Identification of Sites of Mannose 6-Phosphorylation on Lysosomal Proteins*

David E. Sleat‡§¶, Haiyan Zheng‡, Meiqian Qian‡§, and Peter Lobel‡§

Most newly synthesized soluble lysosomal proteins contain mannose 6-phosphate (Man-6-P), a specific carbohydrate modification that is recognized by Man-6-P receptors (MPRs) that direct targeting to the lysosome. A number of proteomic studies have focused on lysosomal proteins, exploiting the fact that Man-6-P-containing forms can be purified by affinity chromatography on immobilized MPRs. These studies have identified many known lysosomal proteins as well as many proteins not previously classified as lysosomal. The latter are of considerable biological interest with potential implications for lysosomal function and as candidates for lysosomal storage diseases of unknown etiology. However, a significant problem in interpreting the biological relevance of such proteins has been in distinguishing true Man-6-P glycoproteins from simple contaminants and from proteins associated with true Man-6-P glycoproteins (e.g. protease inhibitors and lectins). In this report, we describe a mass spectrometric approach to the verification of Man-6-phosphorylation based upon LC-MS of MPR-purified proteolytic glycopeptides. This provided a useful tool in validating novel MPR-purified proteins as true Man-6-P glycoproteins and also allowed identification of low abundance components not observed in the analysis of the total Man-6-P glycoprotein mixture. In addition, this approach allowed the global mapping of 99 Man-6-phosphorylation sites from 44 known lysosomal proteins purified from mouse and human brain. This information is likely to provide useful insights into protein determinants for this modification and may be of significant value in protein engineering approaches designed to optimize protein delivery to the lysosome in therapeutic applications such as gene and enzyme replacement therapies. Molecular & Cellular Proteomics 5:686–701, 2006.

The greatest challenge in “whole cell” proteomic approaches to understanding cellular function is the fact that protein identification techniques are greatly limited by both the large number of individual proteins and isoforms and the wide range of their abundance. Multidimensional sample fractionation can increase proteome coverage in such global approaches, but complexity still remains a significant problem. One approach that addresses this challenge is the proteomic characterization of subcellular organelles (for a review, see Ref. 1). Although by definition more narrow than global studies, the great reduction in sample complexity allows for higher relative proteome coverage. In addition, proteomic analysis of organelles provides useful insights into the biological function of the individual proteins and the organelle in question.

One cellular organelle that has been the focus of a number of proteomic studies is the lysosome, a membrane-delimited compartment that contains over 50 soluble hydrolytic enzymes and associated accessory proteins that play a major role in the catabolism of macromolecules within the cell (2) and are important in such processes as endocytosis, autophagy, and antigen presentation. Lysosomes and lysosomal proteins are of considerable biomedical importance, and defects in lysosomal proteins result in over 40 human genetic diseases, the lysosomal storage disorders. In addition, alterations in the lysosomal system have been implicated with more widespread diseases including cancer, Alzheimer disease, rheumatoid arthritis, and atherosclerosis.

A number of proteomic studies (3–9) rely upon the fact that it is possible to distinguish soluble lysosomal proteins from other classes of proteins based on a specific post-translational modification that is required for lysosomal targeting. Newly synthesized lysosomal proteins differ from other classes of glycoproteins in that they bear one or more N-linked oligosaccharides containing mannose 6-phosphate (Man-6-P).1 This carbohydrate modification is recognized within the cell by two Man-6-P receptors (MPRs) that direct

1 The abbreviations used are: Man-6-P, mannose 6-phosphate; Man, mannose; MPR, Man-6-P receptor; sCl-MPR, soluble cation-independent MPR; Endo H, endoglycosidase H; PNGase F, peptide-N-glycosidase F; N;+203, 203-Da increment resulting from conversion of asparagine to aspartic acid after oligosaccharide cleavage using Endo H; N+1, 1-Da increment resulting from conversion of asparagine to aspartic acid after oligosaccharide removal using PNGase F; GM2, GalNAcβ1→4Gal(3→2sNeuAc) β1→4Glcβ1→1Cer; ER, endoplasmic reticulum; AOAH, acyloxyacyl hydrolase; BTD, bicitinidase; CALU, calumenin; CECR1, cat eye syndrome critical region 1; CREG, cellular repressor of E1A-stimulated gene expression; PO-FUT, protein O-fucosyltransferase; EPD1, ependymin-related protein (up-regulated in colon cancer I); MAG, myelin-associated glycoprotein; RNASET2, ribonuclease T2; SPARC1, SPARC (secreted protein, acidic, and rich in cysteine)-like protein 1; SUMF, sulfatase-modifying factor precursor.
the vesicular trafficking of the newly synthesized lysosomal proteins to a prelysosomal compartment. The Man-6-P marker is removed in most cell types but in others, e.g. neurons, it is retained to some extent (3, 10). In affinity purification protocols, immobilized purified MPRs can be used to bind Man-6-P glycoproteins that can be specifically eluted using free Man-6-P.

As expected, many of the proteins isolated by MPR affinity chromatography are known lysosomal proteins. However, this approach has also led to the tentative lysosomal classification of both novel proteins and previously characterized cellular proteins that were not originally assigned lysosomal function. These proteins are of considerable interest as potential novel components of the lysosomal system, and their identification as such will clearly extend understanding of this organelle. In addition, novel lysosomal proteins may represent candidates for defective genes in lysosomal storage diseases of unknown etiology, and this is an approach that has led to the identification of two human disease genes (11, 12). However, a significant difficulty in interpretation of the results of MPR affinity chromatography lies in distinguishing true Man-6-P glycoproteins from proteins that do not contain this modification. Some contaminants simply represent highly abundant proteins that, although greatly diminished, are still detected using highly sensitive analytical methods. Other contaminants represent non-lysosomal proteins that associate with and copurify with true Man-6-P glycoproteins. These include protease inhibitors bound to lysosomal proteases, lectins that recognize the carbohydrate moieties of lysosomal proteins, and molecular chaperones.

Bioinformatic approaches can help in the classification of lysosomal proteins and identification of these false positives (7). Some MPR-purified proteins, for example, can be eliminated because they lack potential N-linked glycosylation sites (e.g. protease inhibitors cystatin B and cystatin C). Others can be considered unlikely because they are not predicted to contain signal sequences (e.g. the cytoplasmic lectin F-box only protein 2). However, these sorts of bioinformatic approaches are imperfect and are likely to generate false negatives. For example, acid sphingomyelinase, a known lysosomal Man-6-P glycoprotein, is not predicted to contain a signal sequence using SignalP 3.0 (13). In addition, although such approaches are useful in eliminating some false positives, they are of limited value as many non-lysosomal proteins contain both potential N-linked glycosylation sites and predicted signal sequences (7). Thus, a direct method to verify the presence of the Man-6-P modification would be extremely valuable in these studies.

Previously we attempted to directly validate the presence of Man-6-P by fractionation of the proteins purified by MPR affinity chromatography by two-dimensional gel electrophoresis and detection of proteins containing Man-6-P using a radiolabeled MPR probe (12, 7). However, a significant technical problem with this approach was that many spots were found to contain more than one protein species, thus the source of a positive signal was frequently ambiguous. In addition, a conceptual problem with this approach is that many lysosomal proteins consist of multiple subunits or chains of which some contain Man-6-P and some do not. Another approach has been to express and purify candidate proteins containing His tags and demonstrate Man-6-P-dependent binding to MPRs in vitro and Man-6-P-dependent internalization and lysosomal targeting in cell lines (8), but again this is an indirect approach that cannot distinguish between true Man-6-P glycoproteins and potential contaminants that have been identified by virtue of interactions with true Man-6-P glycoproteins. In addition, none of these approaches provides any information with respect to the actual sites of the Man-6-P modification, and each must be pursued on an individual, rather than global, basis.

A recent focus of interest in the proteomic field is the use of glycopeptide capture to identify glycosylated peptides from N-linked glycoproteins using either affinity capture with lectins (14, 15) or chemical capture (16). In this report, we describe a sensitive and unambiguous method to globally identify peptides containing Man-6-P, and we used this approach to validate candidates for novel lysosomal proteins. In addition, we provide a database for Man-6-phosphorylation sites in known lysosomal proteins that is likely to prove valuable in terms of both the understanding of lysosomal cell biology and in bioengineering of lysosomal proteins for therapeutic applications.

**EXPERIMENTAL PROCEDURES**

**Purification of Human and Mouse Brain Man-6-P Glycoproteins—** Man-6-P glycoproteins were purified from normal human brain autopsy samples and from brain samples collected from mixed strain BALB/c/129SvEv/C57BL6 mice as described previously (7).

**LC-MS/MS—** Reduction and alkylation of MPR-purified proteins were as described previously (7), and digestion with trypsin (Promega, Madison, WI), Glu-C (Wako Pure Chemical Industries, Osaka, Japan), and Lys-C (Roche Applied Science) was conducted overnight at 37 °C in 50 mM ammonium bicarbonate. Digestion with Glu-C was also performed in 100 mM sodium phosphate, pH 7.8. Proteolytic digests of MPR-purified proteins were loaded onto a 300-nm × 5-cm Pepmap C18 trap column (Dionex/LC Packings, Amsterdam, the Netherlands) and desalted for 5 min using 0.1% formic acid at a flow rate of 25 µl/min. Peptides were then back-flushed onto a fritless nanoscale column (75 µm × 15 cm) packed with Poros 10, R2 (Applied Biosystems, Foster City, CA) and eluted using a linear 2–45% gradient of solvent B (0.1% formic acid in acetonitrile) in solvent A (0.1% formic acid) over 40 min (tryptic digest of gel slices) or 150 min (solution tryptic digest of affinity-purified proteins) with a flow rate of 300 nl/min using an Ultimate nano-LC system (Dionex/LC Packings). Peptides were analyzed by ESI-MS/MS using an LTQ ion trap mass spectrometer (ThermoFinnigan, San Jose, CA) equipped with a nanospray source (Proxeon Biosystems, Odense, Denmark). Each MS scan was followed by MS/MS of the 15 most intense ions with a dynamic exclusion of 2 min and a repeat count of two. For analysis of MPR-purified Man-6-P glycopeptides, the sample was first loaded onto the trap column and washed with 0.1% formic acid for 40 min at a flow rate of 25 µl/min before back-flushing to the nanoscale column. Peptides were eluted using a gradient of 2–45% B over 50 min. Each MS scan (m/z = 400–1800) was followed by
Mannose 6-Phosphorylation Sites

subsequent zoom scans and MS/MS scans of the four most abundant ions with a dynamic exclusion of 1 min and repeat count of two. Peak lists were generated from raw files using TurboSequest. Parameters were: peptide mass range of 500–5000, threshold intensity of 1000, precursor mass tolerance of 1.4 m/z, minimum ion count of 50, and automatic charge state determination.

