granger *et al.*, 1990). Amino acid sequence comparisons show that lgp-A from one species is more similar to lgp-A from other species than to lgp-B from the same species, which suggests that lgp-A and lgp-B diverged from each other in evolution prior to the divergence of mammals and birds. The most conserved features of the polypeptides are the eight cysteine residues, a stretch of about 30 amino acids around the 5th cysteine, the transmembrane domain and the cytosolic tail. It is presumed that the conserved features of these proteins are critical for their targeting and/or functioning (Granger *et al.*, 1990).

The central region of lgp-A is rich in proline and serine, while the central region of lgp-B is rich in proline and threonine. This region is variable in length among birds and mammals, and separates the two homologous domains of each polypeptide (Granger *et al.*, 1990). The two homologous domains are thought to have arisen by gene duplication: Each has four absolutely conserved cysteine residues, successive pairs of which form disulfide bonds (Carlsson and Fukuda, 1989; Arterburn *et al.*, 1990), and introns in the gene are in corresponding positions in the two domains.

Newly synthesized lgps appear as precursor proteins of 90-100 kDa, which upon treatment with Endo H (endo- β -N-acetylglucosaminidase H, which removes unprocessed N-linked carbohydrates) are reduced to 40-45 kDa core polypeptides (Lewis *et al.*, 1985; Green *et al.*, 1987; Granger *et al.*, 1990). Partial digestion of the precursor forms of lgps with Endo H revealed 16-18 N-linked glycans in mouse lgp-B, human lgp-A and lgp-B, and chicken lgp-A (Lewis *et al.*, 1985; Viitala *et al.*, 1988; Fambrough *et al.*, 1988; Granger *et al.*, 1990). The mature lgps are largely resistant to Endo H digestion; they are also extremely acidic, with an isoelectric point (pI) of 2-4 (Lewis *et al.*, 1985; Chen *et al.*, 1985; Lippincott-Schwartz and Fambrough, 1986; Granger *et al.*, 1990). The pI is raised

to near-neutrality by neuraminidase treatment, suggesting that the acidity is due to sialic acid residues (Lewis *et al.*, 1985; Granger *et al.*, 1990).

Gene Structure of lgp-A and lgp-B

The chicken lgp-A gene is 17 kbp long and has 9 exons (Zot *et al.*, 1990). The human lgp-B gene is 40 kbp long, has 9 exons and is located on the X chromosome at q24-25. The human lgp-A gene is located on chromosome 13 at q34, and 8 exons have been identified (intron one was not found) (Mattei *et al.*, 1990).

Each exon encodes almost identical portions of the polypeptides in all the genes examined (mouse lgp-B, chicken lgp-A and human lgp-A and lgp-B) (Granger *et al.*, 1990; Zot *et al.*, 1990; Sawada *et al.*, 1993). The similarity of lgp-A and lgp-B in genetic exon organization as well as in amino acid sequence further supports the idea that they were derived from a common ancestor (Granger *et al.*, 1990; Zot *et al.*, 1990; Sawada *et al.*, 1993).

Chloramphenicol acetyltransferase (CAT) reporter assays using the 5'-flanking region of the human lgp-B gene revealed that sequences from 20 to 170 nucleotides upstream from the transcription initiation site display maximal promoter activity (Sawada *et al.*, 1993); similar studies on the chicken lgp-A gene indicated that the lgp-A might be a constitutively synthesized protein (Zot *et al.*, 1990).

Intracellular Protein Transport

The eukaryotic cell is composed of several morphologically, biochemically and functionally distinct compartments. The multitude of proteins synthesized by the cell must reach their specific compartments to be able to perform their functions. Transport of

proteins is largely by diffusion in the case of cytosolic proteins, but is highly vectorial (*i.e.*, unidirectional) in the case of most secreted and membrane proteins. Eukaryotic membranes undergo nearly constant fusion and fission events that result in considerable membrane flux through each compartment, yet the molecular identities of the different compartments are preserved.

The sorting of membrane proteins is beginning to be understood. The signal(s) for the identification and sorting of these proteins must be contained within each protein, either as primary, secondary or tertiary structures or as post-translational modifications. In order for proteins to traverse the many discrete steps of the pathway successfully and reach their final destinations, each protein is required to possess many features. Lysosomal membrane proteins, for example, take the following path: They are cotranslationally translocated across the rough endoplasmic reticulum (ER) membrane (Verner and Schatz, 1988) and folded within the cisternae of the ER (Hurtley and Helenius, 1989). In the ER and Golgi, they are co- and post-translationally glycosylated (Kornfeld and Kornfeld, 1985). Sorting takes place in the *trans*-Golgi network (Griffiths and Simons, 1986), followed by delivery to lysosomes (von Figura and Hasilik, 1986; D'Souza and August, 1986; Green *et al.*, 1987; Kornfeld and Mellman, 1989). Between the different compartments, the proteins are carried within free carrier vesicles that pinch off from one compartment and fuse with the next (reviewed by Hopkins, 1992). Certain proteins are transported backwards or recycled, also in vesicles, especially if they are receptors of some sort that are to be preserved for future use. While some proteins are translocated across the ER membrane completely into the lumen of this organelle, others are inserted into the ER membrane and function as transmembrane proteins after being transported to their destinations. Resident proteins of each compartment must somehow be retained in those compartments.

Transport of Lysosomal Proteins

Significant advances have been made in understanding the biosynthesis and targeting of soluble lysosomal hydrolases. Post-translationally, these enzymes acquire a mannose-6-phosphate (M6P) marker that is recognized by the M6P receptors (von Figura *et al.*, 1986). From the *trans*-Golgi network (TGN), these enzymes are transported by the receptors to the pre-lysosomes and ultimately to the lysosomes (Kornfeld *et al.*, 1989). The M6P receptors are recycled from the pre-lysosomes to the TGN for further rounds of transport.

The transport and targeting of lysosomal *membrane* proteins, on the other hand, is not well understood. Transport of lgps to lysosomes seems to be independent of the M6P receptors and attached N-linked glycans (Barriocanal *et al.*, 1986; D'Souza and August, 1986; Lippincott-Schwartz and Fambrough, 1986; Granger *et al.*, 1990). The bulk of human lgp-A is transported directly to lysosomes, although a minor part of lgp-A is transported to the cell surface, internalized, and eventually delivered to lysosomes via the endocytic pathway (Carlsson and Fukuda, 1992). Harter and Mellman (1992) showed that, in CHO cells, surface expression is not a requirement for lysosomal transport of lgp-A, though a very small portion of the molecules do go to the plasma membrane. When lgps are over-expressed, their appearance on the surface increases (Granger *et al.*, 1989; Harter and Mellman, 1992). Canine lgp-B has been shown to pass through the basolateral plasma membrane of polarized cells before delivery to lysosomes (Nabi *et al.*, 1991). Chicken lgp-A has been shown to shuttle between the plasma membrane and endosomes before being delivered to lysosomes (Lippincott-Schwartz and Fambrough, 1986; Lippincott-Schwartz and Fambrough, 1987). Human lysosomal acid phosphatase (LAP), another lysosomal membrane protein, has been shown to be transported to lysosomes via

the plasma membrane (Waheed *et al.*, 1988; Braun *et al.*, 1989). cDNAs for a few other non-lgp lysosomal membrane proteins have been cloned and sequenced (see Discussion), but targeting and transport studies have not yet been done.

Efficient endocytosis of membrane proteins depends on the presence of a critical aromatic amino acid, usually a tyrosine, within the cytoplasmic tail (reviewed by Trowbridge, 1991). This tail tyrosine is essential for the delivery of human lgp-A to lysosomes in COS-1 cells (Williams and Fukuda, 1990). All known lgps have this tyrosine residue in their cytoplasmic tail. Other amino acids in the tail have also been shown to be important for targeting. Deletion, or substitution to a polar residue, of the isoleucine at the carboxy-terminus of mouse lgp-A results in missorting, and the lgps accumulate in the plasma membrane (Guarnieri *et al.*, 1993). Substitution to a hydrophobic amino acid does not affect targeting. Guarnieri *et al.* also demonstrated that the amino acids Tyr-Gln-Thr-Ile, in a specific context, were sufficient to target plasma membrane proteins to lysosomes. They also showed that the two amino acids at the carboxy terminus, Thr and Ile, were cleaved once the mouse lgp-A reached lysosomes, thus preventing the mouse lgp-A from entering the endocytic pathway again. Though Williams *et al.* also demonstrated that the cytoplasmic tail enabled a reporter molecule to be targeted to lysosomes, there is still some doubt that the tail is sufficient for lysosomal transport under all conditions (Granger *et al.*, unpublished).

Role of Carbohydrate in lgps

Malignant transformation has been associated with an increase in the amount of tetraantennary and triantennary N-linked glycans in rodent and human cell lines (Saitoh *et al.*, 1992). Human lgp-A and lgp-B exhibit an altered glycosylation pattern in human myelogenous leukemia cells. Saitoh *et al.* showed that highly metastatic cell lines

synthesize more N-linked oligosaccharides containing poly-N-acetyllactosamine than poorly metastatic cell lines. In addition, the amount of poly-N-acetyllactosamine in human lgps was found to be positively correlated with tumorigenicity. Conversely, glycosylation inhibitors such as swainsonine and castanospermine were reported to reduce tumor metastasis in animals (Saitoh *et al.*, 1992).

It has been demonstrated that lgps contain significant amounts of poly-N-acetyllactosamine, which in granulocytes and monocytes serves as ligands for adhesive molecules (Lowe *et al.*, 1990; Phillips *et al.*, 1990; Larsen *et al.*, 1990). Saitoh *et al.* also showed that highly metastatic sublines express more lgp molecules on their plasma membranes than their poorly metastatic counterparts. Taken together, the above evidence suggests that the lgps in the highly metastatic cell lines play a role in tumor metastasis, especially by influencing cell adhesion.

Functions of lgps

The presence of lgps in various tissues (kidney, liver, heart, gizzard and brain), in both embryonic and adult stages of life (Heffernan *et al.*, 1989; Zot *et al.* 1990) and their highly conserved nature suggest that they may be essential for cell viability. The high glycosylation and sialylation and the abundance of these proteins support the idea that the lgps might be providing a resistant barrier against the acid hydrolases inside lysosomes (Schauer, 1985; Lewis *et al.*, 1985; Granger *et al.*, 1990). Ultrastructural and cytochemical studies have shown that the interior of the lysosomal membrane is indeed lined with a carbohydrate-rich layer (Neiss, 1984). The correlation found between the amounts of lgps and lysosomal membranes in various cell types (more in macrophages, liver and I-cells, less in brain and muscle) further supports the idea of a protective function for the lgps (Ho *et al.*, 1983; Flotte *et al.*, 1983; Sandoval *et al.*, 1989).

The presence of both lgp-A and lgp-B in individual lysosomes, and their conserved amino acid sequences in birds and mammals, however, suggests the possibility of distinct functions for lgp-A and lgp-B. The observed amino acid sequence similarities may be critical for folding, targeting and stability similarities rather than for functional similarities. Alternatively, the presence of the two homologous, luminal domains, which might be involved in the binding of other molecules, raises the possibility that the lgps might themselves be acid hydrolases or receptors of some sort.

lgps as Tools for Studying Transport and Targeting

Regardless of their functions, the lgps are useful for studying membrane protein targeting and transport to lysosomes. The lgps are currently the best markers for lysosomal membranes and are also the best characterized lysosomal membrane proteins. Many polyclonal and monoclonal antibodies have been made against lgp-A and lgp-B of different species. These antibodies are useful for further characterizing and investigating lgps using a variety of techniques such as immunofluorescence, immunoelectron microscopy, ELISA, immunoblotting and immunoprecipitation.

Membrane protein targeting and transport can be studied by expressing mutant lgp cDNAs. Many stably-transfected CHO cell lines expressing normal and mutant mouse and rat lgp-A have already been generated (Granger *et al.*, 1989, 1992; Harter and Mellman, 1992) and used for studying transport and targeting.

CHO cells have been shown to be very hardy, fast growing and more easily transfected than many other cell types. In addition, CHO cells are among the best characterized lines for subcellular fractionation, in terms of optimized methods (Marsh *et al.*, 1987). Subcellular fractions can be examined to define the different compartments traversed by the mutant lgps, and the compartments that accumulate them. CHO

glycosylation mutants, including CHO-15B and CHO-lec2, have defects in complex carbohydrate maturation and can be used to study the role of sugars in lgp transport.

Monoclonal antibodies are available that bind to mouse and rat lgps but not hamster lgps (Hughes and August, *et al.*, 1982; Lewis *et al.*, 1985; Granger *et al.*, 1990; Harter and Mellman, 1992). In CHO cells expressing mouse or rat lgp, it is possible to specifically localize the foreign lgp using the above antibodies. Quantitation of the lgps is possible using immunoprecipitation, SDS-PAGE and autoradiography of cells metabolically labeled with ^{35}S .

It is important to find out whether a foreign lgp can interfere with the targeting and transport of the corresponding endogenous lgp. In cells overexpressing a foreign lgp, the targeting machinery might get saturated, such that excess lgp molecules cannot be sorted; they might thus spill over to the cell surface by a default pathway. An overexpressed mutant lgp might also inhibit the function of the normal lgps by combining with and inactivating them. Investigation of these possibilities requires the characterization of the hamster lgps themselves, generation of antibodies against them and sequencing the cDNAs encoding them. The sequences of the CHO lgps will also help to better define the conserved features of lgps. Here, we describe the cloning and sequencing of the hamster lgp-A and lgp-B cDNAs. In addition, we report the expression of hamster lgp-A and lgp-B cDNAs in mouse cells, and the production of monoclonal antibodies specific for hamster lgps.

METHODS

Library ScreeningPreparation of Probes

Three plasmids, MØλ1C6, MØλ1C4 and R2G-CR1, consisting of cDNAs inserted into the *EcoR* I site of Bluescript, were kindly provided by Dr. B. L. Granger. MØλ1C6 and MØλ1C4 have the same insert in opposite orientations. *Hind* III digestion of MØλ1C6 gives a 1.37 kb fragment of a mouse lgp-A cDNA with 86 bp of the 5' untranslated region (UT), all of the translated region, 57 bp of 3' UT, and 12 bp of vector polylinker. *Hinc* II digestion of R2G-CR1 gives a 1 kb fragment of a rat lgp-B cDNA corresponding to the protein segment from mid-signal peptide to midway between the 6th and 7th cysteine. *Hinc* II digestion of MØλ1C4 gives a 330 bp fragment at the 3' end of the mouse lgp-A cDNA. DNA fragments were gel purified and nick translated using [α - 32 P]dCTP or [α - 35 S]dATP and used as probes to screen a Chinese hamster ovary (CHO) cDNA library (see below).

Nick translation was performed by combining 100-200 ng of DNA with 4 μ l of a DNA polymerase I / DNase I mixture (0.4 U/ μ l DNA polymerase I, 40 pg/ μ l DNase I; Gibco BRL), 1.25 μ l of 10X nick translation buffer (see Appendix for concentrations), 150 μ M each of dATP, dTTP and dGTP (1.25 μ l 10X nucleotide mixture), and 4 μ l [α - 32 P] dCTP (10 nCi/ μ l at 3000 Ci/mmol) in a volume of 12 μ l. Reactions were allowed to proceed at 15°C for one hour, and the DNase activity was preferentially stopped by addition of 0.5 μ l of 20 mM EGTA (800 μ M, final; EGTA chelates Ca $^{++}$ in preference to Mg $^{++}$); incubation was then continued at 15°C for another 15 min. To repair any remaining nicks, 0.5 μ l of 1 mM NAD (38 μ M, final) and 2.5 units (0.5 μ l) of *E. coli* DNA ligase (New England Biolabs) were added and incubated at 15°C for an additional 15 min. The

reaction was stopped by addition of 2 μ l of 0.5 M EDTA (74 mM, final). 150 μ l of medium salt buffer was added, and the DNA was purified by using an Elutip-D^R (Schleicher & Schuell), following the manufacturer's protocol. To assess the efficiency of labeling, radioactivity in the purified probe was counted in a scintillation counter before being used in screening experiments. It was always found to have a specific activity of 2-7 x 10⁸ cpm/ μ g of DNA (more than the minimum recommended 10⁸ cpm/ μ g of DNA).

Library Screening

A CHO-K1 cDNA library was purchased from Clontech (cat. # JL1001b, lot # 7541; cloning vector: λ gt11; first strand synthesis primed with oligo dT and random primers). An overnight culture of *E. coli* Y1088 was grown in LB + 10 mM MgCl₂ + 0.2% maltose + 50 μ g/ml ampicillin; cells were pelleted and resuspended in 20 mM MgCl₂, stored at 4°C and used as host cells for λ gt11. Different dilutions of the phage library were used to obtain a range of plaque densities (3-300 x 10³ per 15 cm plate) by the top agar method (Sambrook *et al.*, 1989) on LB-agar containing 10 mM MgCl₂ and 20 μ g/ml ampicillin. Circular 132 mm HATF filters (Millipore) with a pore size of 0.45 μ m were used for plaque lifts. After the plates were cooled to 4°C, pre-marked filters were gently laid on the plates. After a few minutes, 3 or 4 needle stabs were made through the filter and agarose, and the filters were peeled off and placed successively (DNA side up) on single sheets of paper saturated with:

- 1) 1.5 M NaCl / 0.5 M NaOH (for 3m)
- 2) 1.5 M NaCl / 0.5 M Tris-Cl (pH 7.4) (twice for a total of 3-5 m)
- 3) 2 X SSPE (for around 5 m).