**Database Searching**—Searching of the human (ENSEMBL 28.35a.1, National Center for Biotechnology Information (NCBI) 35, May 2005) and mouse (ENSEMBL, NCBIm34, May 2005) was conducted using a local implementation of X! Tandem version 2005.02.1, which is open source software for protein identification using MS/MS data (17). For the unfractonated proteolytic digest mixtures, LTQ data was searched using a precursor ion mass error of +4 and −0.5 Da and a fragment mass error of 0.4 Da. Errors in assignment of monoisotopic mass were not permitted. Cysteine carbamidomethylation was specified as a complete modification, and methionine oxidation was a permitted variable modification during development of the preliminary model with one missed cleavage site allowed. Methionine oxidation and deamidation at asparagine and glutamine residues were allowed during model refinement. Database searching for deglycosylated peptides was as described for the unfractonated proteolytic digests except that mass increments on asparagine as a result of deglycosylation with PNGase F and Endo H, respectively.

**Criteria for Protein Identification**—In our analysis of the unfractonated MPR affinity-purified mixtures, the threshold score for acceptance of non-lysosomal proteins was set as the lowest score observed for a known lysosomal protein. This resulted in a highly stringent cutoff of protein log(e) > −16.5. The threshold score for acceptance of purified Man-6-P glycopeptides from proteins not previously classified as lysosomal were also identified in the unfractonated mixture was set as the lowest score observed for a peptide derived from a known lysosomal species (peptide log(e) = −1.3 and −1.1 for human and mouse, respectively), although spectra of log(e) > −3 were also inspected manually and removed from our analysis if ambiguous. Although these thresholds for assignment are relatively low, identification of the corresponding protein in the unfractonated mixture together with the observation of N+203 in the context of a potential N-linked glycosylation site (NX(S/T)) greatly increased the confidence of assignment. Some peptides were observed from proteins not previously classified as lysosomal that were not observed by LC-MS analysis of the unfractonated mixtures. In these cases, these spectra were inspected manually and accepted only if a contiguous ion series included the modified NX(S/T) and if the peptide log(e) score was less than −2.

**Purification and Identification of Proteolytic Peptides Containing Man-6-P**—Eppendorf (Westbury, NY) BioPur tubes were used throughout. MPR affinity-purified proteins were digested with trypsin (Promega), Lys-C (Calbiochem), or with a combination of trypsin and Glu-C (Calbiochem). Ten micrograms of peptides in ammonium bicarbonate were heated twice for 10 min at 100 °C to inactivate proteases, lyophilized, dissolved in water, lyophilized twice to remove trace ammonium bicarbonate, and then dissolved in 300 μl of PBS containing 1 mM Pefabloc (Pentafarm, Basel, Switzerland). Peptides were transferred to another tube containing a 50-μl bed volume of Affi-Gel 15 (Bio-Rad) containing scI-MPR coupled at a density of 1 mg/ml and incubated with rocking for 60 min on ice. The peptide-scI-MPR/Affi-Gel mixture was then placed in a microcolumn at 4 °C containing an additional 50-μl bed volume of scI-MPR/Affi-Gel 15. The flow-through was reloaded 10 times, and then the microcolumn was placed in a 1.5-ml microcentrifuge tube and centrifuged for 2 s in a bench-top microcentrifuge to remove flow-through but not completely dry the resin. The resin was washed with 500 μl of PBS, with 500 μl of PBS containing 10 mM glucose 6-phosphate, and finally with 500 μl of 20 mM sodium phosphate buffer, pH 6.9, and then centrifuged again to remove excess buffer. The beads were resuspended in 100 μl of 20 mM sodium phosphate buffer, pH 6.9, containing 10 mM Man-6-P and incubated on ice for 5 min. The column was then centrifuged briefly, and the eluate was reloaded three times. The column was then centrifuged to dryness (10 s in microcentrifuge), washed with 50 μl of water to remove any remaining glycopeptides, and centrifuged again. Pooled eluates were concentrated to 50 μl in a vacuum concentrator resulting in a final buffer concentration of 40 mM sodium phosphate buffer, pH 6.9. To deglycosylate the eluted peptides, 0.5 μl of Endo H (1000 units/μl) or PNGase F (500 units/μl) (New England Biolabs) was added, and the sample was incubated at 37 °C for 2 h. Another aliquot of glycosidase was then added, and the incubation was repeated.

**RESULTS**

In our previous study of proteins purified from human brain by MPR affinity chromatography, we analyzed a trypptic digest of the total mixture. In this study, we digested proteins purified from human brain using several different proteolytic strategies (digestion with trypsin, Lys-C, Glu-C, and a mixture of trypsin and Glu-C) to provide wider proteome coverage. In addition, we also analyzed proteins purified from mouse brain after digestion with trypsin. Data for human brain were compiled from over 20 individual LC-MS/MS experiments on solution digests of MPR affinity-purified proteins. Data for mouse brain were obtained from individual LC-MS/MS runs on trypptic digests of 29 consecutive gel slices from 100 μg of MPR affinity-purified proteins fractionated in a single lane by SDS-PAGE. Forty-eight known lysosomal proteins were identified from human brain (Table I), representing the 44 proteins identified previously (7) as well as four additional known lysosomal proteins (galactocerebrosidase, cathepsin B, GM2 activator, and palmitoyl-protein thioesterase 2). In mouse, 42 known lysosomal proteins were found (Table I), all of which were observed in the human sample with the exception of cathepsin S. In total, analysis of proteins purified from mouse and human brain resulted in the identification of 49 known lysosomal proteins.

Forty-three proteins that are not currently thought to be lysosomal were identified in human brain (Table II), and of these, 13 were also identified in mouse brain. An additional 22 proteins not currently thought to be lysosomal were found only in mouse brain (Table II). Identification of proteins from both sources not only increases the confidence of peptide assignment but also identifies them as proteins that are less likely to represent simple contaminants. Conversely, proteins that are species-specific might be regarded to be more likely to represent contaminants, although differences in the data-
ical proteins LOC196463 and FLJ22662 are predicted to have syndrome critical region 1, phospholipase D3, and hypothet-
erase 3A, epididymis-specific.

For example, acid sphingomyelinase-like phosphodies-
tively considered lysosomal have known or predicted bi-
mouse sample (see “Experimental Procedures”).

A number of the mouse and human proteins that are not
Cathepsin D
CTSB

Cathepsin B

Cathepsin D

CTSD

Cathepsin F

CTSF

Cathepsin H

CTSH

Cathepsin L

CTSL

Cathepsin S

CTSS

Cathepsin Z

CTSZ

CLN5 protein

CLN5

Di-N-acetylchitobiase

CTBS

Sialic acid 9-O-acetylatedesterase

CSE-C

Deoxyribonuclease II

DNASE2

Dipeptidyl-peptidase I

CTSC

Dipeptidyl-peptidase II

DPP7

NPC2 protein

NPC2

Galactocerebrosidase

GALC

γ-Glutamyl hydratase

GGH

Ganglioside GM2 activator

GM2A

Iduronate 2-sulfatase

IDS

Legumain

LGMN

Lysosomal acid lipase

LIPA

Lysosomal α-glucosidase

GAA

Lysosomal α-mannosidase

MAN2B1

Lysosomal protective protein

PPGB

Lysosomal Pro-X carboxypeptidase

PRCP

Myeloperoxidase

MPO

N-Acetylglactosamine-6-sulfatase

GALNS

N-Acetylglucosamine-6-sulfatase

GNS

N-Sulfoglucosamine sulfatohydrolase

SGSH

Palmitoyl-protein thioesterase 1

PPT1

Palmitoyl-protein thioesterase 2

PPT2

Plasma glutamate carboxypeptidase

PGCP

Proactivator polypeptide

PSAP

Sphingomyelin phosphodiesterase

SMPD1

Tissue α-L-fucosidase

FUCA1

Tripeptidyl-peptidase I

TPP1

sets from mouse and human may also reflect the fact that the
human sample was more extensively characterized than the
mouse sample (see “Experimental Procedures”).

A number of the mouse and human proteins that are not
currently considered lysosomal have known or predicted bi-
ological properties that are consistent with lysosomal func-
tion. For example, acid sphingomyelinase-like phosphodies-
terase 3A, epididymis-specific α-mannosidase, and plasma
α-fucosidase are similar to known lysosomal proteins. Cat eye
syndrome critical region 1, phospholipase D₃, and hypothet-
ical proteins LOC196463 and FLJ22662 are predicted to have
hydrolytic activities; this is typical of most soluble lysosomal
proteins. Candidates for novel lysosomal proteins are indi-
cated in Table II, but these assignments are tentative, reflect-
ing the difficulties in classifying MPR affinity-purified proteins
when the presence of Man-6-phosphorylation is not verified.