The filters were dried and baked in a vacuum oven at 80°C for 1-2 hours.

The filters were hydrated in 5X NET (see Appendix) at room temperature for a few minutes, and pre-hybridized at 60°C in hybridization solution for one hour with gentle

agitation. Blocking, hybridization and washing were performed in the same solution (except the hybridization solution also contained the radioactive probe). Hybridization was performed at 60°C with [α - 32 P]dCTP labeled cDNA probe ($1-3 \times 10^6$ cpm/ml) for 8-15 hours. The filters were washed at 60°C for 1 hour with 3-4 changes of solution, and exposed to X-ray film overnight.

The same filters were used to screen for both lgp-A and lgp-B. After hybridization and autoradiography with one of the probes, bound probe was removed by washing the filters five times in 90-95°C water for 2-3 minutes each. These filters were then hybridized with the other probe.

Plaque Purification

Positive plaques were isolated by removing appropriate plugs from the agar plates with the large end of a Pasteur pipette, and transferring to a 1.5 ml microcentrifuge tube containing 1 ml of SM buffer (see Appendix). Another round of screening was done with the eluted phage as just described, except that 10 cm plates and filters were used and the plugs were removed with the narrow end of the Pasteur pipette. Positive plaques were harvested as before, and the phage eluted in 0.5 ml of SM buffer. This was repeated a third time to obtain clonal phage.

Preparation of Lambda DNA and Identification of Inserts

400 μ l of the host cell (Y1088) suspension was mixed with 200 μ l of eluted phage suspension. This was incubated 15 min at 37°C, added to 50 ml of pre-warmed LB containing 50 μ g ampicillin and 20 mM MgCl_2 , and incubated at 37°C with shaking. When lysis was observed, 1 ml of chloroform was added and shaking continued for another 30 min to promote complete lysis. Debris was removed by centrifugation at 10,000 x g for 20 min at 4°C. Phage were precipitated in 10% polyethylene glycol (PEG

8000) and 1 M NaCl for 1-12 hours at 0°C, and collected at 9000 x g for 15 min at 4°C. The pellet was dissolved in SM (0.5 ml/culture). 200 µl of this suspension was mixed with 2 µl of DNase I (1 mg/ml) and 2 µl of RNase-A (10 mg/ml) to digest the host nucleic acid. Phage were disrupted in 0.5% SDS, 20 mM EDTA and Proteinase K (50 µg/ml) for 1 hour at 65°C. The DNA was purified by phenol chloroform extraction and ethanol precipitation, and dissolved in 200-400 µl of Tris/CDTA (see Appendix). Yield of DNA and insert size were determined by agarose gel electrophoresis after digesting a 5 µl aliquot of the DNA solution with *EcoR* I.

Purification of Inserts

cDNA inserts from *EcoR* I digested phage DNA were separated by horizontal agarose gel electrophoresis using low gelling temperature agarose (Agarose II from Amresco). The gels were stained with ethidium bromide and the inserts cut out, equilibrated in water, and phenol-chloroform extracted. Purity and yield of the isolated DNA were determined by agarose gel electrophoresis.

Subcloning of cDNA Inserts into Plasmid Vectors

Ligations

Bluescript SK+ or KS+ plasmids (Stratagene) were digested with *EcoR* I, and the ends dephosphorylated with calf intestine alkaline phosphatase (to prevent religation). The vector DNA was then purified by phenol-chloroform extraction. Insert and vector DNA were combined in a 3 : 1 molar ratio, and ligation was performed with one unit of T4 DNA ligase (Promega) in a volume of 10 µl by incubating at 15°C for 12 hours or at room temperature for 2 hours.

Transformation of Bacteria

E. coli strains NM522 and XL1-Blue were rendered competent by the SEM (simple and efficient method) method (Inoue *et al.*, 1990). Transformation of bacterial hosts with plasmid DNA was performed as follows (Sambrook, *et al.*, 1989): The DNA (in 5 μ l of ligation reaction) was added to 100 μ l of competent cells and incubated on ice for 30 min. The cells were heat shocked at 42°C for one min. The solution was spread on an LB plate containing 100 μ g/ml ampicillin, 0.5 mM IPTG and 40 μ g/ml X-Gal (5-bromo-4-chloro-3-indolyl- β -D-galactoside). Bacteria that produced white rather than blue colonies harbored plasmids with inserts. The blue/white color selection was ultimately discontinued because the dephosphorylation of vector ends almost eliminated religation, so nearly all colonies had plasmids with inserts.

Preparation and Analysis of Plasmid DNA

Plasmid purification followed the alkaline lysis procedure of Birnboim (1983), as modified by B. L. Granger. Transformed NM522 or XL1-Blue colonies were picked and grown in LB with 100 μ g/ml ampicillin overnight at 37°C with shaking. A 1.5 ml microcentrifuge tube was filled with culture, and cells were collected by centrifugation at 8000 x g for one min. Cells were resuspended in 150 μ l of cold Tris/CDTA/dextrose (see Appendix for concentrations) and left on ice for 10 min. Lysis of the cells was performed by adding 300 μ l of NaOH/SDS and immediately capping and mixing gently by inversion, and leaving on ice for 5 min. Genomic DNA was precipitated with 225 μ l of KOAc/formic acid and 5 min incubation on ice. The precipitate was removed by centrifugation and the DNA remaining in the supernatant was ethanol precipitated. The plasmid DNA was dissolved in 100 μ l of Tris/CDTA and centrifuged to remove any debris. To this was added 400 μ l of NaOAc/MOPS + DNase-free RNase-A at 100 μ g/ml, and incubated at

room temperature for 30 min. The plasmid DNA was purified by phenol-chloroform extraction and ethanol precipitation. The DNA pellet was dissolved in 50 µl of Tris/CDTA.

For larger plasmid preparations, 50 ml cultures were used and the above procedure was modified. In addition to adjusting the volumes, the following steps were also performed before the RNase digestion step: The precipitated genomic DNA was removed by centrifugation and filtration through glass wool. The DNA (and RNA) was precipitated with isopropanol and dissolved in 500 µl of Tris/CDTA. The bulk of the RNA was then removed using 500 µl LiCl/MOPS. Plasmid DNA was precipitated with ethanol and dissolved in 250 µl of NaOAc/MOPS/CDTA, and RNase was added to digest the remaining RNA. Subsequently, a plasmid purification kit (MagicTM Minipreps; Promega) was used for small-scale purifications.

Plasmids were digested with *EcoR* I, and inserts verified for presence and size. Whenever two inserts were obtained by *EcoR* I digestion of a lambda clone, the two resulting plasmids were differentiated by suffixing an 'L' or 'S' (to denote the long or short fragment). If part of a insert DNA was removed by restriction enzyme digestion and the remaining DNA religated, the resultant plasmid was given the name of the parent plasmid with 'd' followed by the name of enzyme used (example- pHB10 d*Spe* I was derived from pHB10 after *Spe* I digestion and religation of the ends). Plasmids with the insert in the opposite orientation have the letter R suffixed (to denote reverse). Plasmid pHB9 2H has the part of HB9 insert 3' to the *Pst* I site.

Sequencing of cDNA Inserts

Sequencing Reaction

Sequencing of recombinant plasmids was carried out by the Sanger method of dideoxy nucleotide mediated chain termination (Sanger *et al.*, 1977). Sequenase® 2.0 modified T7 DNA polymerase (United States Biochemicals) was used for sequencing. 1-4 µg of plasmid DNA in a volume of 5-10 µl was denatured with 1-2 µl of 2 M NaOH at room temperature for 5 min. The solution was neutralized with 5-10 µl of 5 M ammonium acetate, and the denatured plasmid DNA was precipitated with ethanol. Sequencing reactions were performed using 0.3-0.5 µl [α -³⁵S] dATP (10 mCi/ml) per reaction. To obtain sequence information close to the primer, a manganese-containing buffer in the Sequenase kit was used. Reaction products were stored at -20°C until electrophoretic analysis.

Electrophoresis and Autoradiography

A sequencing gel apparatus from BRL (Model S2) was used for sequencing. Standard 6% polyacrylamide electrophoresis gels were poured using the Sequagel™ sequencing system solutions from National Diagnostics, following the manufacturer's procedure. Gels were pre-run until the gel temperature reached 55°C. Sequencing reaction samples (1.5 µl) were loaded (8 lanes/template), and about 2.5 hours later a second loading of each reaction was sometimes performed to maximize sequence data. Gels were generally run for 3 hours for one loading with an additional 5-6 hours for a second loading. After electrophoresis, the gels were fixed in 15% methanol / 1% acetic acid for one hour, then dried onto a thick filter paper at 80°C for 2 hours under vacuum. Dried gels

were exposed to X-ray film (Kodak X-OMATTM AR; 35 cm x 43 cm) for 12-48 hours at room temperature.

Sequence Analysis

Sequence data were analyzed with the DNASTAR software package (DNASTAR, Inc.) on a Macintosh LC II personal computer. The hamster lgp nucleotide sequences were compared with the already known lgp sequences from other species using the Needleman-Wunsch algorithm (Devereux *et al.*, 1984). The data obtained were compiled, open reading frames established, and the translated protein compared to other known lgps by the Lipman-Pearson method (Pearson and Lipman, 1988) using DNASTAR. The nucleotide sequences reported in this thesis have been submitted to GenBankTM with accession numbers L18986 (hamster lgp-A) and L19357 (hamster lgp-B).

Sequencing Strategy

Inserts from the purified lambda DNAs were subcloned into plasmids (Bluescript) and sequenced from both ends using the standard plasmid primers (T3, T7, SK and KS). The relative position of each insert with respect to the hamster mRNA was inferred by comparing the sequence data to other known lgp sequences, using the Align program in DNASTAR. Using this information, appropriate clones were chosen for further sequencing. To obtain internal sequences, new plasmids were constructed by deleting selected restriction fragments. Primers already available in the lab (made for site directed mutagenesis of mouse lgp-A, or for PCR) were also found to be useful for sequencing internal stretches of the hamster lgp-A cDNA. In addition, two internal primers were synthesized by Patrice Mascolo (Veterinary Molecular Biology, Montana State University), using an Applied Biosystems 381A DNA synthesizer.

Preparation of DNA and RNA

Genomic DNA Extraction and Southern Blot Analysis

Genomic DNA was extracted from bovine liver, mouse kidney and CHO cells using SDS and proteinase K to liberate the DNA, followed by purification by phenol-chloroform extraction. The DNAs were digested to completion with restriction enzymes (3.32 µg of mouse DNA with *EcoR* I; 3, 1.5 and 0.6 µg of CHO DNA with *EcoR* I; and 4 µg of bovine DNA with *EcoR* I and *Hind* III) and electrophoretically resolved in a 0.8% agarose gel. Unlabeled DNA probes (see below) were included in the gel as positive controls. Following denaturation for 45 min in 0.5 N NaOH and depurination in 0.2 N HCl, DNAs were neutralized for 30 min in 1 M Tris (pH 7.4), 1.5 M NaCl. DNA was subsequently transferred to a BAS-NC (Schleicher & Schuell) filter by capillary transfer (Sambrook *et al.*, 1989) for 20-22 hours. The DNA was fixed to the filter by baking at 80°C for one hour. The filter was then prehybridized for one hour at 65°C in the hybridization solution. Hybridization was carried out at 65°C for 8-12 hours in fresh hybridization solution containing [α^{32} P]dCTP-radiolabeled mouse lgp-A or rat lgp-B probes (as done for library screening). Following hybridization, membranes were washed for one hour in hybridization solution (without probe), changing the solution 4 or 5 times. The washed membranes were subsequently autoradiographed for 46 hours at -80°C using X-ray film (Kodak X-OMATTM AR; 35 cm x 43 cm).

RNA extraction

Total RNA was extracted from CHO cells by the guanidinium thiocyanate-phenol-chloroform method (Chomczynski and Sacchi, 1987) and stored in 100% ethanol at -20°C before use.

Single Stranded DNA Production for Sequencing

Bacterial colonies from M9 minimal plates were inoculated into 2XTY medium with ampicillin (100 µg/ml), and mini-cultures grown. These cultures were used to inoculate 50 ml 2XTY medium and incubated at 37°C with shaking. After one hour, M13K07 helper phage were added and incubation continued with shaking for 30 min. Kanamycin was added to a final concentration of 70 µg/ml and incubation continued with shaking for 9 hours. Cultures were centrifuged, and to 40 ml of the supernatant was added 10 ml of 20% PEG 8000/3.75 M ammonium acetate, and the mixture was incubated on ice for 3 to 12 hours. Phage particles were pelleted and phenol-chloroform extracted. Phagemid DNA was precipitated with ethanol, dissolved in Tris/CDTA and used for sequencing.

RACE (Rapid Amplification of cDNA Ends)

A 5'-AmpliFINDER RACE kit (Clontech # K1800-1) was used to clone the 5' end of the hamster lgp-B cDNA, as follows (Fig. 1):

Primer design

Two nested primers, P1 (5' CACTTCCTTTAGGGATAGCACCAGG 3') and P2 (5' TGTTGTACACGAGCCAGATGCTGCC 3'), were designed using the available sequence information of the hamster lgp-B cDNA. The primers were designed to anneal ~500 bp downstream from the predicted 5'-end the mRNA, and 230 bp downstream from the 5'-end of the already-determined sequence. The known sequence had upstream *Bam*H I and *Pvu* II sites which were used for mapping and cloning of the final PCR product.

cDNA Synthesis and Purification

cDNA synthesis was primed by preincubating 100 µg of total RNA from CHO cells with 1 µg of cDNA synthesis primer (P1) at 65°C for 5 min in a total volume of 10 µl. To this, 20 µl of reverse transcription master mix was added and incubated at 52°C for 30 min.

master mix

9.2 µl DEPC treated water

9 µl 4X reverse transcriptase buffer

1.6 µl RNase inhibitor (40 units/µl)

3.7 µl ultrapure dNTP mix (10 mM each)

0.5 µl AMV reverse transcriptase (25 units/µl)

2 µl of 6 N NaOH was added and the mixture incubated at 65°C for 30 min to hydrolyze the RNA. Then 2 µl of 6 N acetic acid was added to neutralize the base.

80 µl of 6 M NaI was then added and the cDNA was bound to 8 µl GENO BIND™ (proprietary matrix), centrifuged, and the pellet washed twice by resuspending in 500 µl 80% ethanol and spinning down again. The air-dried pellet was resuspended in 50 µl DEPC-treated water, incubated at 65°C for 5 min to elute the cDNA, centrifuged, and the supernatant containing the cDNA transferred to a new tube.

The cDNA in the above preparation was precipitated by adding 5 µl of 2 M sodium acetate and 100 µl of 95% ethanol in the presence of 15 µg of glycogen as carrier, and incubated at -20°C for 30 min. The cDNA was pelleted, rinsed with 80% ethanol, air dried, and the pellet resuspended in 6 µl DEPC treated water.

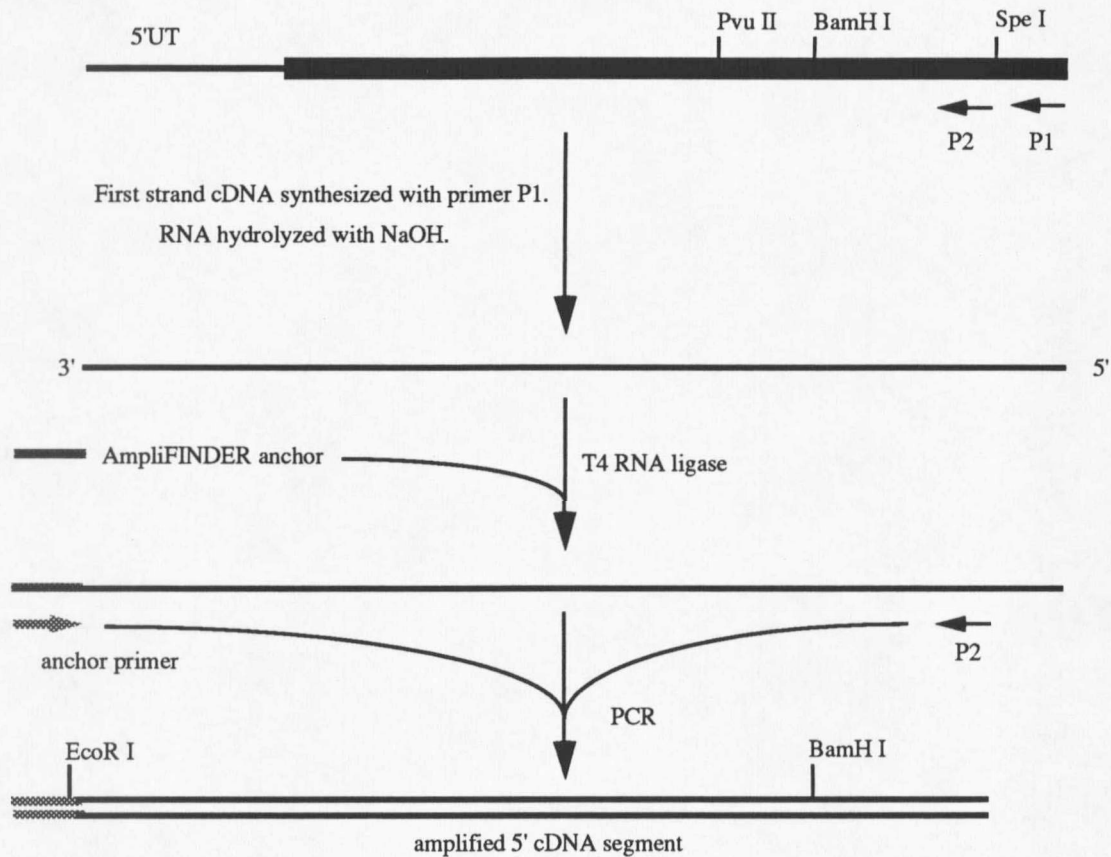


Figure 1. Schematic diagram of 500 bp of the 5' end of the hamster Igp-B mRNA and the AmpliFINDER RACE method. The location of the primers, P1 and P2, are indicated. Using primer P1, the first strand cDNA was synthesized. The AmpliFINDER anchor was ligated to the 3' end of the cDNA. PCR was performed using primers P2 and the AmpliFINDER anchor primer. The *Bam*H I and *Pvu* II sites were used for mapping. The *Eco*R I and *Bam*H I sites were used to clone the PCR-amplified product into a plasmid.