Our approach to the identification of sites of N-linked oli-
gosaccharides containing Man-6-P was as follows. 1) A mix-
ture of proteins was purified from human and mouse brain by
affinity chromatography on immobilized MPRs. 2) Aliquots of
the mixture were digested with trypsin or other proteases. 3) A
portion of each digest was analyzed by LC-MS/MS, and

| Protein                        | Gene      | Homo sapiens | Mus musculus |
|--------------------------------|-----------|--------------|--------------|
| Asparaginase                   | AGA       | −147         | −114         |
| 1-O-Acylseramide synthase      | LYPLA3    | −86          | −228         |
| Acid ceramidase                | ASAH      | −303         | −126         |
| α-Galactosidase A              | GLA       | −129         | −160         |
| α-L-iduronidase                | IDUA      | −93          | −172         |
| α-N-Acetylglactosaminidase      | NAGA      | −135         | −122         |
| α-N-Acetylglucosaminidase       | NAGLU     | −134         | −97          |
| Arylsulfatase A                | ARSA      | −212         | −275         |
| Arylsulfatase B                | ARSB      | −116         | −310         |
| β-Galactosidase                | GLB1      | −187         | −277         |
| β-Glucuronidase                | GUSB      | −121         | −73          |
| β-Hexosaminidase α chain       | HEXA      | −246         | −203         |
| β-Hexosaminidase β chain       | HEXB      | −240         | −178         |
| β-Mannosidase                  | MANBA     | −98          | −134         |
| Cathepsin B                    | CTSB      | −18          | −201         |
| Cathepsin D                    | CTSD      | −296         | −256         |
| Cathepsin F                    | CTSF      | −45          | −79          |
| Cathepsin H                    | CTSH      | −58          | −27          |
| Cathepsin L                    | CTSL      | −82          | −172         |
| Cathepsin S                    | CTSS      | −54          |              |
| Cathepsin Z                    | CTSZ      | −120         | −146         |
| CLN5 protein                   | CLN5      | −34          | −24          |
| Di-N-acetylchitobiase           | CTBS      | −61          | −113         |
| Sialic acid 9-O-acetylatedesterase | CSE-C  | −258         | −64          |
| Deoxyribonuclease II            | DNASE2    | −57          | −16          |
| Dipeptidyl-peptidase I          | CTSC      | −117         | −55          |
| Dipeptidyl-peptidase II         | DPP7      | −245         | −152         |
| NPC2 protein                   | NPC2      | −96          | −36          |
| Galactocerebrosidase            | GALC      | −39          |              |
| γ-Glutamyl hydratase            | GGH       | −127         |              |
| Ganglioside GM2 activator       | GM2A      | −16          |              |
| Iduronate 2-sulfatase           | IDS       | −227         | −114         |
| Legumain                       | LGMN      | −31          | −44          |
| Lysosomal acid lipase           | LIPA      | −65          |              |
| Lysosomal α-glucosidase         | GAA       | −281         | −231         |
| Lysosomal α-mannosidase         | MAN2B1    | −205         | −91          |
| Lysosomal protective protein    | PPGB      | −156         | −178         |
| Lysosomal Pro-X carboxypeptidase| PRCP      | −160         | −38          |
| Myeloperoxidase                | MPO       | −164         |              |
| N-Acetylglactosamine-6-sulfatase| GALNS     | −126         | −95          |
| N-Acetylglucosamine-6-sulfatase | GNS       | −205         | −234         |
| N-Sulfoglucosamine sulfatohydrolase | SGSH  | −99          | −85          |
| Palmitoyl-protein thioesterase 1 | PPT1    | −302         | −255         |
| Palmitoyl-protein thioesterase 2 | PPT2    | −28          |              |
| Plasma glutamate carboxypeptidase | PGCP   | −172         | −309         |
| Proactivator polypeptide        | PSAP      | −307         | −71          |
| Sphingomyelin phosphodiesterase | SMPD1    | −118         | −66          |
| Tissue α-L-fucosidase           | FUCA1     | −84          | −45          |
| Tripeptidyl-peptidase I         | TPP1      | −239         | −34          |
Mannose 6-Phosphorylation Sites

Proteins that are not thought to have lysosomal function or localization identified in MPR affinity-purified mixtures from human and mouse brain

| Protein                                                                 | Gene          | Log(e) | H. sapiens | M. musculus |
|------------------------------------------------------------------------|---------------|--------|------------|-------------|
| 60-kDa heat shock protein                                              | HSPD1         | -34    | -32        |             |
| Acid sphingomyelinase-like phosphodiesterase 3A*                      | SMPDL3A       | -28    | -291       |             |
| α2-Macroglobulin                                                       | A2M           | -66    | -35        | -45         |
| Angiotensinogen                                                        | AGT           | -34    | -108       |             |
| Mitochondrial ATP synthase, α chain                                   | ATP5A1        | -78    | -19        |             |
| Mitochondrial ATP synthase, β chain                                   | ATP5B         | -36    | -33        |             |
| ATPase, Na+/K+ - transporting, α subunit                               | ATP1A2        | -40    | -40        |             |
| ATP-binding cassette subfamily E                                       | ABCE1         | -38    | -33        |             |
| Biglycan                                                               | BGN           | -22    | -22        |             |
| Biotinidase precursor (EC 3.5.1.12)*                                    | BTD           | -65    | -65        |             |
| ATPase, Na+/K+ - transporting, α1 polypeptide                          | ATP1A1        | -150   | -150       |             |
| Hypothetical protein DKF/Zp319G1735 (telethon sulfatase)*              | DKF/Zp319G1735| -108   | -108       |             |
| Breccan core protein                                                  | BCAN          | -90    | -142       |             |
| Cat eye syndrome critical region 1*                                    | CECR1         | -52    | -52        |             |
| Cellular repressor of E1A-stimulated gene transcription               | CREG          | -87    | -87        |             |
| Clustatin                                                             | CLU           | -115   | -115       |             |
| Complement component 1, Q subcomponent                                 | C1QB          | -69    | -69        |             |
| Contactin                                                             | CNTN1         | -37    | -37        |             |
| Cystatin B                                                            | CSTB          | -58    | -58        |             |
| Cystatin C                                                            | CST3          | -66    | -66        |             |
| Parkinson disease (autosomal recessive, early onset) 7                 | PARK7         | -16    | -16        |             |
| Epididymis-specific α-mannosidase*                                     | MAN2B2        | -177   | -125       |             |
| F-box only protein 2                                                   | FBXO2         | -90    | -90        |             |
| F-box only protein 6b                                                  | FBXO6B        | -32    | -32        |             |
| Fc fragment of IgG-binding protein                                     | FCGBP         | -46    | -46        |             |
| Ferritin                                                              | FTH1          | -66    | -66        |             |
| Plasma α-L-2-fucosidase*                                               | FUCA2         | -25    | -25        |             |
| Galectin-1                                                            | LGALS1        | -22    | -22        |             |
| GDP-fucose protein O-fucosyltransferase                                | POFTU2        | -41    | -41        |             |
| Glia-derived nexin                                                    | SERPINE2      | -54    | -54        |             |
| Gial fibrillary acidic protein                                         | GFAP          | -20    | -20        |             |
| Glutamine amidotransferase class I                                    | ENSMUSP0000075593 | -21 | -21 |             |
| Glyceraldehyde-3-phosphate dehydrogenase                               | G3CEV2        | -31    | -31        |             |
| Golgi phosphoprotein 2                                                | GOLPH2        | -37    | -37        |             |
| Hyaluronan and proteoglycan link protein                               | HAPLN1        | -20    | -20        |             |
| Hypothetical protein LOC196463*                                        | LOC196463     | -175   | -153       |             |
| Hypothetical protein FLJ22662*                                         | FLJ22662      | -27    | -27        |             |
| Leucine-rich gliala-inactivated protein                                | LG1           | -19    | -19        |             |
| Leucine-rich repeat LGI family member                                  | LG3           | -22    | -22        |             |
| Leucine-rich repeat neuronal 6A                                        | LRRN6A        | 17     | 17         |             |
| Mammalian ependyma-related protein                                     | EPDR1         | 147    | 147        |             |
| Myelin-associated glycoprotein                                        | MAG           | 41     | 41         |             |
| N-Acetyllactosaminide β-1,3-β- acetylglucosaminyltransferase 6          | B3GNT6        | 259    | 259        |             |
| N-Acylphosphoglycerol amidohydrolase 2*                                | ASAH2         | 26     | 26         |             |
| N-Acylphosphoglycerol amidohydrolase-like*                             | ASAH1         | 26     | 26         |             |
| Neurexin 3-α                                                          | NRXN3         | 25     | 25         |             |
| Neurocan core protein                                                  | CSPG3         | -46    | -46        |             |
| RIKEN cDNA 9630046K23Rik gene                                           | 9630046K23Rik | 38     | 38         |             |
| Phospholipase D3*                                                      | PLD3          | -33    | -33        |             |
| Prostaglandin-H2 ω-isomerase*                                         | PTGDS         | -48    | -48        |             |
| Receptor-type tyrosine-protein phosphatase                            | PTPRZ1        | -55    | -55        |             |
| Retinoid-inducible serine carboxypeptidase*                            | SCPEP1        | -31    | -31        |             |
| Ribonuclease T2*                                                       | RNASET2       | -112   | -112       |             |
| Serine (or cysteine) proteinase inhibitor                              | SERPINB8      | -57    | -57        |             |
| Sodium/potassium-transporting ATPase chain 1                          | ATP1A1        | -33    | -33        |             |
| Sodium/potassium-transporting ATPase chain 3                          | ATP1A3        | -18    | -18        |             |
| Sodium/potassium-transporting ATPase chain 4                          | ATP1A4        | -24    | -24        |             |
| SPARC-like protein 1                                                   | SPARCL1       | -33    | -33        |             |
| S phase kinase-associated protein 1                                   | SKP1A         | -65    | -65        |             |
| Sulfatase-modifying factor 2                                          | SUMF2         | -19    | -19        |             |
| T-complex protein 1, β subunit                                        | CCT7          | -18    | -18        |             |
| Tenascin R                                                            | TNR           | -37    | -37        |             |
| Versican core protein                                                  | CSPG2         | -20    | -20        |             |
| Carbamoyl-phosphate synthase                                          | CPS1          | -24    | -24        |             |
| β-3-Glycosyltransferase-like                                          | B3GTL         | -18    | -18        |             |

* Proteins with known or predicted biological function that may be consistent with lysosomal function.
results were collated to identify components of the mixture. 4) Ten micrograms of each digest were loaded onto a microcolumn of immobilized MPR. 5) The column was washed, and Man-6-P glycoproteins were eluted with free Man-6-P. 6) Man-6-P glycopeptides were enzymatically deglycosylated. 7) Peptides were analyzed by LC-MS/MS. 8) Spectra were assigned to peptides using X! Tandem. 9) Data were analyzed for the presence of mass modifications at asparagine residues that are indicative of deglycosylation and inspected to determine whether they fell within the context of an N-linked glycosylation motif. For the Endo H-deglycosylated human brain glycopeptides, data were compiled from 16 individual LC-MS/MS runs under different sample load and protease digestion conditions. Endo H-deglycosylated mouse brain glycopeptide data were compiled from two LC-MS/MS experiments using trypsin digestion.