Anchor Ligation and PCR Amplification

The anchor (see below) was ligated to the cDNA by mixing 2.5 µl of the above cDNA, 2 µl of AmpliFINDER Anchor (4 pmol), 5 µl 2X single-stranded ligation buffer, and 0.5 µl T4 RNA ligase (20 units/µl) and incubating at 20°C for 18-20 hours.

The 5' end of the cDNA was PCR amplified by using primer P2, which anneals to the cDNA, and the anchor primer (anchor and anchor primer, which are complementary oligonucleotides, are part of the kit; the 5' end of the anchor has a phosphate group necessary for ligation, and the 3' end is blocked with an amine group to prevent self-ligation; in addition, the anchor has an *EcoR* I sites to facilitate cloning), which anneals to the anchor ligated to the 5' end of the cDNA. The PCR was performed with 1 µl of 10 µM primer each, 2 units of Tfi DNA polymerase (Epicentre) and 1 µl of anchor-ligated cDNA. A "hot start" was performed by adding the primers after the tubes were heated to 82°C for one min. The PCR cycles were as follows:

94°C	Denature	45 sec
60°C	Anneal	45 sec
72°C	Extend	2 min

35-100 cycles were performed with a final extension time of 7 min.

A Robocycler™ 40 (Stratagene) was used for thermal cycling.

Plasmid Constructions

Construction of pBSHA and pBGSHA

None of the hamster lgp-A clones obtained represented the full length mRNA. Two of the clones (pA1L2 and pA17B; Fig. 2), however, overlapped and together covered 135 bp of 5' UT and the whole of the coding and 3' UT regions. A unique *BstX* I site in their region of overlap allowed them to be readily spliced together. The above two plasmids

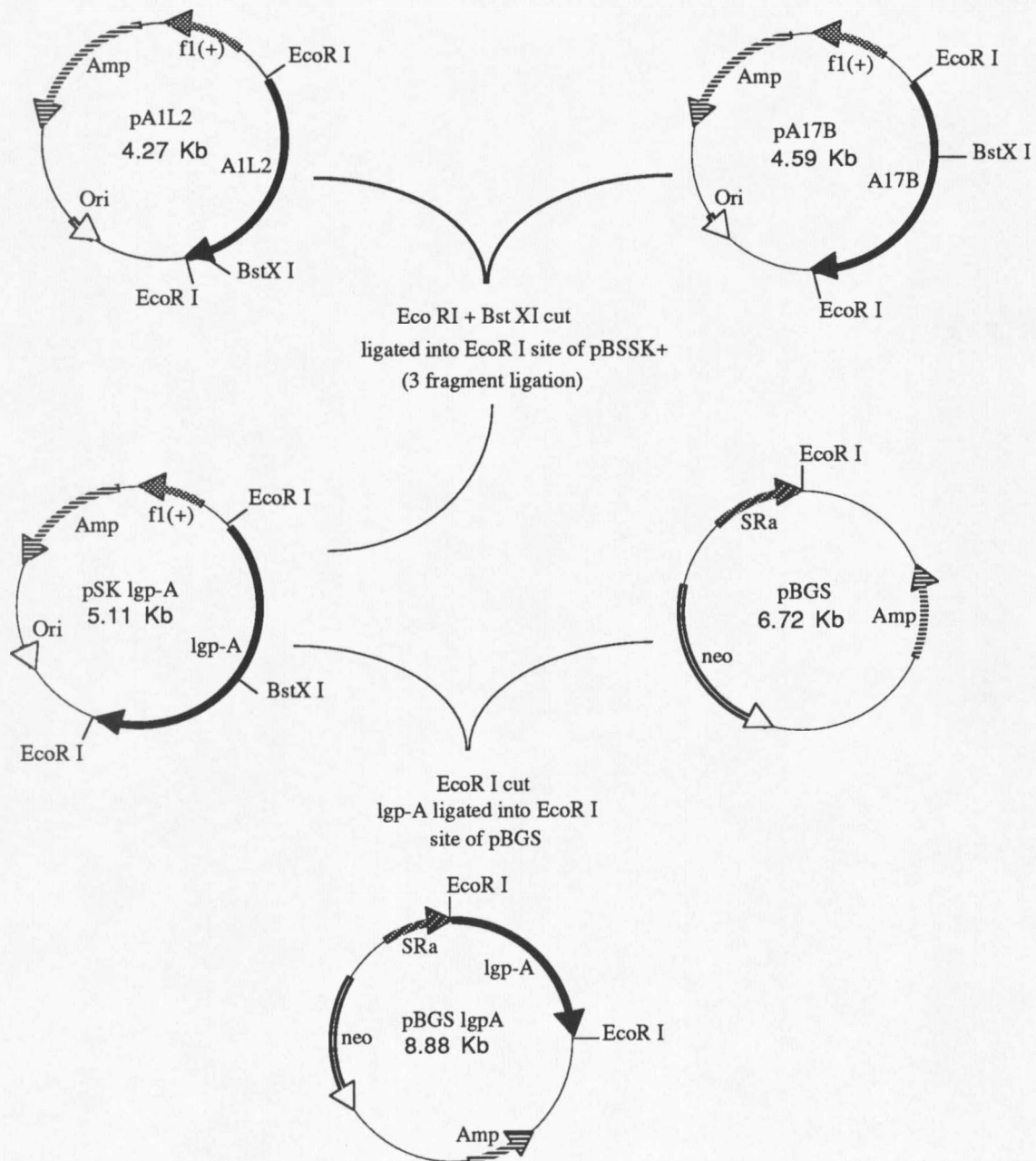


Figure 2. Construction of pBGSHA. Plasmids pA1L2 and pA17B are Bluescript plasmids with partial *lgp-A* cDNA clones; together the cDNA clones cover 135 bp 5'UT and all of the coding sequence and 3' UT of hamster *lgp-A* cDNA. The 5' part of the cDNA was obtained by digesting pA1L2 with *EcoR I* and *BstX I* and the 3' part was obtained by digesting pA17B with *BstX I* and *EcoR I*. The two inserts were ligated together into the *EcoR I* site of Bluescript to obtain pSK *lgpA*. pSK *lgpA* was digested with *EcoR I*, and the *lgp-A* cDNA was cloned into pBGS. Plasmids pBSHA and pBGSHA are denoted by pSK *lgp-A* and pBGS *lgpA*.

were digested with *Bst*X I and *Eco*R I, and the appropriate fragments were gel purified. These two fragments were ligated into the *Eco*R I site at the 3' end of the insert in pA1L2 (pA1L2 linearized by incomplete *Eco*R I digestion was inadvertently used instead of Bluescript as a vector). The resultant plasmid #14 had the full lgp-A cDNA as well as the A1L2 cDNA fragment at its 5' end. The A1L2 fragment was ultimately removed from plasmid #14 by deletion of the segment between the two *Nco* I sites. This plasmid (Bluescript SK+ with hamster lgp-A cDNA) was named pBSHA.

Plasmid #14 was digested with *Eco*R I, and the lgp-A cDNA was gel purified and cloned into the *Eco*R I site of pBGS, a eukaryotic expression vector with the SR α promoter (Takebe *et al.*, 1988) and a G-418-resistance cassette (B. L. Granger, unpublished). The resultant plasmid was named pBGSHA (Fig. 2).

Construction of pBGSHA5' Δ

The aim of this construction was to delete all of the 5' UT sequence of the lgp-A cDNA, except the 5 bp 5' to the start codon. In addition, we decided to change the G 3 bases upstream from the start codon to an A, which is more common at that position (Kozak, 1989). The plasmid pA1Ld*Bam* HI, with only one *Eco*R I site at the 5' end of the cDNA insert (the *Eco*R I site at the 3' end was deleted along with a part of the cDNA and vector for sequencing purposes), was digested with *Eco*R I and *Nco* I to delete the DNA between these two sites (Fig. 3). Since the *Nco* I site overlaps the start codon of the lgp-A cDNA, this deletion removed all of the 5' UT sequence. Two partially complementary 11 base oligonucleotides, BG11 (5' AATTCGTGCAC 3') and BG 12 (5' CATGGTGCACG 3'), were designed to form a double stranded insert with *Eco*R I and *Nco* I sticky ends. The oligonucleotides were dephosphorylated to prevent self-ligation, and the double stranded DNA was ligated into the above plasmid.

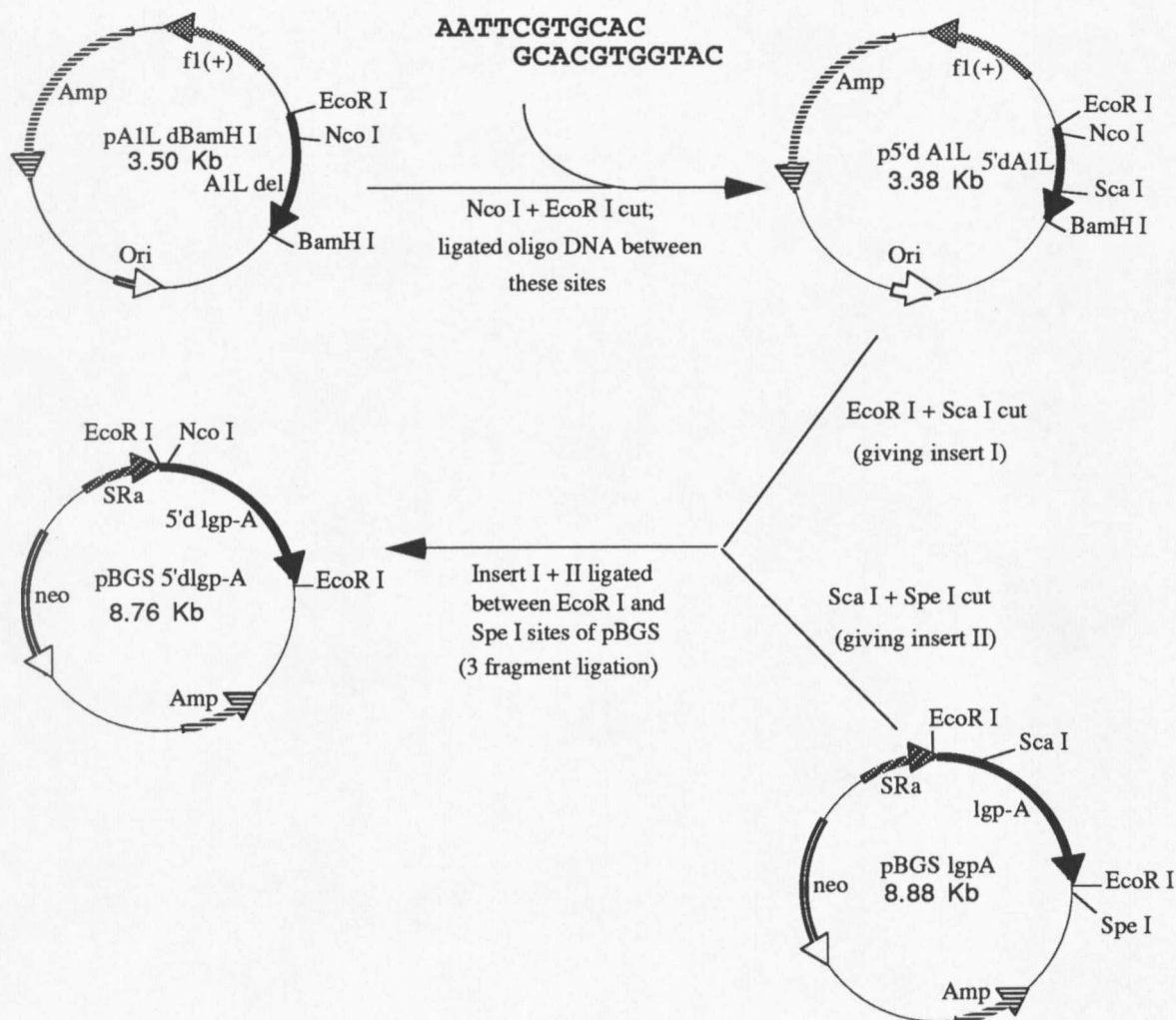


Figure 3. Construction of pBGSHA5'D. pA1LdBamH I (derived from pA1L2 by restriction enzyme deletion), which has only one *EcoR* I site, was digested with *EcoR* I and *Nco* I, deleting the sequences 5' to the *Nco* I site (at the start codon). Into this digested plasmid the oligo DNA (which has *EcoR* I and *Nco* I ends, formed by annealing the two single stranded oligos) was ligated, resulting in 5'dA1L. p5'dA1L was digested with *EcoR* I and *Sca* I and this insert (I) along with insert II obtained by digesting pBGS lgpA with *Sca* I and *Spe* I was ligated between the *EcoR* I and *Spe* I sites. Plasmids pBSHA and pBGSHA5'D are denoted by pSK lgp-A and pBGS 5'dlgpA.

Plasmid p5'dA1L was digested with *EcoR* I and *Sca* I, and the smaller fragment gel purified. Plasmid pBGSHA was digested with *Sca* I and *Spe* I, and the smaller fragment gel purified. These two fragments were ligated between the *EcoR* I and *Spe* I sites of pBGS (3 fragment ligation) to form pBGSHA5'Δ.

Construction of pBSHB and pBGSHB

cDNA inserts from plasmids pPCR, pHB10 and pHB9 were spliced together to make the construct pBSHB (Fig. 4). pHB10 was digested with *Bam*H I and *Xba* I, the cDNA fragment was gel purified and cloned into pPCR between the polylinker *Bam*H I and *Xba* I. The resultant plasmid, pHBP31, was digested with *Xba* I and dephosphorylated, and received the gel-purified fragment from pHB9 after *Xba* I and *Spe* I digestion. (It would have been sufficient to digest pHB9 with *Xba* I alone but an *Xba* I and *Spe* I digested and gel-purified fragment was already available). The resultant plasmid with the desired insert orientation was called pBSHB.

This hamster lgp-B cDNA construct does not possess a polyadenylation signal. To express this cDNA in mammalian cells, it was cloned into a vector, pBGSA, which is a derivative of pBGS that has a polyadenylation and transcription termination signal inserted. (This insert, from the human growth hormone gene, was PCR amplified and cloned between the *Xba* I and *Sac* I sites of pBGS to produce pBGSA; C. Bruno and B. L. Granger, unpublished.) The lgp-B cDNA was cut out from pBSHB with *EcoR* I and cloned into *EcoR* I digested and dephosphorylated pBGSA (Fig. 4).