Prior to LC-MS/MS analysis of purified Man-6-P glycopeptides, carbohydrate residues must be removed, and two enzymes, PNGase F and Endo H, were used here. PNGase F will remove all N-linked oligosaccharides, resulting in conversion
of asparagine to aspartate with a +1-Da increment in mono-isotopic peptide mass. Endo H removes high mannose and some hybrid-type oligosaccharides leaving a residual isotopic peptide mass. Endo H removes high mannose and a signature mass increment of +203 Da. Examples of spectra assigned to peptides with the N+203 modification are shown in Fig. 1.

Preliminary experiments revealed significant problems with the use of N+1 after deglycosylation with PNGase F as a mass marker for N-linked oligosaccharides when data are analyzed in an automated manner. False positive errors can occur when N+1 is assigned to peptides that have not been deglycosylated. For example, we analyzed a total digest of human brain glycopeptides when data are analyzed in an automated manner. False positive errors can occur when N+1 is assigned to peptides that have not been deglycosylated. For example, we analyzed a total digest of human brain MPR affinity-purified glycopeptides because a high resolution zoom scan protocol was used to accurately determine parent ion mass, but deamidation in the context of N-acetylgalactosamine coupled to asparagine with a signature mass increment of +203 Da. Examples of spectra assigned to peptides with the N+203 modification are shown in Fig. 1.

For details, see Table IV.

---

**Table III**

Automated residue modification assignment in known lysosomal proteins after analysis of MPR affinity-purified glycopeptides deglycosylated with either Endo H or PNGase F

| Deglycosylation | Number of N+203 assignments |  |
|-----------------|-----------------------------|---|
|                 | Not in NXS(T)               | In NXS(T) |
| Endoglycosidase H | 1*                          | 2450    |
| PNGase F         | 0                           | 0       |

| Deglycosylation | Number of N+1 assignments |  |
|-----------------|---------------------------|---|
|                 | Not in NXS(T)             | In NXS(T) |
| Endoglycosidase H | 18                        | 15   |
| PNGase F         | 86*                       | 737   |

* For details, see Table IV.

---

**Table IV**

N+1 or N+203 automatically assigned outside of NXS(T)

Potential N-linked glycosylation sites are underlined where present. Eight peptides assigned N+1 that lacked potential N-linked glycosylation sites probably represent chemical deamidated contaminants and are italicized.

| Sequence | Modified Asn | Number of assignments |
|----------|--------------|-----------------------|
| Endo H   | LIDNKTEKEILDAFDK | 331NNK 1 |
| PNGase F | EVTKLIDNNKTEK | 331NNK 1 |
|          | LIDNKTEKEILDAFDK | 331NNK 1 |
|          | FDGILGMAYPRISVNNNLVPFDNLMOQQK | 209NNV 2 |
|          | INLTFASNVTDFMDPDLPTDE | 362NNV 1 |
|          | ISVNVLPVFDNLMOQQK | 209NNV 2 |
|          | ISVNVLPVFDNLMOQQK | 210NVL 2 |
|          | IVYQNYINYSR    | 174NQP 1 |
|          | KLGVYLDRLNEKNS TK | 422NLE 5 |
|          | LIDNKTEK       | 331NNK 18 |
|          | LIDNKTEKEILDAFDK | 331NNK 25 |
|          | LIDNKTEKEILDAFDK | 331NNK 3 |
|          | LNNTYIFTSNGHYHTQFSLPIDK | 318NTY 1 |
|          | LNNTYIFTSNGHYHTQFSLPIDK | 318NTY 2 |
|          | LNNTYIFTSNGHYHTQFSLPIDK | 327NGY 1 |
|          | LVYGILDRNLEKNS TK | 422NLE 3 |
|          | LWSSLQTHCCSONK | 244NKC 2 |
|          | NLEKNSTKOEILAALEK | 422NLE 11 |
|          | YPHNHVVNNTLEGNCSSK | 106NHH 1 |

---

Given these problems with the automated assignment of N+1, we also examined data from peptides deglycosylated with Endo H for the same type of potential errors. In terms of
false positives, no peptides derived from known lysosomal proteins were assigned N+203 when data derived from a non-deglycosylated sample were analyzed automatically. In terms of false negatives, the N+203 modification was assigned to a total of 2451 peptides of which only one was assigned the N+203 modification outside of NX(S/T). In this case, N+203 was assigned to an asparagine adjacent to NX(S/T) (Tables III and IV), and this assignment error was easily detected and corrected manually. This is a significantly lower incidence of false negative errors than observed with PNGase F, thus we chose to conduct subsequent analyses with Endo H. This glycosidase also has advantages for our purposes in that the N+203 modification occurs only by deglycosylation and because Endo H specifically cleaves only those classes of carbohydrate that are capable of receiving the Man-6-P modification (see below and “Discussion”). A number of peptides deglycosylated with Endo H were assigned N+1 both in NX(S/T) and outside of this motif (Tables III and IV). N+1 modification of these peptides could reflect spectra interpretation error or deamidation, although the possibility of a low level of contaminant PNGase F-like activity within the Endo H preparation cannot be excluded.

During the synthesis of N-linked oligosaccharides containing Man-6-P, mannose residues in the α1,6 branch are normally phosphorylated before mannose residues positioned within the α1,3 branch (18) as part of a highly ordered process reflecting the spatial organization of the oligosaccharide processing enzymes in the ER, transitional elements, and the Golgi apparatus. Phosphorylation of mannose residues within the α1,6 branch prevents high mannose-type units from being converted to Endo H-resistant structures by impairing α-mannosidase action. However, a mutant mouse lymphoma cell line that has a glycosylation defect resulting in the synthesis of aberrant Endo H-resistant N-linked glycans (19) that have a truncated core mannose α1,6 branch (20) is still able to confer the Man-6-P modification to the α1,3 branch of the oligosaccharide (21). Therefore, although most Man-6-P-containing glycans from normal mouse cells are Endo H-sensitive (22, 23), we were concerned that in brain there may exist some oligosaccharides containing Man-6-P that are Endo H-resistant. We therefore compared the efficiency of Endo H and PNGase F in removing Man-6-P oligosaccharides. We deglycosylated human brain MPR-purified proteins with each enzyme, fractionated the digest by SDS-PAGE, transferred to nitrocellulose, and used radiolabeled MPR as an affinity probe to detect proteins retaining Man-6-P (Fig. 2). Treatment with both PNGase F and Endo H resulted in a complete loss of detectable Man-6-P glycoproteins indicating that Endo H-resistant Man-6-P oligosaccharides are not present.

In our analysis of Endo H-deglycosylated MPR affinity-purified human brain glycopeptides, 2973 peptides were assigned to known lysosomal proteins, and of these, 522 (17.6%) lacked N+203. These peptides likely represent abundant peptides that are not fully removed but are presumably depleted by affinity purification. To estimate the extent of this depletion, we measured their abundance in both the unfractionated digest of purified brain proteins and in the purified Man-6-P glycopeptide fraction by the integration of extracted ion chromatograms representing the different charge states of select contaminant peptides (peaks were assigned on the basis of retention time and MS/MS). Many of the contaminant peaks were abundant in the unfractionated digest but present at levels approximating background in the Man-6-P glycopeptide fraction and were thus not quantifiable. For peptides whose abundance could be estimated in the Man-6-P glycopeptide fraction, their levels were found to range from 0.06 to 0.45% of starting levels, representing a significant depletion (more than ~200-fold).

In total, 99 different sites of Man-6-phosphorylation were identified from the mouse and human MPR affinity-purified mixtures after deglycosylation with Endo H (Table V) corresponding to 42 of the 49 known lysosomal proteins identified in the unfractionated mixtures (Table I). For a number of these lysosomal proteins, the location of N-linked oligosaccharides that contain Man-6-P have been experimentally determined in site-directed mutagenesis experiments, allowing us to evaluate our approach. Human GUSB (see Table I for gene nomenclature) is glycosylated at all four potential N-linked glycosylation sites, but only two are important for intracellular trafficking and are thus thought to contain Man-6-P (24). The single GUSB peptide identified here contains one of these two N-linked glycosylation sites. ASAH has six potential N-linked glycosylation sites of which the N-terminal five are used, and sites 1, 3, and 5 from the N terminus are thought to contain...
### Mannose 6-Phosphorylation Sites

**TABLE V**

| Protein | Human site (score) | Mouse site (score) |
|---------|--------------------|--------------------|
| 1-O-Acylceramide synthase | 289NYT (−5.3)a | 289NYT (−8.3) |
| Acid ceramidase | 168NDT (−8.8) | 258NTT (−1.6) |
| α-Galactosidase A | 150NCT (−3.9) | Not present |
| α-L-Iduronidase | 192NRT (−3.9) | 288NSS (−5.7) |
| α-N-Acylgalactosaminidase | 201NYS (−8) | 241NYT (−1.1) |
| Arylsulfatase A | 356NVT (−4.1) | 186NVT (−6.8) |
| Arylsulfatase B | 426NTS (−5.3) | Not present |
| β-Galactosidase | 241NIT (−2.3) | Not present |
| β-Glucuronidase | 446NNT (−2.9)a | 416NES (−7.8) |
| β-Hexosaminidase α chain | 272NGT (−3.4) | 157NKT (−3.9) |
| β-Hexosaminidase β chain | 195NKT (−3.3) | 261NTT (−8.9) |
| Cathepsin D | 124NKT (−3.8) | 261NTT (−10.2) |
| Cathepsin F | 266NT (−4.6) | 345NDS (−4.9) |
| Cathepsin H | 89NCS (−7.4) | 440NRS (−7.1) |
| Cathepsin L | 221NDT (−9.5) | 221NDT (−12.6) |
| Cathepsin Z | 224NYT (−5.5) | 286NYT (−2.4) |
| CLN5 protein | 329NIT (−7.4) | 186NYT (−2.1) |
| Deoxyribonuclease II | 299NKS (−7.9) | Not present |
| Di-N-acetylchitotriose | 299NSS (−7) | 215NGS (−5.5) |
| Dipeptidyl-peptidase I | 53NCS (−5.1) | 53NCS (−6) |
| Dipeptidyl-peptidase II | 275NSS (−8.6) | 325NSS (−7.7) |
| γ-Glutamyl hydrolase | 318NAS (−9) | 162NFT (−1.1) |
| Iduronate 2-sulfatase | 307NIS (−9.4) | 238NIT (−4.1) |

**TABLE V — continued**

| Protein | Human site (score) | Mouse site (score) |
|---------|--------------------|--------------------|
| Legumain | 91NGT (−3.8) | 93NGT (−7.2) |
| Lysozyme | 269NTS (−2.3)b | 269NTS (−2.8)a |
| Lysozyme α-glucosidase | 277NKT (−2.3)b | 277NKS (−2.8)a |
| Lysozyme α-mannosidase | 318NHT (−5.3) | 393NLS (−5.5) |
| Lysozyme protective protein | 140NDT (−11.8) | 140NDT (−7.8) |
| Myeloperoxidase | 318NHT (−5.3)a | Not present |
| N-Acetylglucosamine-6-sulfatase | 205NLT (−7.3)a | 203NLT (−5.5) |
| N-Acetylglucosamine-6-sulfatase | 199NYS (−8.7)b | 210NVS (−8.7)b |
| N-Acetylgalactosaminidase | 206NAT (−8.8) | Not present |
| N-Asparaginase | 386NAT (−4.7) | 386NAT (−4.7) |
| N-Sulfogalactosamine sulfohydrolase | 130NIS (−6.6)a | Not present |
| Palmitoyl-protein thioesterase 1 | 266NDT (−4.9) | 266NDT (−3) |
| Palmitoyl-protein thioesterase 2 | 195NHS (−12.6)a | 195NVS (−11.5) |
| Palmitoyl-protein thioesterase 2 | 235NDS (−11.4)a | 235NDS (−13.1) |
| Plasma glutamate carboxypeptidase | 60NTE (−3) | Not present |
| Plasma glutamate carboxypeptidase | 289NRT (−6.3) | 289NRT (−1.6) |
| Plasma glutamate carboxypeptidase | 395NVT (−1.7) | Not present |
| Proactivator polypeptide | 179NYS (−8.5)a | 354NIS (−3.9) |
| Sialic acid 9-O-acetylesterase | 80NAT (−10) | 80NAT (−5.6) |
| Sialic acid 9-O-acetylesterase | 335NKT (−8.3) | 335NKT (−5.3) |
| Sialic acid 9-O-acetylesterase | 426NAT (−8.1) | 426NAT (−4.5) |
| Sialic acid 9-O-acetylesterase | 259NAT (−3.6) | 259NAT (−3.6) |
| Sphingomyelin phosphodiesterase | 88NLT (−4.7) | 84NLT (−3.3) |
| Sphingomyelin phosphodiesterase | 377NT (−4.2) | 335NQS (−8.7) |
| Tissue α-L-fucosidase | 215NLT (−11)b | Not present |
| Tripeptidyl-peptidase I | 222NPS (−4.1) | 222NPS (−2.9) |

---

*a* Sites of Man-6-phosphorylation contained within peptides that were also assigned during analysis of the non-deglycosylated, unfractonated proteolytic digests of purified brain proteins. The presence of oligosaccharides at these sites is therefore heterogeneous. One spectra was assigned to murine palmitoyl-protein thioesterase 2, a protein not identified in the unfractonated proteolytic digests. In this case, we could not determine whether there exists glycosylation heterogeneity at the site of Man-6-phosphorylation.

*b* Potential sites of Man-6-phosphorylation that were found on doubly glycosylated peptides for which singly glycosylated equivalents were not identified. These sites clearly contain Endo H-sensitive carbohydrates (indicated by the N+203 modification), but assignments of Man-6-phosphorylation should be regarded as tentative.
Man-6-P (25); all three were identified here. Similarly Man-6-P glycopeptides corresponding to known sites for Man-6-phosphorylation were also found for HEXA (26), CLN1 (27), CTSL (28), PPGB (29), CTSC (30), and HEXB (31). In some cases, we also identified glycosylation sites that were not found to contain Man-6-P in previous studies. Most likely, these sites are Man-6-phosphorylated at levels below the threshold of detection of site-directed mutagenesis studies, but it is possible that these differences reflect differences in the source of lysosomal proteins (cell culture medium versus brain). Regardless it is clear that our global results and earlier studies of Man-6-phosphorylation of individual proteins are in excellent agreement.

Of the 99 Man-6-P glycosylation sites identified in mouse and human brain, 34 were assigned in both species, 14 were assigned only in mouse, and 51 were assigned only in human. (The degree of similarity is probably higher between the two species as the data here likely reflect the more extensive “Experimental Procedures”). The majority of the N-linked glycosylation sites containing Man-6-phosphorylated oligosaccharides are conserved between mice and humans, although 10 were not, and these are indicated in Table V.

In the analysis of the human brain sample, a small number of peptides were identified that contained more than one N(S/T) motif with an N+203 residue modification, indicating the presence of two Endo H-sensitive N-linked carbohydrates. In such cases, either or both carbohydrates can contain Man-6-P. For some of these doubly glycosylated peptides, glycosylation was found to be heterogeneous, and we also found the same peptide containing only a single N+203 modification, allowing us to unambiguously assign one of the two potential Man-6-phosphorylation sites. In other cases, cleavage with alternate proteases between the two potential sites of Man-6-phosphorylation allowed unambiguous assignment of at least one site. Some potential Man-6-P sites were identified only in doubly glycosylated peptides. Although it is likely that both of such sites actually contain Man-6-P, we cannot formally prove this, and such cases are indicated in Tables V and VI.

A number of the MPR-purified proteins that have not been assigned lysosomal function or localization were found to contain Man-6-phosphorylated oligosaccharides. These are listed in Table VI and described individually below.

Acid Sphingomyelinase-like Phosphodiesterase 3A (SMPDL3A)—SMPDL3A was first identified in a yeast two-hybrid screen for proteins interacting with a bladder tumor suppressor (32). SMPDL3A is 29% identical and 44% similar to acid sphingomyelinase, a known lysosomal protein, and we identified a single, conserved site of Man-6-phosphorylation in both mouse and human brain. This fact together with the observation that it contains Man-6-P makes SMPDL3A an excellent candidate for a protein with novel lysosomal function.

Acyloxyacyl Hydrolase (AOAH)—A tryptic Man-6-P glycopeptide derived from AOAH was found here, although the protein itself was not detected in the total unfractionated mixture of MPR affinity-purified proteins from human brain. AOAH is a highly conserved lipase found in immune cells of myeloid lineage that deacylates bacterial lipopolysaccharides and other substrates (33). AOAH is also found in urine where it is secreted by proximal tubule cells and taken up by bladder cells in a process that in vitro is inhibited by Man-6-P (34), which is in accordance with the identification of a Man-6-P glycopeptide here. The hydrolytic function of AOAH, partial similarity to lysosomal sphingolipid activator proteins (saposins), presence of Man-6-P, and the fact that the protein is processed upon cellular uptake (presumably upon MPR-mediated delivery to the lysosome (35)) make AOAH another promising lysosomal candidate.

Arylsulfatase G (6330406P08Rik)—A single site of Man-6-phosphorylation was identified for predicted protein arylsulfatase G in mouse brain.

Biotinidase (BTD)—BTD (36) generates free biotin, a cofactor required for four cellular carboxylases, by hydrolysis of biocytin or biotinyl peptides derived from the diet or from the turnover of biotinylated cellular proteins. BTD deficiency causes multiple carboxylase deficiency in humans. We identified two sites of mannose 6-phosphorylation in BTD, and a potential lysosomal role for BTD is supported by the observations that BTD has an acidic pH optima (37) and that biotinylated carboxylases are known to be degraded in the lysosome (38). Immunohistochemical detection of BTD reveals a punctate cytoplasmic distribution that is consistent with lysosomal localization, although results reported from subcellular fractionation were interpreted to suggest a microsomal localization for the bulk of cellular BTD activity (38).

Calumenin (CALU)—CALU is a calcium binding protein that is both secreted and found within the cell throughout the secretory system (39). The single Man-6-P site identified here in mouse brain is known to be glycosylated (39).

Cat Eye Syndrome Critical Region 1 (CECR1)—Two sites of Man-6-phosphorylation were identified for CECR1, a protein of unknown function with some similarity to invertebrate growth factors and to adenosine deaminase (40). CECR1 has a predicted signal sequence and maps to a region that is duplicated in a developmental disorder, cat eye syndrome. Overexpression of CECR1 has been suggested to play a role in this disease (40).

Cellular Repressor of E1A-stimulated Gene Expression (CREG)—Two sites of Man-6-phosphorylation were identified for CREG. The presence of Man-6-P-containing oligosaccharides was inferred in a previous study in which CREG was shown to bind to the CI-MPR (41).

Epididymis-specific α-Mannosidase (MAN2B2)—Two Man-6-phosphorylation sites were observed for MAN2B2. This observation together with its catalytic function suggests that this protein almost certainly has lysosomal function.
**TABLE VI**

Sites of mannos-6-phosphorylation in Endo H-deglycosylated brain MPR affinity-purified proteins in proteins that are not known to have lysosomal function

In most cases, *N*-linked glycosylation sites are conserved between mouse and humans; however, in some cases, this is not the case and non-conserved sites are indicated by “Not present.” Blank entries indicate that spectra were not assigned to the peptide containing the indicated site of Man-6-phosphorylation. The score representing the most confident peptide assignment (lowest log(e)) is indicated in parentheses.