Monoclonal Antibodies

Partial Purification of Antigen

CHO cells were grown to confluence in αMEM with 5% fetal bovine serum (FBS) in five 15 cm plates. Soluble proteins were removed by permeabilizing the cells on the

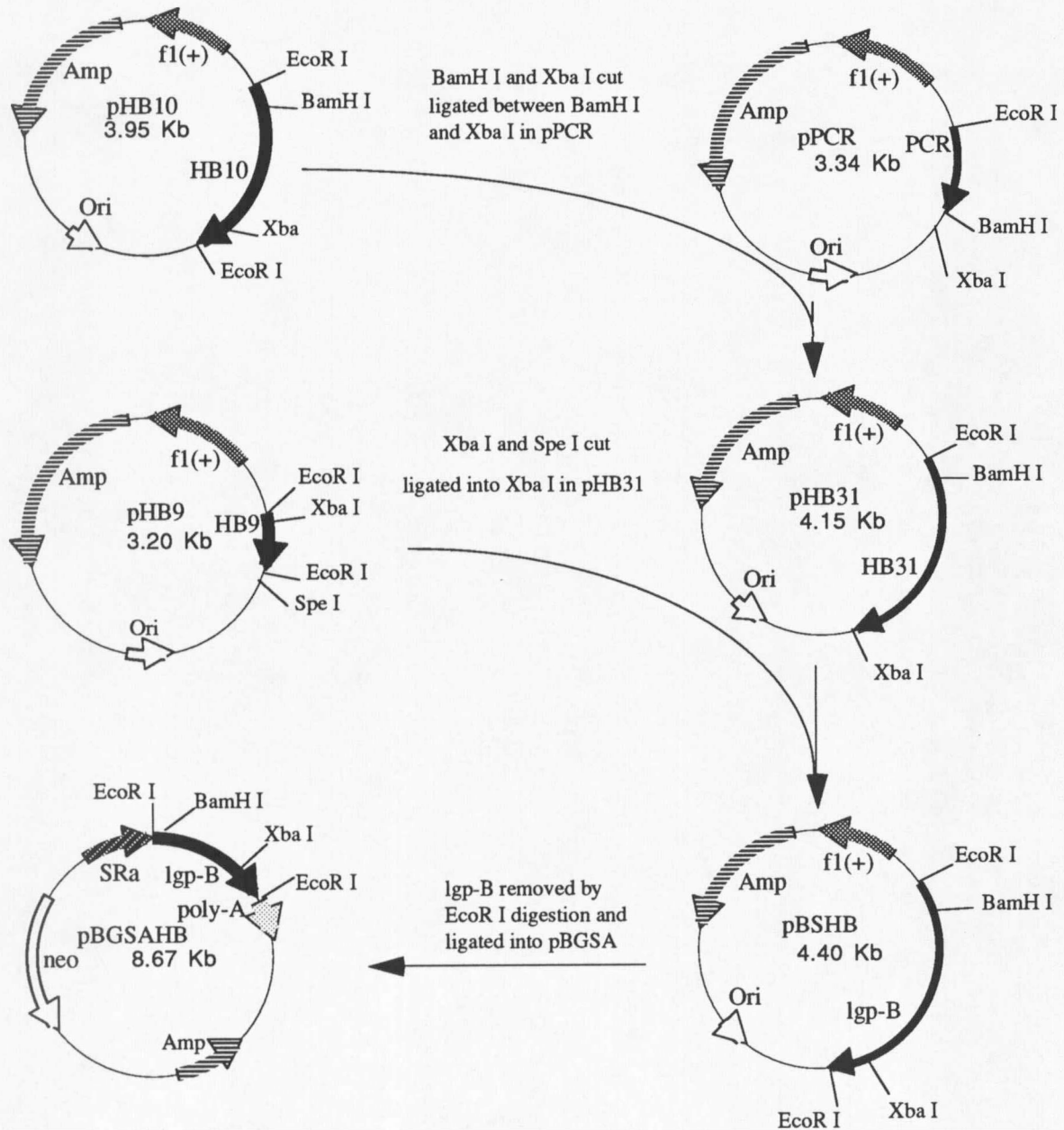


Figure 4. Construction of pBGSAHB. pPCR, pHB10 and pHB9 are Bluescript plasmids with partial *lgp-B* cDNA clones; together the cDNA clones cover 107 bp 5'UT, all of the translated coding sequence, and 82 bp 3'UT of hamster *lgp-B* mRNA. The insert in pHB10 was digested with *BamH* I and *Xba* I and cloned between the polylinker *BamH* I and *Xba* I sites of pPCR to obtain pHB31. pHB31 digested with *Xba* I received the *Xba* I-Spe I fragment from pHB9, resulting in pBSHB. *lgp-B* cDNA was removed from pBSHB by *EcoR* I digestion and cloned into the *EcoR* I site of pBGSA.

plates with cold extraction buffer (see Appendix) containing 0.1% saponin and 1mM PMSF. Membrane proteins were then solubilized with cold extraction buffer containing 1% Triton X-100. This extract was clarified by centrifugation and passed through a wheat germ agglutinin affinity column (Vector Laboratories, Inc.). Glycoproteins bound to the column were eluted with 3 ml of 0.5 M N-acetyl glucosamine, precipitated with 9ml of 100% ethanol, and the pellet was stored at -80°C until use.

Hybridomas and Monoclonal Antibodies

The pellet of partially purified membrane glycoproteins from CHO cells was dispersed in 600 µl of cmfPBS (calcium and magnesium free PBS). 200 µl of the protein suspension was emulsified in an equal volume of Hunter's TiterMax™ adjuvant (CytRx Corporation) and injected into the peritoneal cavities of two mice on day one. One booster injection (100 µl of the aqueous suspension of the protein per mouse) was given intraperitoneally, without adjuvant, on day 78, 3 days prior to sacrifice of one of the mice and removal of its spleen. Fusion of a portion of the spleen cells with SP2/0 mouse myeloma cells did not produce any viable hybridomas. Some of the remaining spleen cells, which had been frozen in aliquots, were later fused with P3U1 mouse myeloma cells, generating viable hybridomas. Culture supernatants were screened by indirect immunofluorescence of methanol-fixed or periodate-lysine-paraformaldehyde-fixed (McLean and Nakane, 1974) CHO cells; hybridomas secreting anti-CHO antibodies were propagated and preserved. The culture supernatants that reacted with CHO cells were again screened using NIH-3T3 (mouse) cells transfected with the hamster Igp-A or Igp-B cDNAs described earlier (see below), with untransfected NIH-3T3 cells serving as controls. Those that reacted with the transfected cells and not with untransfected cells were identified as being anti hamster Igp-A or Igp-B antibodies. These hybridomas were ultimately cloned by limiting dilution.

Expression in Heterologous Cells

Expression of hamster lgp-A, lgpA5' Δ and lgp-B in mouse cells

The hamster lgp-A and lgp-A5' Δ cDNAs were cloned into pBGS, and the lgp-B' cDNA was cloned into pBGSA (Figs. 2-4). NIH-3T3 (mouse fibroblast) cells were grown in DMEM containing 6% FBS. Cells were transfected by the calcium phosphate method (Chen and Okayama, 1988) with 20 μ g of plasmid DNA. The cells were subjected to selection in a medium containing 400 μ g/ml Geneticin (G-418 sulfate, GIBCO Laboratories) and cloned two weeks later. Individual colonies were harvested by scraping and aspirating simultaneously with a pipette tip.

Immunofluorescence

Cells (CHO, NIH-3T3 and transfected NIH-3T3) to be used for immunofluorescence were grown on coverslips or ten-well slides, and fixed in methanol or paraformaldehyde. All washings and incubations of the methanol-fixed cells were done in Tris-buffered saline containing 0.1% gelatin. The paraformaldehyde-fixed cells were permeabilized and washed in Tris-buffered saline containing 0.1% gelatin and 0.005% saponin. The cells were incubated for at least one hour with the primary (mouse anti-hamster) monoclonal antibody. Unbound primary antibody was then washed off, and a secondary, fluorescein-conjugated goat-anti mouse IgG was added and incubated for another hour. The unbound secondary antibody was also washed off, and the cells examined by fluorescence microscopy. Cells were photographed on 35mm Kodak Tri-X film and developed in Diafine (Acufine).

RESULTS

Library ScreeningSouthern Blot Analysis to Test Probes

To determine whether our existing nucleic acid probes for lgp-A and lgp-B would be useful for screening libraries from other species, we performed Southern blot analysis using hamster, bovine and mouse genomic DNAs. The DNAs were digested to completion with *EcoR* I (bovine, mouse and hamster) or *Hind* III (bovine) and electrophoretically resolved and transferred to a filter. The filter was probed with ³²P-labeled mouse lgp-A or rat lgp-B probes and autoradiographed. The bovine DNA showed two bands and the mouse and hamster DNAs each showed one band in the autoradiograms, suggesting specific cross-hybridization of the probes. This encouraged us to go ahead with the screening of a hamster cDNA library with mouse and rat nucleic acid probes.

Screening a Hamster cDNA Library for lgp-A

A commercial hamster (CHO-K1 cell) cDNA library in λ gt11 was plated with *E. coli* Y1088. Plaque lifts were prepared and probed with a ³²P-labeled mouse lgp-A cDNA. After the primary screening (3-300 x 10³ plaques per plate; 6 x 10⁵ plaques total), ten potential lgp-A clones were identified. The lambda lgp-A clones were named λ HA1 to λ HA10, (to denote that they are lambda, hamster, lgp-A clones). Six of these, λ HA1, λ HA3, λ HA4, λ HA7 and λ HA10, were chosen (on the basis of greatest signal intensity) for rescreening, and five were plaque-purified; inserts were cloned into the *EcoR* I site of Bluescript plasmids and sequenced. Phage subclones were given names that have letters and numbers following the names of parent clones, to denote their lineage and their

difference in purity from their parents (*e.g.*, λ HA1C1 is a twice-subcloned plaque from λ HA1).

Sequence data from the clones obtained after the first round of screening showed that none of the clones included a polyadenylation signal; comparison of the sequences with the known lgp-A sequences suggested that about 150 bp cDNA at the 3' end of the mRNA was absent. The library was therefore screened again using a smaller probe corresponding to the 3' terminal 330 bp of the mouse cDNA. Nine 15 cm plates with 300,000 plaques each were screened. Fifteen positives were identified, and were named λ HA11 to λ HA25. Nine of them were plaque purified, and DNA inserts from the four clones that gave the strongest signals were subcloned and sequenced.

A total of 3.3×10^6 plaques (in two separate screening) were screened, yielding 25 potential positive lgp-A clones; nine were used for sequencing.

Screening the Same Library for Hamster lgp-B

The filters used for screening for lgp-A were stripped of the lgp-A probe and re-probed with a ^{32}P -labeled rat lgp-B cDNA probe. Seven potential lgp-B clones were identified, and named λ HB1 to λ HB7. Only one of these (λ HB3) gave a strong signal, and the rest were found to be false positives upon rescreening. The insert from λ HB3 was subcloned into Bluescript and sequenced, as for the lambda lgp-A subclones.

Sequence data obtained from λ HB3 showed that the clone covered only a part of the coding sequence for the hamster lgp-B cDNA. The cDNA library was therefore screened again (see screening for lgp-A; same filters were used), yielding ten potential lgp-B clones, named λ HB8 to λ HB17. λ HB14 was found to be a false positive. All the other lgp-B clones were plaque purified and phage DNA extracted. Upon *EcoR* I digestion, λ HB13 did not yield an insert; perhaps one *EcoR* I site was lost during the construction of the library. Further mapping of λ HB13 revealed that it would not give any sequence not

covered by the other clones. The cDNA inserts from the other lgp-B clones were subcloned and sequenced.

A total of 3.3×10^6 plaques (in two separate screenings) were screened, yielding 10 positive lgp-B clones; nine were used for sequencing.

Gel Purification and Cloning of cDNA Inserts

Purified recombinant λ gt11 DNA was digested with *EcoR* I, and cDNA inserts were visualized by agarose gel electrophoresis. λ HA1, λ HB3 and λ HB11 gave two bands each, indicating the presence of internal *EcoR* I sites in the inserts. Inserts were given the names of the lambda clones from which they were excised, except that the λ sign was removed. Competent *E. coli* strains NM522 or XL1-Blue were used for growth of recombinant plasmids.

Sequencing

Sequencing Strategy of Hamster lgp-A and lgp-B

Sequencing of the cDNA inserts was carried out by the Sanger method of dideoxy nucleotide chain termination (Sanger *et al.*, 1977). Usually, 250 to 300 bp were read from each sequencing reaction from the autoradiograms. The cDNA inserts used for sequencing, and their relative locations with respect to the cDNA, are shown in Figures 5 and 6. Most of the sequencing was done with double stranded DNA. Some of the cDNA inserts gave sequences that were unrelated to any lgp; these sequences are denoted by dotted lines, and are presumably artifacts of library construction. Both strands of the lgp-A cDNA were sequenced. Both strands of the lgp-B cDNA were also sequenced, except for the 5' 135 bases (for which only one strand was sequenced).

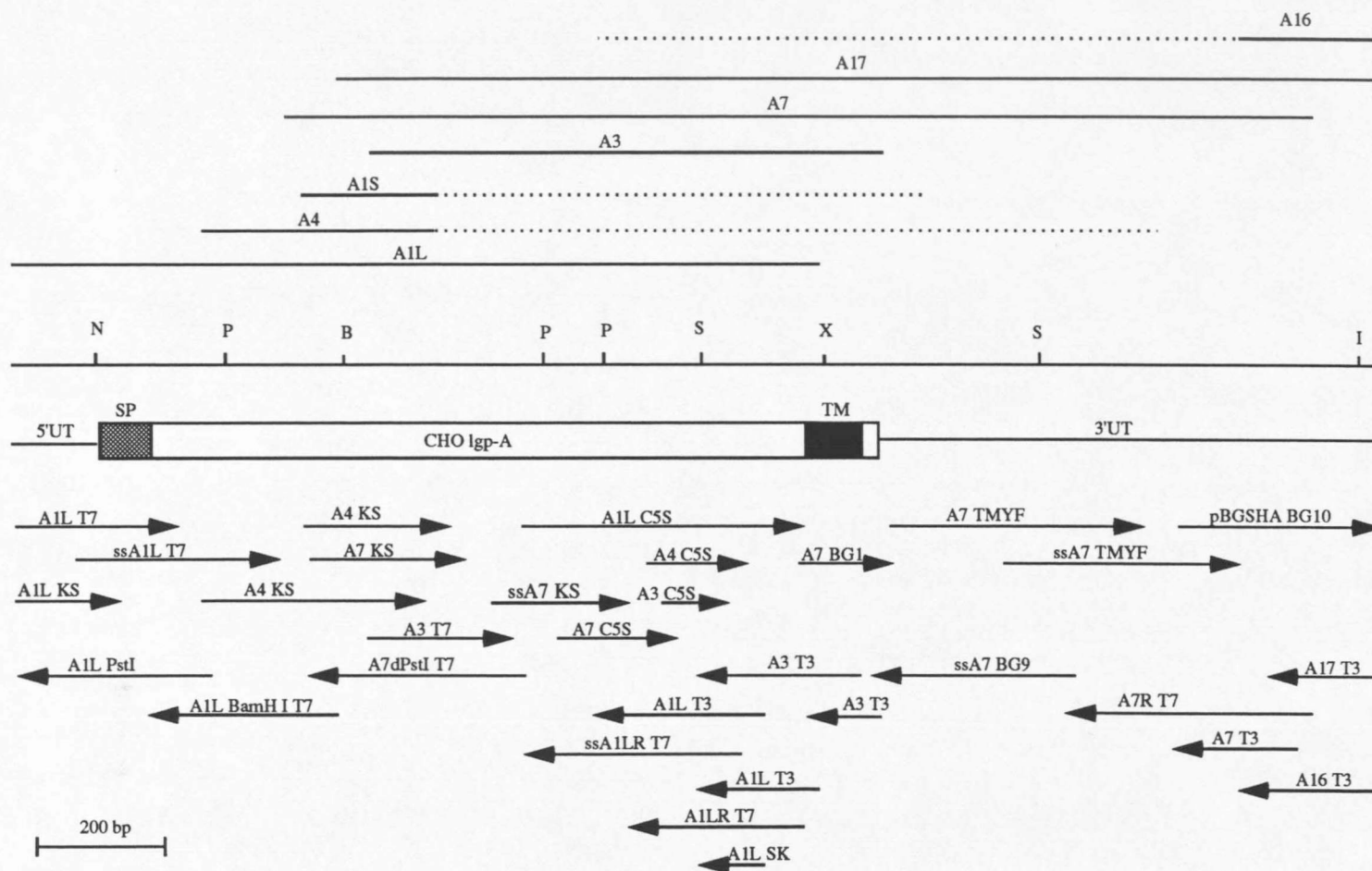


Figure 5. Sequencing strategy of hamster lgp-A cDNA. The translated region of the cDNA is shown as a box, with the untranslated (UT) regions on either side. SP and TM denote the signal peptide and transmembrane domain of hamster lgp-A. The arrows represent the directions and lengths of individual sequencing runs, with the names of the clones and primers indicated. The letters "ss" denote that single stranded DNA was used for sequencing. The lines above the cDNA show the names and relative positions of the lambda inserts used for sequencing after cloning into Bluescript. The dotted lines denote sequences that were unrelated to lgp-A and presumably artifacts of library construction. N, P, B, S, X and I denote restriction enzyme sites for *Nco* I, *Pst* I, *Bam*HI, *Sac* I, *Bst*X I, and *Nsi* I.

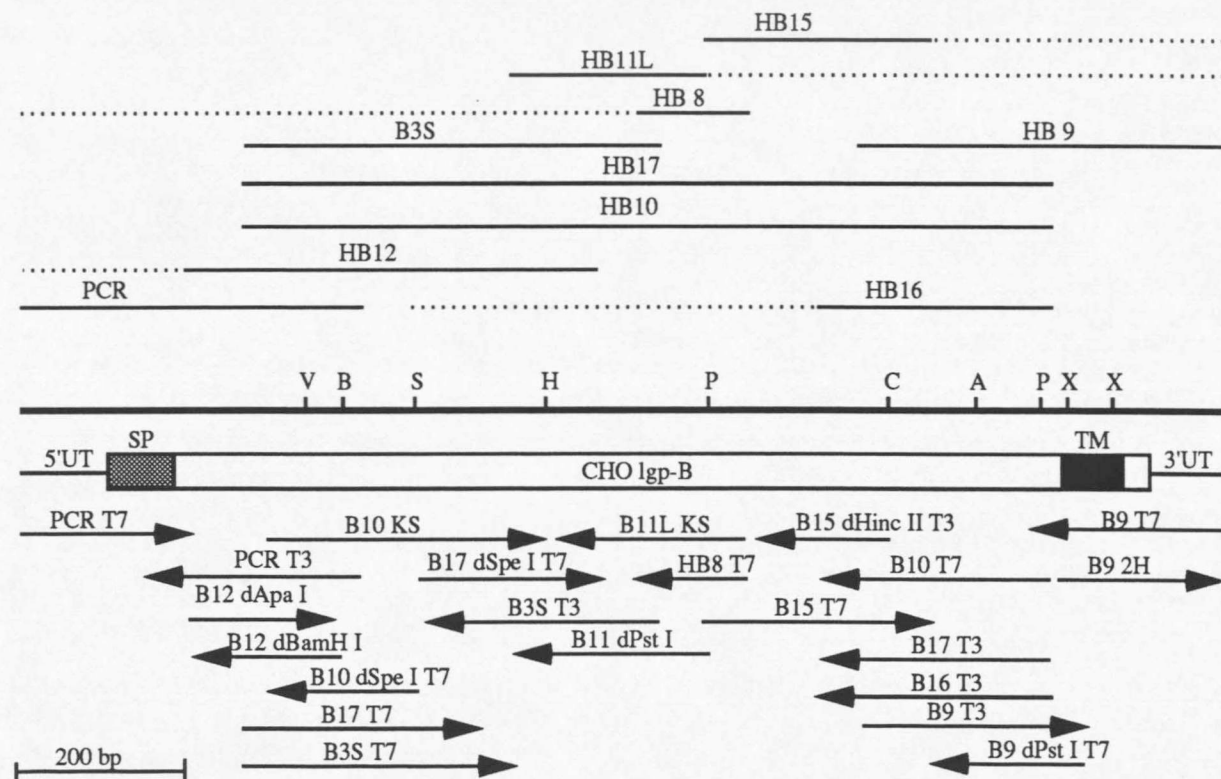


Figure 6. Sequencing strategy of hamster lgp-B cDNA. The translated region of the cDNA is shown as a box, with the untranslated (UT) regions on either side. SP and TM denote the signal peptide and transmembrane domain of hamster lgp-B. The arrows represent the directions and lengths of individual sequencing runs, with the names of the clones and primers indicated. The lines above the cDNA show the names and relative positions of the lambda inserts used for sequencing after cloning into Bluescript. The dotted lines denote sequences that were unrelated to lgp-B and presumably artifacts of library construction. V, B, S, H, P, C, A and X denote restriction enzyme sites for *Pvu* II, *Bam*H I, *Spe* I, *Hind* III, *Pst* I, *Hinc* II, *Xba* I and *Bst*X I.