| Protein                                                | Human site (score) | Mouse site (score) |
|--------------------------------------------------------|--------------------|--------------------|
| Acid sphingomyelinase-like phosphodiesterase 3A        | 238NNS (−3.6)      | 235NSS (−10.7)     |
| Acylxoyacyl hydrolase                                  | 207NDS (−5)        |                    |
| Biotinidase                                            | 150NDT (−2.1)      |                    |
| 349NAT (−4.7)                                         |                    |                    |
| Arylsulfatase G                                        | 174NVT (−4.9)      |                    |
| Calumenin                                              | 160NET (−4.7)      |                    |
| 216NVT (−8.4)*                                        |                    |                    |
| Epidysmystis-specific α-mannosidase                    | 229NRS (−3.6)      |                    |
| 748NNS (−7.3)                                         |                    |                    |
| GDP-fucose protein O-fucosyltransferase                | 250NST (−2.9)      | 256NST (−3.7)      |
| Golgi phosphoprotein 2                                 |                    | 150NIT (−2.2)      |
| Hypothetical protein FLJ22662                          | 411NWS (−2.4)      |                    |
| Hypothetical protein LOC196463                         | 68NLT (−13.4)      | 98NLT (−8.9)       |
| 231NKT (−3.3)                                         |                    |                    |
| 436NAS (−8.4)                                         |                    |                    |
| 465NOS (−3.7)                                         |                    |                    |
| 515NGS (−6)                                           |                    |                    |
| 520NST (−6.3)                                         |                    |                    |
| Mammalian ependymin-related protein                    |                    |                    |
| Microsomal stress 70 protein ATPase                    | 184NST (−6)        | 150NSS (−9.7)*     |
| Myelin-associated glycoprotein                         | 223NTT (−7.5)      | 97NLT (−11.9)      |
| N-Acetylglactosaminidase β-1,3-N-acetylglicosaminyltransferase | 204NV (−3)     |                    |
| Neuroserpin                                            | 157NNT (−5)        |                    |
| 321NLT (−3.7)                                         |                    |                    |
| Plasma α-L-2-fucosidase                                | 377NDT (−2.1)      |                    |
| Proenkephalin A                                        |                    |                    |
| Prostaglandin-H$_2$ D-isomerase                        | 78NLT (−14.2)      |                    |
| Ribonuclease T2                                        | 76NRS (−4.7)       |                    |
| 108NRS (−7.8)                                         |                    |                    |
| 212NCT (−2.4)                                         |                    |                    |
| SPARC-like protein 1                                   | 498NYS (−2.6)      | 496NQT (−3.7)      |
| Bone Rel sulfatase                                     | 191NRT (−6)        | 496NYS (−4.9)*     |
| Sulfatase-modifying factor 2                           |                    |                    |
| Vasopressin-neurophysin 2 copeptin                     |                    | 135NAT (−2.6)*     |

* Sites of Man-6-phosphorylation contained within peptides that were also assigned during analysis of the non-deglycosylated, unfraccionated proteolytic digests of purified brain proteins. The presence of oligosaccharides at these sites is therefore heterogeneous.

* Some spectra were assigned to proteins that were not identified in the unfraccionated proteolytic digests. In these cases, we could not determine whether there exists glycosylation heterogeneity at the sites of Man-6-phosphorylation.

**GDP-fucose Protein O-Fucosyltransferase 2 (POFUT2)—**

POFUT2 is an uncharacterized protein with some resemblance (22% identical, 37% similar) to protein O-fucosyltransferase 1 (POFUT1), a soluble protein thought to be localized within the endoplasmic reticulum (42). A single conserved site of Man-6-phosphorylation was found in both mouse and human POFUT2.

**Hypothetical Protein LOC196463 (LOC196463)—**

LOC196463 has been identified as a potential Man-6-P glycoprotein in human brain (7) and is a predicted protein of unknown function with some similarity to Dictyostelium discoideum phospholipase B. We found here that five of the six potential *N*-linked glycosylation sites contain Man-6-P. LOC196463 has also been identified recently in neuromelanin granules from human brain that are subcellular compartments rich in lysosomal proteins (43) and as an MPR affinity-purified protein secreted from murine MPR-deficient fibroblasts (8). These observations together with the finding that LOC196463 contains numerous Man-6-phosphorylated oligosaccharides are highly suggestive that this protein actually represents a lysosomal protein. Two of the sites of Man-6-phosphorylation of LOC196463 were conserved between mouse and human.
LOC196463 and FLJ22662 (see below) may well represent a novel class of lysosomal phospholipases.

Hypothetical Protein FLJ22662 (FLJ22662)—FLJ22662 is a predicted protein observed previously (7) that has considerable sequence similarity to LOC196463 (33% identical, 48% similar). A single Man-6-phosphorylation site was identified here in FLJ22662 purified from human brain.

Mammalian Ependymin-related Protein; Up-regulated in Colon Cancer I (EPDR1)—EPDR1 was first identified as a potential Man-6-P glycoprotein in rat brain (3) and is a protein of unknown function that bears some similarity to fish ependymins, which are soluble proteins reported to be associated with extracellular matrix (44) and involved in memory consolidation (45). The human ortholog has been proposed to be a transmembrane protein (46), but here we verified in both mouse and human that EPDR1 is a soluble protein that contains Man-6-P. A lysosomal function for EPDR1 is supported by the observation that this protein, like many known lysosomal proteins, is also identified in neuromelanin granules from human brain (43) and by the fact that it is detected as an MPR affinity-purified protein that is secreted from MPR-deficient mouse fibroblasts (8).

Microsomal Stress 70 Protein ATPase (STCH)—STCH is a microsomal membrane-associated member of the heat shock protein 70 family (47) found here to contain a single N-linked glycosylation site containing Man-6-P. Given its likely function as an endoplasmic reticulum chaperone, the significance of Man-6-phosphorylation of STCH is not clear.

Myelin-associated Glycoprotein (MAG)—MAG is a minor but necessary component of myelin (48). As a sialic acid-binding lectin, we originally classified MAG as a likely specific contaminant (7) in human brain MPR-purified samples; however, we find here that this protein actually appears to contain Man-6-P.

N-Acetyllactosaminide β-1,3-N-Acetylglucosaminyltransferase (B3GNT6)—B3GNT6 is involved in the synthesis of poly-N-acetyllactosamine (49). It is a predicted type II transmembrane protein, although detection here as a soluble Man-6-P glycoprotein suggests that this assignment may not be correct or that B3GNT6 may be proteolytically cleaved into transmembrane and soluble fragments.

Neuroserpin (SERPIN11)—Not detected by LC-MS analysis of the total unfractionated mixture of MPR-purified proteins from human brain, two sites of Man-6-phosphorylation were identified for this protein. Missense mutations in SERPIN11 can result in polymeric aggregation of defective protein in cytoplasmic inclusions with a subsequent autosomal dominant familial encephalopathy (50). Identification of neuroserpin as a Man-6-P glycoprotein raises the intriguing possibility that such disease may actually represent a lysosomal storage disease.

Plasma α-Fucosidase (FUA2)—FUA2 is a secreted hydrolase that is genetically distinct from lysosomal tissue α-L-fucosidase (residing at locus 6q24 versus 1p34) but very similar at the sequence level (55% identical, 69% similar). A single site of Man-6-phosphorylation was identified here.

Proenkephalin A (PENK)—PENK is a glycosylated neuroendocrine precursor that is known to be glycosylated (51).

Prostaglandin H2 Isomerase (PTGDS)—PTGDS is a multifunctional protein that is responsible for the conversion of prostaglandin H2 to prostaglandin D2 (52) and has also been reported to be a retinoic acid transporter (53). Identification of a single Man-6-phosphorylation site that is conserved between mouse and human suggests a lysosomal localization for this protein, and in support of this, an electron microscopic study (54) detected PTGDS in lysosomes, multivesicular bodies, and endocytic vesicles in some cell types including meningeal macrophages and parenchymal microglia.

Ribonuclease T2 (RNASET2)—RNASET2, given its catalytic function, represents a promising candidate for an enzyme with lysosomal function.

SPARC-like Protein 1 (SPARCL1)—SPARCL1 is a secreted acidic protein thought to play a role in modulating lymphocyte adhesion to basement membranes (55).

Hypothetical Protein DKFZp313G1735, Telethon Sulfatase, Bone Rel Sulfatase (DKFZp313G1735)—DKFZp313G1735 is a predicted protein with homology to sulfatases. The presence of a number of other sulfatases within the lysosome is well documented; thus this protein represents another promising lysosomal candidate.

Sulfatase-modifying Factor 2 Precursor (SUMF2)—SUMF2 is a paralog of SUMF1, an enzyme responsible for generation of the unique amino acid C-formylglycine in the active site of a number of microsomal and lysosomal sulfatases and in which defects are responsible for multiple sulfatase deficiency. SUMF2 is known to be glycosylated at its only potential N-linked glycosylation site (56), which we found also to contain Man-6-P. Although most of the cellular SUMF2 appears to be located within the ER, it is possible that some proportion is transported to the lysosome to play a specific role in lysosomal sulfatase function.

Vasopressin-Neurophysin 2 Copeptin (AVP)—AVP is a glycosylated peptide hormone precursor (57).

DISCUSSION

A principal aim of this study was to develop and validate a mass spectrometric method for the direct demonstration of the Man-6-P modification on glycoproteins and for identification of the sites of modification. Given that the Man-6-P modification is characteristic of most soluble lysosomal proteins, these represent important steps in the characterization of the lysosomal proteome and will allow selection of candidate proteins for further, in-depth investigation of lysosomal localization. Although we identified a number of excellent lysosomal candidates (see “Results”), we also found Man-6-P on a number of proteins for which lysosomal function or localization appears highly unlikely. For example, several Man-6-P glycoproteins identified here have been proposed to
Mannose 6-Phosphorylation Sites

function in the ER, including STCH, SUMF2, and POFUT2, thus the significance of the Man-6-P modification is unclear. If these do actually represent authentic ER proteins, one possibility is that they may escape the ER retention apparatus and then become Man-6-phosphorylated in the transitional elements/cis-Golgi apparatus before retrieval back to the ER prior to encountering the MPRs in the trans-Golgi. Although the three proteins identified here lack the ER retrieval signal KDEL, the demonstration that cathepsin D engineered to contain this sequence is both ER-resident and Man-6-phosphorylated (58) provides general support for this possibility.

Sites of Man-6-phosphorylation have been experimentally determined for a number of individual lysosomal proteins, generally by site-directed mutagenesis of potential N-linked glycosylation sites of proteins expressed in vitro and analyses of the effects of these mutations on protein targeting and processing (24–29, 31). These studies have provided useful data on individual lysosomal proteins on a case-by-case basis but have drawbacks. They are limited to a small subset of lysosomal proteins and frequently rely on surrogate markers for Man-6-phosphorylation (e.g. secretion versus retention within the cell) that may not clearly identify sites where either glycosylation or the presence of Man-6-P may be variable modifications. An additional potential problem is that Man-6-phosphorylation and intracellular targeting of lysosomal proteins overexpressed in cell culture systems is not necessarily the same as seen in vivo (59, 60). Therefore, another major aim of this study was to globally map Man-6-phosphorylation sites of lysosomal proteins purified from a source that is more physiologically relevant.