Obtaining the 5' End of the lgp-B cDNA Using the RACE Procedure

Even though 3.3×10^6 phage from the cDNA library (having 1.8×10^6 independent clones) were screened, the 5' end of lgp-B cDNA (including all of the 5' UT and the region corresponding to the first 28 amino acids of the protein) was not found. A PCR based, RACE (Rapid Amplification of cDNA Ends) method was therefore employed to obtain these sequences. We used a 5'-AmpliFINDER RACE Kit (Clontech), which employs a modified SLIC (Single Strand Ligation to Single Stranded cDNA) procedure, designed specifically for amplifying the 5' end of a cDNA (see Methods for details).

A cDNA specific primer, P1, was annealed to total RNA from hamster cells, and first strand cDNA was synthesized using AMV reverse transcriptase. The anchor was ligated to the 3' end of the purified cDNA with T4 RNA ligase. PCR was performed using the anchor primer and nested cDNA primer (P2), and a 500 bp product was obtained. The RACE product was restriction mapped and found to have the expected *Bam*H I and *Pvu* II sites. The product was digested with *Bam*H I and *Eco*R I, cloned into Bluescript, and sequenced. The sequence obtained was aligned with the sequence of the hamster lgp-B cDNA and found to be identical in their region of overlap. The non-overlapping region was found to be highly similar to the corresponding region of lgp-B cDNAs from other known species. It was possible to establish a contiguous open reading frame when the sequence data from the RACE product was appended to the existing lgp-B cDNA sequence, and the translated sequence was found to be similar to other lgp-B sequences.

Hamster lgp-A Sequence

Overlapping inserts from two plasmids, pA112 and pA17B, were digested with *Bst*X I and spliced together. The total length of hamster lgp-A cDNA thus constructed was 2147 bp. Bases 1-133 comprise the 5'-untranslated (UT) region and 1358-2147 comprise the 3' UT region (Fig. 7). The coding sequence including the stop codon comprise bases

```

1  AATCGGGACA GGCCTGCGG GCCGCAGGGT CCCCTCCGTG CCGGCGCCGC
51 AGCTGCCGGG CCTCTCGAGT CGGTGGACGC GTCCTCCCGC CGGCCTCGCA
101 ACCGCCGCAC CCCGGCCTCG GCTGCGTCGC GCCATGGCGG CCCCTGGCGC
151 CCCGCGGTCG CTGCTCCTGT TGCTGCTGGC AGGCCTTGCA CACGGTGCCT
201 CAGCATTGTT TGTGGTGAAA GACAGCAATG GCACAGCTTG TATAATGGCC
251 AACTTCTCTG CCTCCTTCTT CACCATCTAC GAGACTGGAC ATGGTTCTAA
301 GAACTCAACC TTTGAACTGC CGTCCTCTGC AGAAGTACTG AATAGTAACA
351 GTTCTTGTGG TAGAGAAAAT GTTTCTGAAC CCATCCTCAC AATTGCTTTT
401 GGAAGCGGAT ATTTATTGAC GCTCAACTTC ACAAGAAATG CAACACGTTA
451 CAGTGTCCAG GACATGTATT TTGCTTATAA CTTGTCAGAC ACTCAACATT
501 TCCTGAATGC CAGCAACAAG GGGATCCACT CTGTAGACTC CTCAACTGAC
551 ATCAAGGCAG ACATCAACAA AACATACCGG TGTTTGAGTG CTATCCAGGT
601 CCACATGGGC AATGTGACCG TCACGCTCAG TGATGCCACT ATCCAGGCCT
651 ACCTGTTGAA TAGCAACTTC AGCAAGGAAG AGACACGCTG CACGCAGGAT
701 GGACCTTCCC CGACCACTGT GCCGCTAGC CCCTCACCAC CACTTGTGCC
751 CACAAACCCC ACTGTGATTA AGTACAATGT GACTGGTGAA AACGGAACCT
801 GCCTGCTTGC CTCTATGGCC CTGCAGATGA ACATCACCTA CATGAAGAAG
851 GACAACATGA CTGTGACCAG AGCATTGAAC ATCAGCCCAA ATGACACAGC
901 TAGTGGGAGC TGCAGTCCTC ATGTGGTGAC CTTGACAGTG GAGAGCAAGA
951 ACAGTATCCT GGATTTGAAG TTTGGGATGA ATGGCAGTTC TAGCCTGTTT
1001 TTCCTACAAG AAGTCCGGTT GAACATGACT CTTCCCGATG CCAACGTGTC
1051 GTCACATCATG GCCTCCAACC AATCACTGAG AGCTCTTCAG GCCACGGTCG
1101 GGAACCTCCTA CAAGTGCAAC ACCGAGGAGC ACATCTTTGT CACCAAGGAG
1151 TTTTCCCTCA ATGTCTTCAG CGTTCAGGTC CAGGCTTTCA AGGTTGAAAG
1201 TGACAGGTTT GGGTCTGTGG AAGAGTGTAT GCAGGATGGT AACAAACATGC
1251 TGATCCCCAT TGCTGTGGGT GGTGCCCTGG CAGGGCTCGT CCTCATCGTC
1301 CTCATCGCCT ACCTCATCGG CAGGAAGAGG AGTCATGCCG GCTACCAGAC
1351 CATCTAGCCG GTGGACAGGC GGGTGCGCGT GCGCCAGACG CACAGGGGCC
1401 TGTTCTTATG TCCCCGAGCT TAGATAGGTG TGGAAGGGAG GCACACTTTC
1451 TGGCAAACCTG TTTTCAAATC TGCTTTATCA AATGTGAAGT TCATCTTGCA
1501 ACATTTACTA TGCACAAAGG AATAACTATT GAAATGACGG TGTTAATTTT
1551 GCTAACTGGG TTAAATATTT TGCTAACTGG TTAAATGTTA ATGTTACCAA
1601 AGTAGAGCTC TAAGGAGGAC AAGAAGGCTT TCAGAATCGC ACGGCTCCCC
1651 CATTTGACTT TTTTAAGACT TGGTGTGTTG TTTCTTCATC CTGTTACTCA
1701 GATTTAAGCC TTACAAAGGG AATCTCTGGC CTGACACTTG GCGTGGTAA
1751 GGGCGGCTGA TCTTTAGGCT GCACGCTGAA GATGCTAGAA GGAAGAGCA
1801 GGAAGCCCT GCACAGAGGC ATTCACAGGG GCAACCTCTG GACAGGCTTG
1851 GGCTGCCTGG CCTGGGGACC GGAATCTGGC ATCTGGCTGG GCACAACCAC
1901 CAGTGTCTCT GTGCTCTGCC ACCCATGAGC TGGTACTTTT CTAAATAGAA
1951 AAAGGCATTA TTTTTATTTA CTTTTTGTA AGTGATTTC AGTCTTCTGT
2001 TGGTGTGAG GGTGGCCCCG TTTCTGCACT GTGTACATA AAAGATTAC
2051 ACTGCTGACC TGTCCTGCCG CATAGGTGGG TTGTACACTG AGGACCAGCT
2101 CACGTAATGC ATTGCTGTGA ACGATGCTAA TAAAAAGTCT CCTTCTTAA
2151 AAAAAAAAAA AAAAA

```

Figure 7. Nucleotide sequence of hamster Igp-A cDNA. The protein coding region is from 134 to 1357. Polyadenylation signals are found in two places: 2037 to 2042 and 2129 to 2134. The start codon (ATG), stop codon (TAG) and the polyadenylation signals (AATAAA) are shown in bold. (This nucleotide sequence has been submitted to GenBank™ with the accession number L18986)

```

1  MAAPGAPRSL LLLLLAGLAH GASALFVVKD SNGTACIMAN FSASFFTIYE
      .                      *
51  TGHGSKNSTF ELPSSAEVLN SNSSCGRENV SEPILTIAFG SGYLLTLNFT
      *
101 RNATRYSVQD MYFAYNLSDT QHFLNASNKG IHSVDSSTDI KADINKTYRC
      *
151 LSAIQVHMGN VTVTLSDATI QAYLLNSNFS KEETRCTQDG PSPTTVPPSP
      *
201 SPPLVPTNPT VIKYNVTGEN GTCLLASMAL QMNITYMKKD NMTVTRALNI
      =====
251 SPNDTASGSC SPHVVTLTVE SKNSILDLKF GMNGSSSLFF LQEVRLNMTL
      *
301 PDANVSSLMA SNQSLRALQA TVGNSYKCNT EEHIFVTKEF SLNVFSVQVQ
      *
351 AFKVESDRFG SVEECMQDGN NMLIPIAVGG ALAGLVLIVL IAYLIGRKRS
      *
401 HAGYQTI

```

Figure 8. The predicted protein sequence of hamster Igp-A. The translated polypeptide contains 407 amino acids. A dot appears below the probable (by comparison with the mouse sequence) amino terminal amino acid of the mature polypeptide (after signal peptide cleavage). Bold Ns are potential N-linked glycan addition sites (in the consensus sequence -Asn-X-Thr/Ser-). An asterisk (*) appears below each of the 8 cysteines. Two lines (=) appear below the proline-, threonine- and serine-rich variable domain. The putative transmembrane domain is underlined (-).

134-1357. Two potential polyadenylation signals were identified in the 3' UT region: bases 2037-2042 and 2129-2134. An 18 base poly A tail was present after position 2147. The translated protein is 407 amino acids long with 23 potential N-linked glycosylation sites, a putative signal peptide of 24 amino acids, a transmembrane domain of 25 amino acids, and a cytoplasmic tail of 11 amino acids (Fig. 8).

Hamster lgp-B Sequence

The lgp-B cDNA fragments cloned into plasmids pPCR, pHB10 and pHB9 were spliced together in a stepwise manner to construct the full hamster lgp-B cDNA. The total length of lgp-B cDNA was 1422 bp. Bases 1-107 comprise the 5'-untranslated (UT) region and 1341-1422 comprise the 3' UT region (Fig. 9). The coding sequence including the stop codon comprise bases 108-1340. The translated protein is 410 amino acids long, has 17 potential N-linked glycosylation sites, a putative signal peptide of 28 amino acids, a transmembrane domain of 26 amino acids, and a cytoplasmic tail of 10 amino acids (Fig. 10). Bases 1-190 were obtained using the 5'-AmpliFINDER RACE procedure (Fig. 1). The sequence information for bases 1-132 was obtained by sequencing only one strand of the cDNA. The missing portion of the 3' UT region (containing an expected polyadenylation signal and a poly A tail) was not found in the cDNA library.

Monoclonal Antibodies

A partially purified protein fraction from CHO cells, enriched for membrane glycoproteins, was used to immunize mice. One booster injection was given on day 78, 3 days prior to sacrifice of the mouse and removal of the spleen. The fusion of P3U1 mouse myeloma cells with the spleen cells from the immunized mouse resulted in viable hybridomas in 74 wells. The supernatants from the hybridomas were screened by immunofluorescence of CHO cells. Supernatants from 35 hybridoma cultures reacted with

```

1  AGCTTTCTCG..GTGCTCTTGT..CGGTGGTCAC..CTGCGCTGTC..GCCTTCCCCG
51  GGGCTGTGAC..GGTGTTCCTT..AGCTTTGCCG..GCCAGGTCGA..CTCCTGCCCT
101 GCGGATCATG..ATGTGCTTCC..GCCTCTCTCC..GGTTTCGGGC..TCGGGGCTCG
151 TCCTGTCTTG..TCTCCTCCTG..GGAGCTGTTT..AGTCCTATGC..ATTCGAACTT
201 AATTTGCCAG..ATTCAAAAGC..TACTTGCCCTT..TTTGCCAAAT..GGAAGATGAA
251 TTTACAATA..TCCTATGAAA..CTACAACAAA..CAAAACCCTT..AAAAGTGTAA
301 CCATTTTCAGA..ACCTCACAAAT..GTGACATATA..ATGGAAGCAG..CTGTGGTGAT
351 GACCAGGGTG..TGGCTAAAAT..AGCTGTGCAA..TTTGGATCCA..CTGTTTCTTG
401 GAATGTGACT..TTTACTAAGG..AAGAATCTCA..TTATGTTATT..GGCAGCATCT
451 GGCTCGTGTA..CAACACTAGT..GATAACACAA..CATTTCTCTG..TGCTATCCCT
501 AAAGGAAGTG..CTACTGTTAT..CAGTTCTCAG..AGCATCGAAA..TTCCATTGGA
551 TGACATCTTT..AGGTGCAATA..GTTTACTAAC..TTTCAAGACG..GGTAATGTGG
601 TCCAGAACTA..TTGGGATATT..CACCTGCAAG..CTTTTGTTCA..AAATGGCACA
651 GTGAGTAAAG..AAGAATTTGT..GTGTGAGGAA..GACAAATCTG..TCACCACTGT
701 AAGACCCATC..ATTCACACTA..CTGTGCCACC..TCCTACTACG..ACACCCACTC
751 CACTGCCTCC..AAAGGTGGA..AACTATTCCG..TTAGTAATGG..CAATGCTACC
801 TGTCTGCTGG..CTACCATGGG..ACTGCAGCTG..AATGTCAGT..AGGAGAAGGT
851 GCCTTTCATT..TTTAACATCA..ACCCAGTAC..AACTAACTT..ACCGGCAGCT
901 GTCATCCTCA..AACTGCTCAA..CTTAGGCTGA..ACAACAGCCA..AATTAAGTAT
951 CTTGACTTTA..TCTTTGCTGT..GAAAAGTGAA..AGCCATTTCT..ATCTGAAGGA
1001 AGTTAACGTC..AGCATGTATA..TGGCTAATGG..CTCAGTTTTT..AGTGTTGCAA
1051 ATAACAACCT..TAGCTTCTGG..GATGCCCTC..TTGGAAGTTC..TTATATGTGC
1101 AACAAAGAGC..AGGTTGTTTC..AGTGTCTAGA..ACGTTTCAGA..TAAACACCTT
1151 TAACCTAAAG..GTGCAACCTT..TTAATGTGAC..GAAAGGAAAG..TATGCCACAG
1201 CTCAAGACTG..CAGTGCAGAT..GAAGACAACT..TCCTTGTGCC..CATAGCGGTG
1251 GGAGCAGCTC..TGGCAGGAGT..ACTTGCTCTA..GTGTTGCTGG..CTTATTTTAT
1301 TGGCCTCAAG..CGCCATCATA..CTGGATATGA..GCAGTTTTAG..GACCTGCAAT
1351 CCAATTGATT..ATATTATAAA..ATACATTCAA..AGAACGAGGT..TTTCTTGCCCT
1401 GTCAGTGCTC..CTTAAAACAC..AC

```

Figure 9. Nucleotide sequence of hamster lgp-B cDNA. The protein coding region is from 108 to 1340 (start and stop codon are given in bold). Residues 1 to 190 were from PCR amplification (dotted underline). Primers P1 (486 to 510) and P2 (441 to 465) are underlined. The complete 3'UT region was not obtained. (This nucleotide sequence has been submitted to the GenBankTM with the accession number L19357)

```

1  MMCFRLSPVS  GSGLVLSCLL  LGAVQSYAFE  LNLPSKATC  LFAKWKMNFT
      .                *
51  ISYETTTNKT  LKTVTISEPH  NVTYNGSSCG  DDQGVAKIAV  QFGSTVSWNV
      *
101 TFTKEESHYV  IGSIWLVYNT  SDNTTFPGAI  PKGSATVISS  QSIEIPLDDI
151 FRCNSLLTFK  TGNVVQNYWD  IHLQAFVQNG  TVSKEEFVCE  EDKSVTTVRP
      *                *                =====
201 IIHTTVPPPT  TTPTPLPPKV  GNYSVSNGNA  TCLLATMGLQ  LNVTEEKVPF
      =====
251 IFNINPSTTN  FTGSCHPQTA  QLRLNNSQIK  YLDFIFAVKS  ESHFYLKEVN
      *
301 VSMYMANGSV  FSVANNNLSF  WDAPLGSSYM  CNKEQVVSVS  RTFQINTFNL
      *
351 KVQPFNVTKG  KYATAQDCSA  DEDNFLVPIA  VGAALAGVLA  LVLLAYFIGL
      *                -----
401 KRHHTGYEQF

```

Figure 10. The predicted protein sequence of hamster Igp-B. The translated polypeptide contains 410 amino acids. A dot appears below the probable (by comparison with the mouse sequence) amino terminal amino acid of the mature polypeptide (after signal peptide cleavage). Bold Ns are potential N-linked glycan addition sites (in the consensus sequence -Asn-X-Thr/Ser-). An asterisk (*) appears below each of the 8 cysteines. Two lines (=) appear below the proline- and threonine-rich variable domain. The putative transmembrane domain is underlined (-).

hamster antigens, some of which appeared to be lysosomal. These hybridomas were named UH1 through UH35 ("U" for Uthayakumar, "H" for hamster). The 35 supernatants were further screened for the presence of anti-hamster lgp-A and anti-hamster lgp-B antibodies using NIH-3T3 cells transfected (see below) with hamster lgp-A (3T3HA) or lgp-B cDNAs (3T3HB). Untransfected NIH-3T3 cells were used as negative controls. Supernatant from one hybridoma (designated UH1) reacted specifically with 3T3HA cells and not with 3T3HB or control NIH-3T3 cells. This hybridoma thus secreted anti-hamster lgp-A antibodies. Conversely, using the same tests, UH3 was found to secrete anti-hamster lgp-B antibodies.

Expression of Hamster lgp-A and lgp-B cDNAs

The hamster lgp-A cDNA was cloned into the expression vector, pBGS (forming pBGSHA). The lgp-A cDNA from which the 5' UT region was deleted (see Methods) was also cloned into pBGS (forming pBGSHA5' Δ). The hamster lgp-B cDNA was cloned into the expression vector, pBGSA (forming pBGSAHB), which has polyadenylation and transcription termination signals derived from the human growth hormone gene. Mouse NIH-3T3 cells were transfected with the above constructs, using the calcium phosphate method (Chen *et al.*, 1988). The transfected cells were called 3T3HA, 3T3HA5' Δ and 3T3HB.

Immunofluorescence assays were performed on paraformaldehyde-fixed 3T3HA and 3T3HA5' Δ cells using supernatants from UH1. The fluorescence pattern was found to be lysosomal (Fig. 11), comparable to the pattern produced by two control monoclonal antibodies (1D4B and GL2A7, which are rat anti-mouse lgp-A and rat anti-mouse lgp-B antibodies, respectively; Hughes and August, 1982; Granger *et al.*, 1990; data not shown). Whereas the expression level of hamster lgp-A in 3T3HA was comparable to that of the

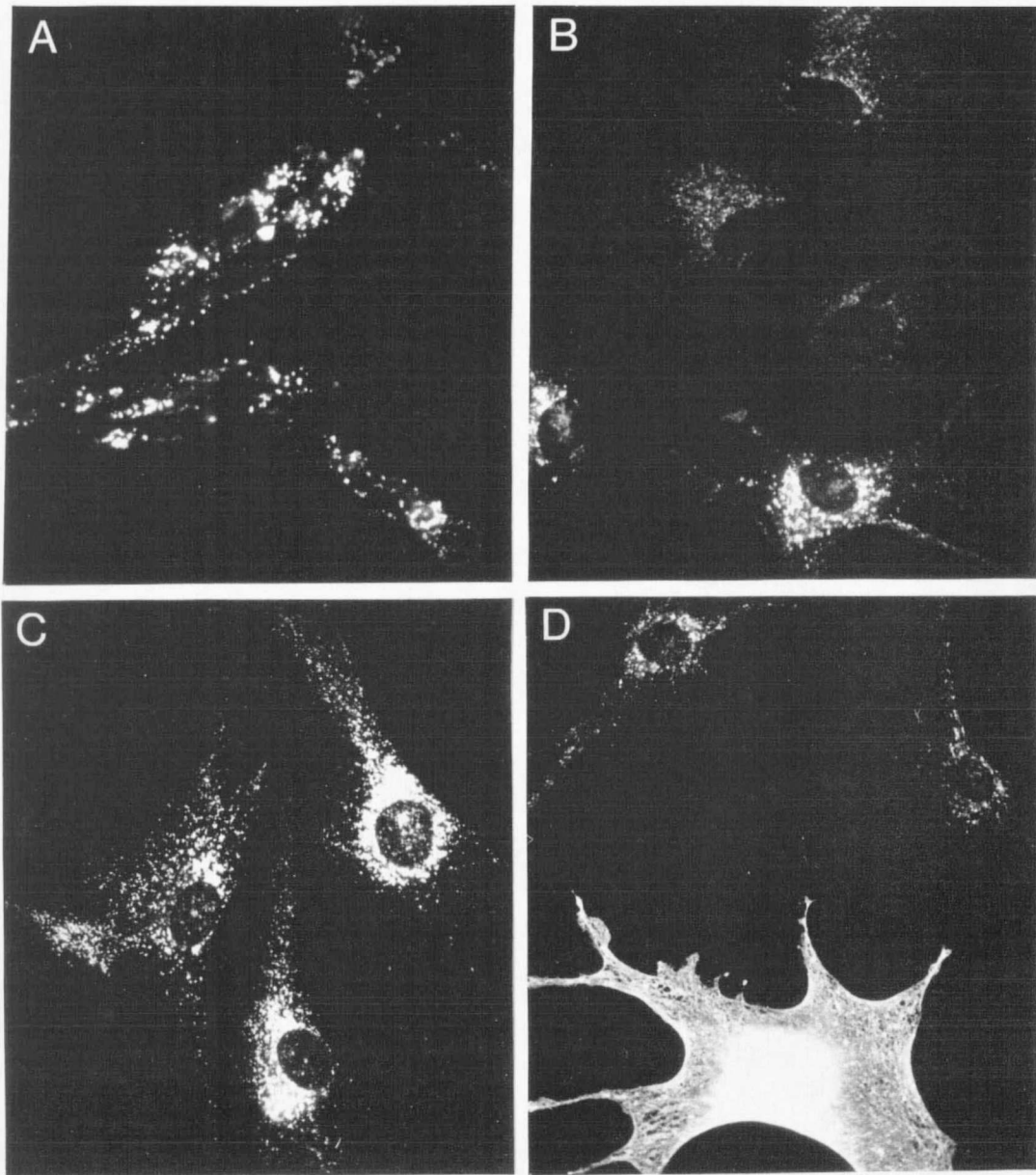


Figure 11. Immunofluorescence micrographs of endogenous and foreign hamster Igps. UH1 hybridoma supernatant (containing mouse anti-hamster lgp-A) was used to label the endogenous hamster lgp-A in CHO cells (A) or foreign hamster lgp-A expressed by transfected mouse NIH-3T3 cells (B). UH3 hybridoma supernatant (containing mouse anti-hamster lgp-B) was used to label hamster lgp-B expressed by transfected mouse NIH-3T3 cells (C and D). The cells in panel B and D were grown in the presence of sodium butyrate to increase expression of the foreign lgp. The CHO cells were fixed with formaldehyde and the NIH-3T3 cells with methanol prior to processing with the monoclonal antibodies and a fluorescein-conjugated goat anti-mouse IgG secondary antibody.

endogenous mouse lgp-A, the expression level in 3T3HA5'Δ was found to be very low, judging from the fluorescence intensity. Similarly, when the immunofluorescence assay was repeated with 3T3HB cells using supernatants from UH3, staining was comparable to the pattern produced by two control monoclonal antibodies, 3E9d9 and 1B3-13, specific for hamster lgp-B (kind gifts of Dr. Sandra Schmid, Research Institute of Scripps Clinic, La Jolla, CA). It was possible to induce over-expression of the transfected hamster lgps in the 3T3HA and 3T3HB cells by growing the cells in the presence of 5-10 mM sodium butyrate (Gorman and Howard, 1983). When the hamster lgps were over-expressed, they were found on the plasma membrane in addition to lysosomes (Fig. 11).

DISCUSSION

With an overall aim of identifying the signals involved in sorting and transport, and characterizing the structures and functions of lysosomal membrane proteins, we cloned and sequenced the hamster lgp-A and lgp-B cDNAs. We also expressed the hamster lgp-A and lgp-B cDNAs in mouse NIH-3T3 cells, and generated monoclonal antibodies specific for hamster lgp-A and lgp-B.

Structure of lgps and Sequence Comparisons

lgps have been characterized and their cDNAs cloned and sequenced from mouse, rat, human and chicken cells (see Introduction). Comparison of the hamster lgp-A and lgp-B cDNA and protein sequences (Pearson and Lipman, 1988) to the corresponding sequences from these other species revealed that they are highly similar. All are type I membrane glycoproteins (single transmembrane segments and cytosolic C-termini) with 380-396 amino acids (not counting the signal peptides; Figs. 12-14). Figures 12 and 13 show that the 8 cysteines, a stretch of 30 amino acids around the 5th cysteine, the transmembrane domain and the cytosolic tail are the most highly conserved regions of the proteins. The central variable region that separates the two luminal domains of hamster lgp-A is rich in proline (45%), threonine (20%) and serine (15%); that of lgp-B is rich in proline (32%) and threonine (32%) (Figs. 12 & 13).

The identity between hamster lgp-A and other known mammalian lgp-As is 66-79% (Table 1), whereas chicken lgp-A is only 46% identical; the identity between hamster lgp-B and the other known lgp-Bs is 78-82% (Table 2). The identity between lgp-A and lgp-B within a species is 33-34% (Table 3). This demonstrates that lgp-A from one species is more similar to lgp-A from other species than to lgp-B from the same species. Similarly,

CH MAAPGAPRSLLLLLLAGLAH...GASALFVVVD.SNGTACIMANFSASFFTIYETGHGSKNSTFELPSSAEVL
 MA ---G---P-----N---T---S---L-T---AN---QIVNIS--A-----
 RA ---G---P-----V-----S=P---E---N-----S---L-T-DA--V--V-NMT--A-----
 HA --PRS---P-----PVAA-RPHALSSA_M-M--N.G-----A-SVN-D-KS-P--M--D---D-T-V
 CA -GG.A--AV--GF-----QASSS...-D-R--T-KV--I--LTVA-SVE-KSSGQKQFAH-F--QN-TS.
 CH NSNSSCGRENVSEPILTIAFGSGYLLTLNFTRNATRYSVQDMYPAYNLSDTQHFLNASNKGIIHSVDSSTDIKAD
 MA KNG---K---D-S---T-R-----K-T-----H---T-----E--P--IS-E-YTM--T-----
 RA KNS---EK-A---T-A-T--E---K-T--K-T-----H---T-----F-P--S--PDT--T-----
 HA LNR---K--T-D-S-V-----HT-----A-----L-S-V-----HL-P--S-E-KT-E-I--R--
 CA Q-H---EG-T-H---ALS--A-H-IS---SKTLDK-Q-EELT-H---ETL-P--TEGKVMVATQKSV-Q-R
 CH INKTYRCLSAIQVHMGNVTVTLSDATIQAAYLLNSNFSKEETRCTQDGPSPPTT.....VPPSPSPPLVPTN.
 MA ---A---V-D-R-Y-K---V-R-----SSG-----H---G-----G-----
 RA ---V-D-R-Y-K---IV-W-----PS-----P-Q-----G-----
 HA -D-K---V-GT---N-----H-----S--S--RG---E--R-----APPA-----SP--KS.
 CA -GTE---INSKY-R-KH-NI-F-NV-LE--PT-DT--ANK-E-RE-MV-T--VAPTTPKHATSOVPTTSPAPTA
 CHPTVIKYNVTGNGTCLLASMALQMNITYMKKDNMTVTRALNISPNDTA.SGSCSPHVVTLTVESKNSILD
 MAS---N-----L---L---K-----F-----S---GINL---K--N--RA-E
 RAS-S-----D-----L-----T---F--N-S-.KY--T-GAQL--K-GN-SRV-E
 HAS-D---S-T-----G--L-L---R--T---L--N--K-SA---GA-L---ELH-EGTTVL
 CA APSS-A-G-----A---V---G--L---V---EKMGLDL--FI-HN-SA--M-ESTSAF-NLAFKTKIT
 CH .LKFGMNGSSSLFFLQEVRLNMTLP.DANVSSLMASNQSLRALQATVGNSYKCNTEEHIFVTKEFSLNVFSVQV
 MA .-Q---A-----G-----L-PTFSI--H--K-----S-ML-----
 RA .-Q---AT-----G-Q-----IEPTFST--Y--K---S-----S-ALA-----
 HA LFQ---A--R---GIQ--TI--RDPAFK-A-G-----A--VR--A--V-I-K-W-
 CA .FH-VL-A-EK---G-NVST--SE-KAPTFE--D-MSES--SA--NFQ--DKALV---N---
 CH QAFKVESDRFGSVEECMQDGNMMLIPIAVGGALAGLVLIIVLIAYLIGRKRSHAGYQTI 383
 MA ---D---V----- 382
 RA ---R-----V----- 386
 HA ---GGQ---LL-E-ST-----V----- 389
 CA ---DG-K--AM---QL-E-----I--A----- 396
 TM

Figure 12. Aligned amino acid sequences of lgp-A. CH, MA, RA, HA and CA denote hamster, mouse, rat, human and chicken lgp-A, respectively. Periods represent gaps that have been introduced to optimize alignments. The amino acid residues identical to hamster lgp-A are denoted by hyphens. Amino acids that are identical in all known lgp-A and lgp-B sequences are denoted by large dots above. The variable domain (VD) and transmembrane domain (TM) are underlined. The amino terminus of the mature proteins is also underlined. (Data are from Chen *et al.*, 1988; Howe *et al.*, 1988; Fukuda *et al.*, 1988; Fambrough *et al.*, 1988; Granger *et al.*, 1990 and this thesis)

```

CB  M M C F R L S P V S G S G L V L S C L L L G A V Q S Y A F E L N L P D S K A . T C L F A K W K M N F T I S Y E T T N K T L K T V T I S E P H N V T
RB  .....--RSD-LK---T---G---Y-E-E---T---AL..-VNE---TV-DK--
MB  .....-K-AK-I-IF-F---N-LIV--T---G---Y-E-E---T---.Q-N--I--AV-DKA-
HB  -VC---F--P-----V--V---R---L-----T-ENA---Y---Q---VR-----Y-----DHGT--

CB  Y N G S S C G D D Q G V A K I A V Q F G S T V S W N V T F T K E E S H Y V I G S I W L V Y N T S D N T T F P G A I P K G S A T V I S S Q S I E I P L
RB  -----KNG---MI-Y---L--A-N---A-Q-F-NN-T-S---N-TK-----V---IL---IPVGSQ L--
MB  HD-----RNS---MI---FA---A-N---A---S-HD-V-S-----S-V-----VA--VH--KNPEN EKV--
HB  ----I-----NGP-----PGF--IAN---AA-T-SND-VSFS--G-----DAED--IL--DELLA-R--

CB  D D I F R C N S L L T F K T G N V V Q N Y W D I H L Q A F V Q N G T V S K E E F V C E E D K S V T T V R P I I H T T V P P P T T T P T P L P P K . .
RB  G V --K-S-V---NLSP---H-G-----H-Q--K--TA--A-----L--TSIP..
MB  -V--K--V--YNLTP--K--G-----N-Q---QTP--A-----A-ST--L--TSTPTP
HB  NDL-R---S-LEKND--H--VLV-----TN--L-DK--.TS--A-T-----S-----KEKPEP

CB  .....VGNYSVSNNGNATCLLATMGLQLNVTEEKVPFIFNINPSTTNFTGSGCHPQTAQLRLNNSQIKYLD F I F
RB  ..VPTPT---TI-----I-----A-----Q-----
MB  TPTPTPT---IR--T-----I-----A-----Q--S-----
HB  .....-T---N--D-----I-QD--ASVI--N--HS--RSH--L--S-T-----V-

CB  A V K S E S H F Y L K E V N V S M Y M A N G S V F S V A N N N L S F W D A P L G S S Y M C N K E Q V V S V S R T F Q I N T F N L K V Q P F N V T K G
RB  ---N-KR-----N--L---A-H-S-----L---A-----
MB  ---N-KR-----Y--L---A-NIS-K-----L---A-----
HB  ---N-NR-----I--LV-----I-----Y-----T---GA-----D-R-----Q-

CB  K Y A T A Q D C S A D E D N F L V P I A V G A A L A G V L A L V L L A Y F I G L K R H H T G Y E Q F
RB  E-S-----G--I-----
MB  Q-S-----G--I-----
HB  --S-----D-----I-----H--A-----

TM