We assigned a total of 85 individual sites of Man-6-phosphorylation to 40 known human lysosomal proteins and 48 sites to 31 mouse lysosomal proteins. Taken together, this represents a total of 99 sites of Man-6-phosphorylation representing 42 individual proteins. As would be predicted, there is a significant incidence of false negatives in terms of overall assignment of Man-6-P-containing glycoproteins from the glycopeptide analysis. We estimate the false negative incidence to be ~8 and 19% for human and mouse brain, respectively, based on the assignment of Endo H-deglycosylated MPR affinity-purified peptides to known lysosomal proteins. Failure to detect any given Man-6-phosphorylation site may simply reflect low peptide abundance. Alternatively the specific properties (e.g. sequence and size) of a given peptide may render it prone to selective loss during sample preparation or cause it to be missed by the mass spectrometer because it is poorly ionized or it exceeds the mass range constraints of the instrument. Such peptide-specific issues may be addressed in part by conducting different proteolytic digests, resulting in peptides with partially overlapping sequences that may increase the probability of detecting a glycosylation site. It is also worth noting that when identifying proteins containing Man-6-P, the relative probability of assigning at least one Man-6-P glycopeptide to a protein increases exponentially in proportion to the number of Man-6-P sites within this protein. For this reason, proteins containing a single site of Man-6-phosphorylation are particularly likely to represent false negatives.

In this study, we evaluated two different enzymes for the deglycosylation of purified glycopeptides, PNGase F and Endo H, and found that the latter has distinct advantages for our purposes. First, the N+203 mass increment observed after Endo H digestion results from an N-acetylgalactosamine residue that remains after the removal of the rest of an oligosaccharide; thus this is a highly specific marker for N-linked glycans that is difficult to incorrectly assign even when data are analyzed in an automated manner. In contrast, the N+1 modification associated with the removal of N-linked glycans by PNGase F can be confused with deamidation, and others have found it necessary to use isotopic labeling to distinguish between N+1 as a result of deamidation versus deglycosylation (14). In addition, automated data analysis can lead to incorrect assignment of N+1 to peptides that may not actually contain this modification (see “Results”). Second, Endo H provides an additional level of specificity in the analysis of Man-6-phosphorylation sites. Previous studies have found that Man-6-phosphorylation blocks further carbohydrate processing to Endo H-resistant structures, and thus the vast majority of Man-6-P-containing glycans are Endo H-sensitive (22, 23) (Fig. 2). PNGase F, in contrast, cleaves all N-linked oligosaccharides including complex type sugars that do not contain Man-6-P.

Manual examination of the spectra derived from glycopeptides deglycosylated with Endo H revealed that in many cases prominent peaks were present that represent peptides after neutral loss of the 203-Da N-acetylgalactosamine (Fig. 1). This represents another significant advantage of Endo H as the presence of this neutral loss ion provides a clear mass fingerprint for glycosylation. In addition, it may be possible in the future to exploit the presence of this characteristic neutral loss as a trigger for MS3 scans that should be specific for deglycosylated peptides.

One potential problem in using Endo H is that it is possible that some peptides containing a Man-6-phosphorylated oligosaccharide may also contain an additional Endo H-resistant glycan. In this case, the glycopeptide would not be detected using our current methods after Endo H treatment. However, in an analysis of MPR affinity-purified glycopeptides that were sequentially deglycosylated with Endo H and then PNGase F, we observed no peptides that contained both N+203 and N+1 in the context of NX(S/T) (data not shown). For our sample at least, the presence of Endo H-resistant glycans does not therefore appear to be a problem.

A number of previous studies have investigated protein determinants for the selection of carbohydrates for Man-6-phosphorylation by site-directed mutagenesis of proteins that

---

2 S. Kornfeld, personal communication.
acquire Man-6-P in cell culture, including DNase 1 (61, 62), cathepsin D (63–66), cathepsin L (67), aspartylglucosaminidase (68), cathepsin D/pepsinogen chimeras (69), and ariysulfatases (70–72). A general consensus of these studies is that spatially distinct amino acids, in particular lysine residues, form part of a broad three-dimensional patch on the surface of protein molecules that interacts with lysosomal protein UDP-N-acetylglucosamine phosphotransferase, orienting this enzyme for the recognition and phosphorylation of specific oligosaccharides (63, 69, 73). The generation of a global database of Man-6-phosphorylation sites in this study, especially when considered in conjunction with structural information, is likely to prove valuable in future studies of determinants for Man-6-phosphorylation and in validating the models arising from individual case studies.

Glycosylation is often a variable modification; thus an individual gene product may be represented both by molecules that are glycosylated at a given site and others that are not. For the known human lysosomal proteins, we found that 19 different peptides of the 99 that were assigned the N glycosylation site of NPC2 was not found to be glycosylated, the second always contains high mannose oligosaccharides, and the third is either not glycosylated or may contain high mannose or complex type oligosaccharides (74) (also data not shown). In this study, the first N-linked glycosylation site of NPC2 was not found to be glycosylated, the second was only detected with Endo-H- cleavable oligosaccharides, and the third was heterogeneous in terms of glycosylation (Fig. 1). These findings are in complete accordance with the previous observations.

The ability to recognize sites of Man-6-phosphorylation on lysosomal proteins may also have therapeutic implications for gene and enzyme replacement therapies for human lysosomal storage diseases. Central to the gene therapy of lysosomal storage diseases is the concept that not all cells will be transduced with a recombinant virus, but surrounding cells will be cross-protected by MPR-mediated uptake of recombinant protein secreted from the transduced cells. In developing gene therapy approaches, it may therefore be possible to optimize protein spread by adjusting the balance between secretion, lysosomal targeting, and uptake. Key to this balance is the level of Man-6-phosphorylation, and lysosomal proteins that contain physiological levels of Man-6-phosphorylated carbohydrates are generally efficiently targeted to the lysosome (75, 76). Expression of highly Man-6-phosphorylated recombinant proteins would promote efficient lysosomal targeting, rather than secretion, in transduced cells as well as highly efficient endocytosis in neighboring cells and would thus result in diminished spread. Selective mutagenesis of some (but not all) Man-6-phosphorylation sites, decreasing the efficiency of lysosomal targeting and thus promoting secretion from transduced cells and reduced uptake by adjacent cells, may therefore allow for greater intracellular spread before eventual MPR-mediated endocytosis and delivery to the lysosome. Similar arguments can be made for this sort of protein engineering to optimize enzyme replacement therapies that rely on MPR-mediated uptake and delivery of recombinant proteins. Conversely an understanding of protein determinants may allow for the increased selection of N-linked glycans for Man-6-phosphorylation; this could also be of benefit, depending on the disease and protein in question, especially for lysosome enzyme replacement therapies (77, 78). Detailed studies of individual lysosomal proteins will be required to thoroughly understand the effects of modulation of Man-6-phosphorylation in terms of protein structure and function and in terms of therapeutic efficacy.

Acknowledgments—We thank Christopher Gabel and Michel Jadot for helpful discussions and Henry Lackland for technical assistance.

* This study was supported by National Institutes of Health Grants DK054317 and S10RR017992 (to P. L.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence may be addressed: Center for Advanced Biotechnology and Medicine, 679 Hoes Lane, Piscataway, NJ 08854. E-mail: sleat@cabm.rutgers.edu.

‡ To whom correspondence may be addressed: Center for Advanced Biotechnology and Medicine, 679 Hoes Lane, Piscataway, NJ 08854. E-mail: lobel@cabm.rutgers.edu.

REFERENCES

1. Taylor, S. W., Fahy, E., and Ghosh, S. S. (2003) Global organellar proteomics. Trends Biotechnol. 21, 82–88
2. Holtzman, E. (1989) Lysosomes, Plenum Press, New York
3. Sleat, D. E., Sohar, I., Lackland, H., Majercak, J., and Lobel, P. (1996) Rat brain contains high levels of mannose-6-phosphorylated glycoproteins including lysosomal enzymes and palmitoyl-protein thioesterase, an enzyme implicated in infantile neuronal lipofuscinosis. J. Biol. Chem. 271, 19191–19198
4. Sleat, D. E., Kraus, S. R., Sohar, I., Lackland, H., and Lobel, P. (1997) γ-Glucosidase and N-acetylgalactosamine-6-sulfatase are the major mannose-6-phosphate glycoproteins in human urine. Biochem. J. 324, 33–39
5. Journet, A., Chapel, A., Kieffer, S., Louwagie, M., Luche, S., and Garin, J. (2000) Towards a human repertoire of monocytic lysosomal proteins. Electrophoresis 21, 3411–3419
6. Journet, A., Chapel, A., Kieffer, S., Roux, F., and Garin, J. (2002) Proteomic analysis of human lysosomes: application to monocytic and breast cancer cells. Proteomics 2, 1026–1040
7. Sleat, D. E., Lackland, H., Wang, Y., Sohar, I., Xiao, G., Li, H., and Lobel, P. (2005) The human brain mannose 6-phosphate glycoproteome: A complex mixture composed of multiple isoforms of many soluble lysosomal proteins. Proteomics 5, 1520–1532

3 H.-L. Liou and P. Lobel, unpublished data.
Mannose 6-Phosphorylation Sites

8. Kollmann, K., Mutenda, K. E., Balleininger, M., Eckermann, E., von Figura, K., Schmidt, B., and Lubke, T. (2005) Identification of novel lysosomal matrix proteins by proteome analysis. *Proteomics* 5, 3966–3978

9. Czupalla, C., Mansukoski, H., Riedl, T., Thiel, D., Krause, E., and Hoflack, B. (2006) Proteomic analysis of lysosomal acidic hydrolases secreted by osteoclasts: implications for tyrosic enzyme transport and bone metabolism. *Mol. Cell. Proteomics* 5, 134–143

10. Jadot, M., Lin, L., Sleat, D. E., Sohar, I., Huang, C. G., Pintar, J., Dubois, F., Coninck, S. W., Wattiaux–De Coninck, S., and Lobel, P. (1999) Subcellular localization of mannose 6-phosphate glycoproteins in rat brain. *J. Biol. Chem.* 274, 21104–21113

11. Sleat, D. E., Donnelly, R. J., Lackland, H., Liu, C. G., Sohar, I., Pullarkat, R. K., and Lobel, P. (1997) Association of mutations in a lysosomal protein, with classical late-infantile neuronal ceroid lipofuscinosis. *Science* 277, 1802–1805

12. Naureckiene, S., Sleat, D. E., Lackland, H., Fensom, A., Vanier, M. T., Wattiaux, R., Jadot, M., and Lobel, P. (2000) Identification of HE1 as the second gene of Niemann–Pick C disease. *Science* 290, 2298–2301