```

382
385
389
380

Figure 13 Aligned amino acid sequences of lgp-B. CB, RB, MB and HB denote hamster, rat, mouse and human lgp-B, respectively. Periods denote gaps introduced to optimize alignments. The amino acid residues identical to hamster lgp-A are denoted by hyphens. Amino acids that are identical in all known lgp-A and lgp-B sequences are denoted by large dots above. The variable domain (VD) and transmembrane domain (TM) are underlined. The amino terminus of the mature proteins is also underlined. (Data are from Fukuda *et al.*, 1988; Noguchi *et al.*, 1989; Cha *et al.*, 1990; Granger *et al.*, 1990; Sawada *et al.*, 1993; and this thesis)

```

lgp-A  ....MAAPGAPRSLLLL.LLAGLAHGASALFVVKDSNGTACI MANFSASFFTIYETGHGSKNSTFELPSS
      |      | | | |      | | | |      | | |      |
lgp-B  MMCFR LSPVSGSGLVLSCLLLGAVQSYAFELNLPDSKAT.CLF AKWKMNFTISYETTTNKT LKTVTISEP

lgp-A  AEVLNSNSSCGRENVSEPI LTIAFGSGYLLTLNFTRNATRYSVQDMYFAYNLSDTQHFLNASNKG IHSVD
      |      | | |      |      | |      |      | | |      | | | |
lgp-B  HNVTYNGSSCGDDQGVAKI.AVQFGSTVSWNVFTTKEESHYVIGSIWL VYNTSDNTTFPGAIPKGSATVI

lgp-A  SSTDIKADINKTYRCLSAIQVHMGNVTVTLSDATI QAYLLNSNFSKEETRCTQD.....GPSPTTVPPS
      |      |      | | |      | | |      | | |      | | |      | | |
lgp-B  SSQSIEIPLDDIFRCNSLLTFKGTGNVVQNYWDIHLQAFVQNGTVSKEEFVCEEDKSVTTVRPIIHTTVPP

lgp-A  PSPPLVPTNPTVIKYNVTGENGTCLLASMA LQMNITYMKKDNMTVTRALNISPNDTA.SGSCSPHVVTLT
      |      | | |      | | | | | | | | |      | | |      | | | |
lgp-B  PTTTPTPLPPKVGNYSVSNGNATCLLATMGLQLNVT.....EEKVPFIFNINPSTTNFTGSCHPQTAQLR

lgp-A  VESKN.SILD LKFGMNGSSSLFFLQEVRLNMTLPDANVSSLMASNQSLRALQATVGNSYKCNTEEHIFVT
      | |      |      | | | |      |      | |      |      | | | | |
lgp-B  LNNSQIKYLD FIFAVKSESH.FYLKEVNVSMY MANGSVF SVANNNL SFW..DAPLGSSYMCNKEQVVSVS

lgp-A  KEFSLNVFSVQVQAFKVESDRFGSVEEQMDGNNMLIPIAVGGALAGLV LIVLIAYLIGRKRSHAGYQTI
      | | |      | | |      |      | | | | | | | | |      | | | | | | |
lgp-B  RTFQINTFNLKVQPFNVTKGKYATAQDCSADEDN FLVPIAVGAALAGVLALVLLAYFIGLKRHHTGYEQF

```

Figure 14. Aligned amino acid sequences of hamster Igp-A and Igp-B. Periods represent gaps that have been introduced to optimize alignments. Vertical lines denote identical residues in both the sequences. The 8 cysteine residues are underlined.

	% identity with hamster lgp-A
mouse lgp-A	79
rat lgp-A	76
human lgp-A	66
chicken lgp-A	46

Table 1. Percentage of amino acid identity between hamster lgp-A and other known lgp-As. Data are from Figure 11. The length of the shorter sequence is used for the calculations. Gaps and gap lengths are not considered.

	% identity with hamster lgp-B
mouse lgp-B	78
rat lgp-B	82
human lgp-B	70

Table 2. Percentage of amino acid identity between hamster lgp-B and other known lgp-Bs. Data are from Figure 12. Calculations are the same as in Table 1.

lgp-A and lgp-B from	% identity
hamster	34
mouse	33
rat	34
human	34

Table 3. Percentage of amino acid identity between lgp-A and lgp-B within the same species. Data for comparing hamster sequences are from Figure 12 (other alignments are not shown). Calculations are the same as in Table 1.

lgp-B from one species is more similar to lgp-B from other species than to lgp-A from the same species. This suggests that the evolutionary divergence of lgp-A and lgp-B preceded the divergence of birds and mammals.

The lgps have been shown to be highly glycosylated (Howe, *et al.*, 1988; Fukuda *et al.*, 1988; Fambrough *et al.* 1988; Granger *et al.*, 1990). The number of potential N-glycosylation sites reported for other species is 16-20, with the highest being 20 for mouse lgp-A. The potential N-glycosylation sites for hamster lgp-A was found to surpass this range, at 23 (with 17 for hamster lgp-B; Howe, *et al.*, 1988; Fukuda *et al.*, 1988; Fambrough *et al.* 1988; Chen *et al.*, 1988; Granger *et al.*, 1990). The mass of the polypeptide backbone of hamster lgp-A is 42 kDa. As evidenced by immunoprecipitation and SDS-PAGE analysis (Granger and Uthayakumar, unpublished observations), the mature hamster lgp-A was indeed found to be larger than the mouse counterpart, correlating with the greater number of N-glycosylation sites. The mass of the polypeptide backbone of hamster lgp-B is 42 kDa; it has the same number of N-linked glycans as mouse lgp-B.

Four other lysosomal membrane proteins, CD68 (110 kDa) /macrosialin (87-115 kDa) (Holness and Simmons, 1993; Holness *et al.*, 1993), CD63 (30-60 kDa; also called ME491 or LIMP I Metzelaar *et al.*, 1991), LGP85 (85 kDa; also called LIMP II; Fujita *et al.*, 1991; Vega *et al.*, 1991; Fujita *et al.*, 1992) and LAP (lysosomal acid phosphatase; Pohlmann *et al.*, 1988) have been identified and their cDNAs cloned and sequenced. CD68 and macrosialin are likely species variants of the same protein in human and mice, and will be referred to as CD68 hereafter. Like the lgps, all of the above proteins localize to lysosomes, and all but LAP have unknown functions. All except LGP85 have a tyrosine residue in their tail, which has been implicated as critical for lysosomal targeting, at about the same position as in lgp-A and lgp-B. The tail of LGP85 is 20-21 amino acids long and does not have a tyrosine residue, indicating that this tyrosine residue is not absolutely required for targeting of membrane proteins to lysosomes. This suggests that the LGP85

might be targeted differently than the other lysosomal membrane proteins characterized thus far.

Whereas lgps (and other lysosomal membrane proteins characterized) are not tissue specific, CD68 is the first example of a lysosomal membrane protein that seems to have a restricted cell-type expression. The luminal domain of CD68 is bipartite with a central hinge-like region separating the two luminal domains, similar to the lgps. The membrane distal domain is unrelated to lgps, but the membrane proximal domain is 26% identical to the membrane proximal domain of human lgp-A (Holness and Simmons, 1993). It contains 4 regularly-spaced cysteines (36-37 residues apart) that align with the cysteines in the equivalent domain of the lgps. The presence of a common domain suggests that it could serve a similar function in both, even though the functions for the two molecules as a whole could be different.

To date, lgps have been characterized only in homeotherms. The lgps have not been identified in poikilotherms or any lower forms of eukaryotes. As a first step to identify lgps from a poikilotherm, we attempted to screen a *Xenopus* cDNA library (a kind gift from Drs. Randall T. Moon and Jan L. Christian, University of Washington, Seattle) using the hamster, mouse and chicken cDNAs as probes. The initial results were encouraging in terms of identification of many positive plaques. When duplicate plaque lifts were done however, none of the positive signals from the two plaque lifts matched, so it was concluded that the signals were background and the attempt was given up temporarily. This failure could be result of hybridization conditions not being ideal for the limited similarity expected between the probes and the target. The above approach could again be tried with some changes in the experimental conditions.

As an alternative to screening libraries to identify lgps from a poikilotherm, we also tried a PCR-based approach to clone and sequence the region encoding the transmembrane domain and cytoplasmic tail of lgp-A from different species. It was anticipated that this

would generate better probes for library screening, in addition to augmenting the existing data on transmembrane and tail sequences of lgps. Two primers were designed to anneal to the sequences 5' (forward) and 3' (reverse) to the regions encoding the transmembrane and cytoplasmic tail of the mouse lgp-A. A control experiment was done using the mouse DNA as a template for PCR amplification, and the procedure was found to give the expected product. When the same was tried with hamster DNA, the sequence of the PCR product was unrelated to any known lgp. This could be due to primer DNA mismatch, especially between the reverse primer and the 3' UT region of the hamster lgp-A gene, which was later found to be significantly different. The next step was to design a primer that would anneal to the region encoding a conserved part of the protein, the cytoplasmic tail, but this experiment remains to be completed.

Antibodies to the lgps

Our ability to produce antibodies against hamster lgp-A and lgp-B by immunizing mice with crude membrane glycoproteins shows that the lgps are highly immunogenic, consistent with earlier observations (Hughes and August, 1982; Lewis *et al.*, 1985; Lippincott-Schwartz and Fambrough, 1986; Granger *et al.*, 1990). Although it was demonstrated that a polyclonal antibody against mouse lgp-A reacted with hamster lgp-A (Do *et al.*, 1990), antibodies against lgp-A have never been shown to react with lgp-B; similarly, antibodies against lgp-B have not been shown to react with lgp-A. Indeed, our anti-hamster lgp-A antibody (UH1) does not recognize hamster lgp-B, and our anti-hamster lgp-B antibody (UH3) does not recognizes hamster lgp-A. Neither UH1 nor UH3 recognizes mouse lgp-A or lgp-B. The absence of cross-reactivity of the lgp antibodies is surprising, especially in view of the high sequence similarity between the lgps, but very useful for distinguishing the different species and types of lgps.

Heterologous Expression of Hamster Igps

Hamster lgp-A and lgp-B cDNAs were cloned into eukaryotic expression vectors and used to transfect mouse NIH-3T3 cells; the resultant cell lines were called 3T3HA and 3T3HB. Immunofluorescence was performed on 3T3HA and 3T3HB cells using anti-hamster lgp-A or anti-hamster lgp-B antibodies, and the cells were found to express the appropriate hamster lgps, which were localized to lysosomes (Fig. 11). In control experiments with untransfected NIH-3T3 cells, no fluorescence could be detected. This is a clear indication that the hamster cDNA constructs are accurate and expressible.

Since both lgp-A and lgp-B localize to lysosomes, it is impossible to distinguish lgp-A antibodies from lgp-B antibodies by fluorescence pattern alone. For this reason, hamster cells lines which express both lgp-A and lgp-B are of no use for differentiating the two antibodies. We used the 3T3HA cells, which expresses hamster lgp-A (and not lgp-B), to identify the UH1 hybridoma as a secretor of hamster lgp-A antibody. Similarly, 3T3HB cells that express hamster lgp-B (and not lgp-A) were useful in identifying UH3, the hybridoma secreting hamster lgp-B antibody.

The known 5' UT region of the lgp-A mRNA has an unusually high GC content of ~ 80% (compared to a GC content of around 65% for lgp-B). To explore the possible regulatory role of the 5' UT segment of the hamster lgp-A mRNA, we deleted the 5' UT sequences from the cDNA, preserving the five bases (consensus sequence among the known lgp sequences) 5' to the start codon. We also changed the G at the -3 position (relative to the start codon) to an A, which is more common at that position (Kozak, 1989). The 5' deleted lgp-A was transfected in to NIH-3T3 cells, and the resultant stable transformants were called 3T3HA5'Δ. The expression level of hamster lgp-A in 3T3HA5'Δ cells was very low compared to the expression level of the same protein in 3T3HA cells. This suggests that a potential role of the 5' UT region in the regulation of

expression, but additional experiments must be done to verify this conclusion and determine likely mechanisms.

5' Sequences of lgp-B cDNAs

In the past, it has been observed that the rat and mouse lgp-B cDNA clones were much less abundant in cDNA libraries than the corresponding lgp-A clones (Granger *et al.*, 1990). The same was also observed with the hamster lgp-B cDNA in this study. The observed ratio between hamster lgp-A and lgp-B clones was 10 : 1.

Previously obtained mouse and rat lgp-B cDNAs lacked the 5' UT region and the coding region for the first 20-40 amino acids (Granger *et al.*, 1990). The hamster lgp-B cDNA obtained in this study also lacked the 5' end; it did not have any of the 5' UT region or coding region corresponding to the first 28 amino acids.

To obtain the 5' end of the hamster lgp-B cDNA, we used a modified RACE (rapid amplification of cDNA ends) method (Frohman *et al.*, 1988). The RACE method originally involved the addition of a homopolymeric tail of dA's (or dG's) to the 3' end of the first-strand cDNA (synthesized by reverse transcription of mRNA using a gene-specific primer), by terminal deoxyribonucleotide transferase (TdT). This reaction, however, was difficult to control. The addition of too few or too many nucleotides diminished the effectiveness of the tail as a PCR template, since primers designed to anneal to the tail contain a homopolymeric sequence at the 3' end, and often annealed to sites within the cDNA. This often generated a high background of non-specific products. An improved method developed later, termed SLIC (single strand ligation to single stranded cDNA), involved ligation of a single stranded oligonucleotide anchor directly to the 3' end of the first-strand cDNA using T4 RNA ligase in the presence of hexamine chloride (Dumas *et al.*, 1991). The 5'-AmpliFINDER RACE procedure (Siebert and Apte, 1993), based on

SLIC, was followed to obtain the 5' sequences of the hamster lgp-B cDNA using a 5'-AmpliFINDER RACE kit (Clontech) for this purpose.

As anticipated (by comparison with mouse lgp-B cDNA sequence), a 500 bp RACE product was generated, which was mapped, cloned and sequenced (see Results). No obvious explanation for the absence (or under-representation) of the 5' region in the cDNA library was evident. It is unlikely that the absence is specific for this cDNA library, since it has also been observed in cDNA libraries from other species (Granger *et al.*, 1990).

The cloned hamster lgp-A cDNA 5' UT region is almost certainly incomplete. To verify whether the 5' UT region of the cDNA has an effect on expression, the full 5' UT region would have to be obtained, which should be possible using the 5'-AmplifinderRACE kit. It remains to be determined whether the observed higher expression of the lgp-B cDNA in NIH-3T3 cells (relative to the lgp-A constructs) is due to its complete 5' UT region, its incomplete 3' UT region, the polyadenylation segment in the expression vector, or just chance.

Future Directions

With the available antibodies to hamster and mouse lgps, targeting and transport of these lgps can now be better studied. CHO cells have been stably transformed with many mutant and normal mouse lgp-A cDNAs for studying the targeting and transport of lgps (Granger *et al.*, 1989, 1992; Harter and Mellman, 1992). In these cells, both the foreign mouse lgp-A and endogenous hamster lgp-A (and lgp-B) are synthesized and transported to lysosomes. The transport machinery is likely to be the same for both mouse and hamster lgp-A, but we do not know whether the hamster lgp-A influences the transport of mouse lgp-A, or vice versa. Potential interference can now be studied by immunofluorescence of

cells over-expressing mouse lgp-A, using both mouse and hamster specific lgp-A antibodies.

Some mouse lgp-A has been found to be transported to the plasma membrane in CHO cells over-expressing mouse lgp-A (Granger *et al.*, 1989, 1992; Harter and Mellman, 1992). We do not know whether the hamster lgp-A is also transported to the plasma membrane in these cells. The proportion of each (mouse and hamster) lgp-A that is transported to the plasma membrane can now be studied by immunofluorescence using our specific monoclonal antibodies. This may shed light on the mechanism: Perhaps there is always a constant proportion of each lgp on the surface (but in amounts that are undetectable unless over-expressed), or over-expression saturates a transport pathway to lysosomes and the excess spills over into a default pathway to the surface.

Using the hamster and mouse lgp cDNA sequence information, oligonucleotide probes specific for hamster or mouse mRNAs can now be designed. Such probes can be used for relating the message and protein levels in transfected cells. This information will be useful for studying the regulation of lgp message expression.

REFERENCES CITED

- Arterburn, L.M., Earles, B.J. and August, J.T. (1990) The disulfide structure of mouse lysosome-associated membrane protein 1. *J. Biol. Chem.* **265**, 7419-7423.
- Barriocanal, J.G., Bonifacino, J.S., Yuan, L. and Sandoval, I.V. (1986) Biosynthesis, glycosylation, movement through the Golgi system, and transport to lysosomes by an N-linked carbohydrate-independent mechanism of three lysosomal integral membrane proteins. *J. Biol. Chem.* **261**, 16755-16763.
- Braun, M., Waheed, A. and von Figura, K. (1989) Lysosomal acid phosphatase is transported to lysosomes via the cell surface. *EMBO J.* **8**, 3497-3506.
- Carlsson, S.R. and Fukuda, M. (1989) Structure of human lysosomal membrane glycoprotein 1. *J. Biol. Chem.* **264**, 20526-20531.
- Carlsson, S.R. and Fukuda, M. (1992) The lysosomal membrane glycoprotein lamp-1 is transported to lysosomes by two alternative pathways. *Arch. Biochem.* **296**, 630-639.
- Cha, Y., Holland, S.M. and August, J.T. (1990) The cDNA sequence of mouse LAMP-2. Evidence for two classes of lysosomal membrane glycoproteins. *J. Biol. Chem.* **265**, 5008-5013.
- Chen, C.A. and Okayama, H. (1988) Calcium phosphate-mediated gene transfer: A highly efficient transfection system for stably transforming cells with plasmid DNA. *Biotechniques* **6**, 632-638.
- Chen, J.W., Cha, Y., Yuksel, K.U., Gracy, R.W. and August, J.T. (1988) Isolation and sequencing of a cDNA clone encoding lysosomal membrane glycoprotein mouse LAMP-1. *J. Biol. Chem.* **263**, 8754-8758.
- Chen, J.W., Murphy, T.L., Willingham, M.C., Pastran, I. and August, J.T. (1985) Identification of two lysosomal membrane glycoproteins. *J. Cell Biol.* **101**, 85-95.
- Chomczynski, P. and N. Sacchi. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**, 156-159.
- D'Souza, M.P. and August, J.T. (1986) A kinetic analysis of biosynthesis and localization of a lysosome-associated membrane glycoprotein. *Arch. Biochem. Biophys.* **249**, 522-532.
- Devereux, J., Haeberli, H. and Smithies, O. (1983) A complete set of sequence analysis programs for the VAX. *Nuc. Acids Res.* **12**, 387-396.

- Do, K.-Y., Smith, D.F. and Cummings, R.D. (1990) LAMP-1 in CHO cells is a primary carrier of poly-n-acetyllactosamine chains and is bound preferentially by a mammalian S-type lectin. *Biochem. Biophys. Res. Comm.* **173**, 1123-1128.
- Dumas, J.B., Edwards, M., Delort, J. and Mallet, J. (1991) Oligodeoxyribonucleotide ligation to single-stranded cDNAs: a new tool for cloning 5' ends of mRNAs and for constructing cDNA libraries by *in vitro* amplification. *Nuc. Acids Res.* **19**, 5227-5232.
- Fambrough, D.M., Takeyasu, K., Lippincott-Schwarz, J., Siegel, N.R. and Somerville, D. (1988) Structure of LEP100, a glycoprotein that shuttles between lysosomes and the plasma membrane, deduced from the nucleotide sequence of the encoding cDNA. *J. Cell Biol.* **106**, 61-67.
- Flotte, T.J., Springer, T.A. and Thorbecke, G.J. (1983) Dendritic cell and macrophage staining by monoclonal antibodies in tissue sections and epidermal sheets. *Am. J. Pathol.* **111**, 112-124.
- Frohman, M.A., Dush, M.K. and Martin, G.R. (1988) Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85**, 8998-9002.
- Fujita, H., Ezaki, J., Noguchi, Y., Kono, A., Himeno, M. and Kato, K. (1991) Isolation and sequencing of a cDNA clone encoding 85 kDa sialoglycoprotein in rat liver lysosomal membranes. *Biochem. Biophys. Res. Comm.* **178**, 444-452.
- Fujita, H., Tadat, Y., Kono, A., Tanaka, Y., Takahashi, T., Himeno, M. and Kato, K. (1992) Isolation and sequencing of a cDNA clone encoding the 85 kDa human lysosomal sialoglycoprotein (hLGP85) in human metastatic pancreas islet tumor cells. *Biochem. Biophys. Res. Comm.* **184**, 604-611.
- Fukuda, M. (1991) Lysosomal membrane glycoproteins. *J. Biol. Chem.* **266**, 21327-21330.
- Fukuda, M., Viitala, J., Matteson, J. and Carlsson, S.R. (1988) Cloning of cDNAs encoding human lysosomal membrane glycoproteins h-lamp-1 and h-lamp-2. Comparison of their deduced amino acid sequences. *J. Biol. Chem.* **263**, 18920-18928.
- Gorman, C.M. and Howard, B.H. (1983) Expression of recombinant plasmids in mammalian cells is enhanced by sodium butyrate. *Nuc. Acids Res.* **11**, 7631-7648.
- Granger, B.L., Green, S.A., Gabel, C.A., Howe, C.L., Mellman, I. and Helenius, A. (1990) Characterization and cloning of lgp 110, a lysosomal membrane glycoprotein from mouse and rat cells. *J. Biol. Chem.* **265**, 12036-12043.
- Granger, B.L., Howe, C.L., Hull, M., Lombillo, S.A., Harter, C., Mellman, I. and Helenius, A. (1989) Cloning and expression of lysosomal membrane glycoproteins from mouse and rat cells. *J. Cell Biol.* **109**, 114a.

- Granger, B.L., Plutner, H., Green, S.A., Mellman, I. and Helenius, A. (1985) Defining the lysosome membrane: Characterization of specific glycoproteins. *J. Cell Biol.* **101**, 54a.
- Granger, B.L., Warwood, S.J., Uthayakumar, S. and Wimer, S. (1992) Transport of a lysosomal membrane glycoprotein. *Mol. Biol. Cell* **3S**, 312a.
- Green, S.A., Zimmer, K.-P., Griffiths, G. and Mellman, I. (1987) Kinetics of intracellular transport and sorting of lysosomal membrane and plasma membrane proteins. *J. Cell Biol.* **105**, 1227-1240.
- Griffiths, G. and Simons, K. (1986) The *trans* Golgi network: Sorting at the exit site of the Golgi complex. *Science* **234**, 438-443.
- Guarnieri, F.G., Arterburn, L.M., Penno, M.B., Cha, Y. and August, J.T. (1993) The motif Tyr-X-X- hydrophobic residue mediates lysosomal membrane targeting of lysosome-associated membrane protein. *J. Biol. Chem.* **268**, 1941-1946.
- Harter, C. and Mellman, I. (1992) Transport of the lysosomal membrane glycoprotein lgp120 (lpg-A) to lysosomes does not require appearance on the plasma membrane. *J. Cell Biol.* **117**, 311-325.
- Heffernan, M., Yousefi, S. and Dennis, J.W. (1989) Molecular characterization of P2B/LAMP-1, a major protein target of a metastasis-associated oligosaccharide structure. *Cancer Research* **49**, 6077-6084.
- Himeno, M., Noguchi, Y., Sasaki, H., Tanaka, Y., Furuno, K., Kono, A., Sakaki, Y. and Kato, K. (1989) Isolation and sequencing of a cDNA clone encoding 107 kDa sialoglycoprotein in rat liver lysosomal membranes. *FEBS Lett.* **244**, 351-356.
- Ho, M.-K. and Springer, T.A. (1983) Tissue distribution, structural characterization, and biosynthesis of Mac-3, a macrophage surface glycoprotein exhibiting molecular weight heterogeneity. *J. Biol. Chem.* **258**, 636-642.
- Holness, C.L. and Simmons D.L. (1993) Molecular cloning of CD68, a human macrophage marker related to lysosomal glycoproteins. *Blood* **6**, 1607-1613.
- Holness, C.L., da Silva, R.P., Fawcett, J., Gordon, S. and Simmons D.L. (1993) Macrosialin, a mouse macrophage-restricted glycoprotein, is a member of the lamp/lgp family. *J. Biol. Chem.* **268**, 9661-9666.
- Hopkins C.R. (1992) Selective membrane protein trafficking: Vectorial flow and filter. *TIBS* **17**, 27-32.
- Howe, C.L., Granger, B.L., Hull, M., Green, S.A., Gabel, C.A., Helenius, A. and Mellman, I. (1988) Derived protein sequence, oligosaccharides, and membrane insertion of the 120-kDa lysosomal membrane glycoprotein (lgp120): Identification of a highly conserved family of lysosomal membrane glycoproteins. *Proc. Natl. Acad. Sci. USA* **85**, 7577-7581.

- Hughes, E.N. and August, J.T. (1982) Murine cell surface glycoproteins. Identification, purification, and characterization of a major glycosylated component of 110,000 daltons by a monoclonal antibody. *J. Biol. Chem.* **257**, 3970-3977.
- Hurtley, S.M. and Helenius, A. (1989) Protein oligomerization in the endoplasmic reticulum. *Annu. Rev. Cell Biol.* **5**, 277-307.
- Inoue, H., Nojima, H. and Okayama, H. (1990) High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**, 23-28.
- Kelly, R.B. (1985) Pathways of protein secretion in eukaryotes. *Science* **230**, 25-32.
- Kornfeld, R. and Kornfeld, S. (1985) Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* **54**, 631-664.
- Kornfeld, S. and Mellman, I. (1989) The biogenesis of lysosomes. *Annu. Rev. Cell Biol.* **5**, 483-525.
- Kozak, M. (1989) The scanning model for translation: An update. *J. Cell Biol.* **108**, 229-241.
- Larsen, E., Palabrica, T., Sajer, S., Gilbert, G.E., Wagner, D.D., Furie, B.C. and Furie, B. (1990) PADGEM-dependent adhesion of platelets to monocytes and neutrophils is mediated by a lineage specific carbohydrate, LNF III (CD15). *Cell* **63**, 467-474.
- Lewis, V., Green, S.A., Marsh, M., Vihko, P., Helenius, A. and Mellman, I. (1985) Glycoproteins of the lysosomal membrane. *J. Cell Biol.* **100**, 1839-1847.
- Lippincott-Schwartz, J. and Fambrough, D.M. (1986) Lysosomal membrane dynamics: Structure and interorganellar movement of a major lysosomal membrane glycoprotein. *J. Cell Biol.* **102**, 1593-1605.
- Lippincott-Schwartz, J. and Fambrough, D.M. (1987) Cycling of the integral membrane glycoprotein, LEP100, between plasma membrane and lysosomes: Kinetic and morphological analysis. *Cell* **49**, 669-677.
- Lowe, J.B., Stoolman, L.M., Nair, R.P., Larsen, R.D., Berhend, T.L. and Marks, R.M. (1990) ELAM-1-Dependent cell adhesion to vascular endothelium determined by transfected human fucosyltransferase cDNA. *Cell* **63**, 475-484.
- Marsh, M., Schmid, S., Kern, H., Harms, E., Male, P., Mellman, I. and Helenius, A. (1987) Rapid analytical and preparative isolation of functional endosomes by free flow electrophoresis. *J. Cell Biol.* **104**, 875-886.
- Mathews, P.M., Martinie, J.B. and Fambrough, D.M. (1992) The pathway and targeting signal for delivery of the integral membrane glycoprotein LEP100 to lysosomes. *J. Cell Biol.* **118**, 1027-1040.

- Mattei, M., Matterson, J., Chen, J.W., Williams, M.A. and Fukuda, M. (1990) Two human lysosomal membrane glycoproteins, h-lamp-1 and h-lamp-2 are encoded by genes localized to chromosome 13q34 and chromosome Xq24-25, respectively. *J. Biol. Chem.* **265**, 7548-7551.
- McLean, I.W. and Nakane, P.K. (1974) Periodate-lysine-paraformaldehyde fixative. A new fixative for immunoelectron microscopy. *J. Histochem. Cytochem.* **22**, 1077-1083.
- Nabi, I.R. Le Bivic, A., Fambrough, D. and Rodriguez-Boulon, E. (1991) An endogenous MDCK lysosomal membrane glycoprotein is targeted basolaterally before delivery to lysosomes. *J. Cell Biol.* **115**, 1573-1584.
- Neiss, W.F. (1984) A coat of glycoconjugates on the inner surface of the lysosomal membrane in the rat kidney. *Histochemistry* **80**, 603-608.
- Noguchi, Y., Himeno, M., Sasaki, H., Tanaka, Y., Kono, A., Sakaki, Y. and Kato, K. (1989) Isolation and sequencing of a cDNA clone encoding 96 kDa sialoglycoprotein in rat liver lysosomal membranes. *Biochem. Biophys. Res. Comm.* **164**, 1113-1120.
- Orci, L., Ravazzola, M., Amherdt, M., Perrelet, A., Powell, S.K., Quinn, D.L. and Moore, H.-P.H. (1987) The *trans*-most cisternae of the Golgi complex: A compartment for sorting of secretory and plasma membrane proteins. *Cell* **51**, 1039-1051.
- Pearson, W.R. and Lipman, D.J. (1988) Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci.* **85**, 2444-2448.
- Pfeffer, S.R. and Rothman, J.E. (1987) Biosynthetic protein transport and sorting by the endoplasmic reticulum and Golgi. *Annu. Rev. Biochem.* **56**, 829-852.
- Phillips, M.L., Nudelman, E., Gaeta, F.C.A., Perez, M., Singhal, A.K., Hakomori, S.-I., and Paulson, J.C. (1990) ELAM-1 mediates cell adhesion by recognition of a carbohydrate ligand, sialyl-Le^x. *Science* **250**, 1130-1132.
- Saitoh, O., Wang, W.-C., Lotan, R. and Fukuda, M. (1992) Differential glycosylation and cell surface expression of lysosomal membrane glycoproteins in sublines of a human colon cancer exhibiting distinct metastatic potentials. *J. Biol. Chem.* **267**, 5700-5711.
- Sambrook, J., Fritsch, E.F., Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. 2nd edition. Cold Spring Harbor Laboratory, New York.
- Sandoval, I.V., Chen, J.W., Yuan, L. and August, J.T. (1989) Lysosomal integral membrane glycoproteins are expressed at high levels in the inclusion bodies of I-cells disease fibroblasts. *Arch. Biochem. Biophys.* **271**, 157-167.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* **74**, 5463.

- Sawada, R., Jardine, K.A. and Fukuda, M. (1993) The genes of major lysosomal membrane glycoproteins Lamp-1 and Lamp-2. *J. Biol. Chem.* **268**, 9014-9022.
- Schauer, R. (1985) Sialic acids and their role as biological masks. *TIBS* **10**, 357-360.
- Siebert, P.D. and Apte, A. (1993) A new method for amplification of 5'-cDNA ends. *Clontechiques* **8**, 1-4.
- Trowbridge, I. (1991) Endocytosis and signals for internalization. *Curr. Op. Cell Biol.* **3**, 634-641.
- Vega, M.A., Segui-Real, B., Garcia, J.A., Cales, C., Rodriguez, F., Vanderkerckhove, J. and Sandoval, I.V. (1991) Cloning, sequencing, and expression of a cDNA encoding rat LIMP II, a novel 74-kDa lysosomal membrane protein related to the surface adhesion protein CD36. *J. Biol. Chem.* **266**, 16818-16824.
- Verner, K. and Schatz, G. (1988) Protein translocation across membranes. *Science* **241**, 1307-1313.
- Viitala, J., Carsson, S.R., Siebert, P.D. and Fukuda, M. (1988) Molecular cloning of cDNAs encoding lamp A, a human lysosomal membrane glycoprotein with apparent Mr = 120,000. *Proc. Natl. Acad. Sci. USA* **85**, 3743-3747.
- von Figura, K. and Hasilik, A. (1986) Lysosomal enzymes and their receptors. *Annu. Rev. Biochem.* **55**, 167-193.
- Waheed, A., Gottschalk, S., Hille, A., Drentler, C., Pohlmann, R., Bräulke, T., Hauser, H., Geuze, H. and von Figura, K. (1988) Human lysosomal acid phosphatase is transported as a transmembrane protein to lysosomes in transfected baby hamster kidney cells. *EMBO J.* **7**, 2351-2358.
- Williams, M.A. and Fukuda, M. (1990) Accumulation of membrane glycoproteins in lysosomes requires a tyrosine residue at a particular position in the cytoplasmic tail. *J. Biol. Chem.* **265**, 955-966.
- Zot, A.S. and Fambrough, D.M., (1990) Structure of a gene for a lysosomal membrane glycoprotein (LEP100). *J. Biol. Chem.* **265**, 20988-20995.

APPENDIX

Composition of Solutions20X SSC

3 M NaCl
0.3 M sodium citrate
NaOH to pH 7.0

20X SSPE

3 M NaCl
25 mM EDTA
200 mM NaH₂PO₄
NaOH to pH 7.4

Hybridization solution

5X NET
1% SDS
0.5% NFDM
0.1% PVP-40

20X NET

3 M NaCl
20 mM EDTA
300 mM Tris
HCl to pH 8

10X nick-translation buffer

1 M Tris-Cl pH 7.5
100 mM MgCl₂

High salt buffer

1.0 M NaCl
20 mM Tris-Cl pH 7.5
1.0 mM EDTA

Medium salt buffer

0.4 M NaCl
20 mM Tris-Cl pH 7.5
1.0 mM EDTA

Tris/CDTA/dextrose

25 mM Tris
10 mM CDTA
50 mM dextrose
NaOH to pH 8.0

NaOH/SDS

0.2 M NaOH
1% SDS

KOAc/formic acid

3 M potassium acetate
1.8 M formic acid

NaOAc/MOPS

100 mM sodium acetate
50 mM MOPS
NaOH to pH 8.0

LiCl/MOPS

5 M LiCl
50 mM MOPS

NaOAc/MOPS/CDTA

100 mM sodium acetate
50 mM MOPS
1 mM CDTA
NaOH to pH 8.0

Tris/CDTA

10 mM Tris
1 mM CDTA
HCl to pH 8.0

2X BBS

50 mM BES
280 mM NaCl
1.5 mM Na₂HPO₄
NaOH to pH 6.95

SM

100 mM NaCl
50 mM Tris-Cl pH 7.5
10 mM MgSO₄

LB

1% Bacto tryptone
0.5% Bacto yeast extract
1% NaCl
NaOH to pH ~7.4

Extraction buffer

120 mM KCl
10 mM PIPES
5 mM EGTA
30 mM NaOH
10 mM MgCl₂
pH 6.9-7.3

Abbreviations:

BES	N, N-bis (2-hydroxyethyl)-2-amino-ethanesulfonic acid
CDTA	<i>trans</i> -1, 2-diaminocyclohexane-N, N, N', N'-tetra acetic acid
MOPS	3-[N-morpholino]propanesulfonic acid
NFDM	non-fat dry milk
PVP-40	polyvinylpyrrolidone-40
PIPES	piperazine-N, N'-bis[2-ethane-sulfonic acid]

MONTANA STATE UNIVERSITY LIBRARIES



3 1762 10203133 1