13. Bendtsen, J. D., Nielsen, H., von Heijne, G., and Brunak, S. (2004) Improved prediction of signal peptides. SignaPro 3.0. *J. Mol. Biol.* 340, 783–795

14. Kaji, H., Saito, H., Yamauchi, Y., Shinkawa, T., Taoka, M., Hirabayashi, J., Kimura, Y., Takasato, N., and Ishihara, T. (2003) Lectin affinity capture, isoform-coded tagging and mass spectrometry to identify N-linked glycoproteins. *Nat. Biotechnol.* 21, 667–672

15. Fan, X., She, Y. M., Bagshaw, R. D., Callahan, J. W., Schachter, H., and Mahuran, D. J. (2004) A method for proteomic identification of membrane-bound proteins containing Asn-linked oligosaccharides. *Anat. Biochem.* 322, 178–186

16. Zhang, L. H., Li, J., Martin, D. B., and Aebersold, R. (2003) Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol.* 21, 660–666

17. Craig, R., Cortens, J. P., and Beavis, R. C. (2004) Open source system for analyzing, validating, and storing protein identification data. *J. Proteome Res.* 3, 1234–1242

18. Lehn, D. E. A., and Gabel, C. A. (1989) Mannose processing is an important determinant in the assembly of phosphorylated high mannose-type oligosaccharides. *J. Biol. Chem.* 264, 5015–5023

19. Trowbridge, I. S., and Hyman, R. (1979) Abnormal lipid-linked oligosaccharides in class E Thy-1-negative mutant lymphomas. *Cell* 17, 503–508

20. Kornfeld, S., Gregory, W., and Chapman, A. (1979) Class E Thy-1 negative mouse lymphoma cells utilize an alternate pathway of oligosaccharide processing to synthesize complex-type oligosaccharides. *J. Biol. Chem.* 254, 11649–11654

21. Gabel, C. A., and Kornfeld, S. (1982) Lysosomal enzyme-phosphorylation in the expression of active human beta-1,3-N-acetylglucosaminyltransferase that is essential for poly-N-acetyllactosamine synthesis. *Science* 21104–21113

22. Varki, A., and Kornfeld, S. (1980) Structural studies of phosphorylated high mannose-type oligosaccharides. *J. Biol. Chem.* 255, 10847–10858

23. Palmitoyl- and Farnesyl-protein thioesterase deficiency in the U.S. *J. Biol. Chem.* 271, 11216–11225

24. Pradel, G., Schachner, M., and Schmidt, B. (1999) Inhibition of memory consolidation by antibodies against cell adhesion molecules after active avoidance conditioning in zebrafish. *J. Neurobiol.* 39, 197–206

25. Gregorio-King, C. C., McLeod, J. L., Collier, F. M., Collier, G. R., Bolton, K. A., Van Der Meer, G. J., Apostolopoulos, J., and Kirkland, M. A. (2002) Identification of a lysosomal membrane protein that is essential for the processing of asparagine-linked oligosaccharides. *J. Biol. Chem.* 278, 11289–11294

26. Ottesen, G. A. F., Gunn, G. A., Coxon, A., Johnston, P. G., and Kaye, F. J. (1994) A chimeric mouse stem cell factor gene encoding a functional human stem cell factor receptor. *Eur. J. Immunol.* 24, 2471–2476

27. Cramer, C., and Ballard, F. J. (1983) Inhibition of pyruvate carboxylase degradation and total protein breakdown by lysosomotropic agents in 3T3-L1 cells. *Biochem. J.* 210, 845–853

28. Holmberg, D. A., and Schriver, C., Beaudet, A., Sly, W., and Valle, D., eds (8th Ed., pp. 3935–3962, McGraw-Hill, New York

29. Vorum, H., Hager, H., Christensen, B. M., Nielsen, S., and Honore, B. (1999) Human calumenin localizes to the secretory pathway and is secreted to the medium. *Exp. Cell Res.* 248, 473–481

30. Davis, R. L., Shrimpton, A. E., Holohan, P. D., Bradshaw, C., Feiglin, D., Collins, G. H., Sonderegger, P., Kinter, J., Becker, L. M., Labcan, F., Krasnewich, D., Muenke, M., Lawrence, D. A., Yerby, M. S., Shaw, C. M., and McDermid, H. E. (2000) The human homolog of insect-derived factor II receptor. *Oncogene* 22, 5436–5445

31. Luo, Y., and Halttunen, R. S. (2005) C-Fucosylation of notch occurs in the endoplasmic reticulum. *J. Biol. Chem.* 280, 11289–11294

32. Tribi, F., Gerlach, M., Marcus, K., Asan, E., Tatschner, T., Arzberger, T., Meyer, H. E., Bringmann, G., and Riederer, P. (2005) “Subcellular proteomics” of neuromelanin granules isolated from the human brain. *Mol. Cell. Proteomics* 4, 945–957

33. Shashoua, V., Hesse, G. W., and Milinazzo, B. (1990) Evidence for the in vivo polymerization of epdymyn: a brain extracellular glycoprotein. *J. Cell Biol.* 118, 1219–1225

34. Sasaki, K., Kurata-Mura, K., Uijta, M., Angata, K., Nakagawa, S., Sekine, S., Nakata T., and Fukuda, M. (1997) Expression cloning of CD26 encoding a human beta-1,3-N-acetylgalactosaminyltransferase that is essential for poly-N-acetyllactosamine synthesis. *Proc. Natl. Acad. Sci. U. S. A.* 94, 14294–14299

35. Davis, R. L., Shirampton, A. E., Holohan, P. D., Bradshaw, C., Feiglin, D., Collins, G. H., Sonderegger, P., Kinter, J., Becker, L. M., Labcan, F., Krasnewich, D., Muenke, M., Lawrence, D. A., Yerby, M. S., Shaw, C. M., and McDermid, H. E. (2000) Identification of acylacyl hydrolase, a lipopolysaccharide-deactivating enzyme, in the murine urinary tract. *Infect. Immun.* 72, 3171–3178

36. Kant, A. F., Nieuwenhuizen, A. J., Hofmann, S. L., and Vonkova, V. (1994) Evidence for the in vivo polymerization of epdymyn: a brain extracellular glycoprotein. *J. Cell Biol.* 118, 1219–1225

37. Konttinen, Y. T., and Freden, L. (1982) Bone formation in vitro by osteogenic cells derived from human bone. *Proc. Natl. Acad. Sci. U. S. A.* 79, 7077–7081

38. Ortega, A., and Martínez, E. (2000) Expression of CD26 encoding a human beta-1,3-N-acetylgalactosaminyltransferase that is essential for poly-N-acetyllactosamine synthesis. *Proc. Natl. Acad. Sci. U. S. A.* 94, 14294–14299
Mannose 6-Phosphorylation Sites

Gooptu, B., Elliott, P. R., Finch, J. T., Carrell, R. W., and Lomas, D. A. (1999) Familial dementia caused by polymerization of mutant neuroserpin. Nature 401, 376–379

51. Thomas, G., Thomas, L., Thorne, B. A., and Herbert, E. (1988) Cell-type specific processing of neuroendocrine precursor proteins using vaccinia recombinants. Biochimie (Paris) 70, 89–97

52. Nagata, A., Suzuki, Y., Igarashi, M., Eguchi, N., Toh, H., Urade, Y., and Hayashi, O. (1991) Human brain prostaglandin D synthase has been evolutionarily differentiated from lipophilic-ligand carrier proteins. Proc. Natl. Acad. Sci. U. S. A. 88, 4020–4024

53. Tanaka, T., Urade, Y., Kimura, H., Eguchi, N., Nishikawa, A., and Hayashi, O. (1997) Lipocalin-type prostaglandin D synthase (β-trace) is a newly recognized type of retinoid transporter. J. Biol. Chem. 272, 15789–15795

54. Nagasaka, T., Hiraide, M., Sugimoto, T., Shindo, K., Shiozawa, Z., and Yokota, S. (2004) Localization of lipocalin-type prostaglandin D synthase in rat brain: immunoelectron microscopic study. Histochem. Cell Biol. 121, 483–491

55. Girard, J. P., and Springer, T. A. (1995) Cloning from purified high endothelial venule cells of hevin, a close relative of the antiadhesive extracellular matrix protein SPARC. Immunity 2, 113–123

56. Mairappan, M., Preusser-Kunze, A., Balleininger, M., Eiseit, N., Schmidt, B., Gande, S. L., Wenzel, D., Dierks, T., and von Figura, K. (2005) Expression, localization, structural, and functional characterization of pFGE, the paralog of the Cu-formylglycine-generating enzyme. J. Biol. Chem. 280, 15173–15179

57. Ivell, R., Schmalle, H., and Richter, D. (1981) Glycosylation of the arginine phosphorylation. I. Protein recognition determinants in both lobes of the insulin-like growth factor II mannose 6-phosphate receptor. J. Biol. Chem. 256, 393–403

58. Pelham, H. R. (1988) Evidence that luminal ER proteins are sorted from secreted proteins in a post-ER compartment. EMBO J. 7, 913–918

59. Lee, S. J., and Nathans, D. (1988) Proliferin secreted by cultured cells binds to mannose 6-phosphate receptors. J. Biol. Chem. 263, 3521–3527

60. Kovacina, K. S., Steele-Perkins, G., Purcho, A. F., Loubin, M., Miyazono, K., Heldin, C. H., and Roth, R. A. (1989) Interactions of recombinant and platelet transforming growth factor-β1 precursor with the insulin-like growth factor II mannose 6-phosphate receptor. Biochem. Biophys. Res. Commun. 160, 393–403

61. Nishikawa, A., Gregory, W., Frenz, J., Cacia, J., and Kornfeld, S. (1997) The phosphorylation of bovine DNase I Aα-linked oligosaccharides is dependent on specific lysine and arginine residues. J. Biol. Chem. 272, 19408–19412

62. Nishikawa, A., Nanda, A., Gregory, W., Frenz, J., and Kornfeld, S. (1999) Identification of amino acids that modulate mannose phosphorylation of mouse DNase I, a secretory glycoprotein. J. Biol. Chem. 274, 19309–19315

63. Baranski, T. J., Koelsch, G., Hartsuck, J. A., and Kornfeld, S. (1991) Mapping and molecular modeling of a recognition domain for lysosomal enzyme targeting. J. Biol. Chem. 266, 23365–23372

64. Baranski, T. J., Cantor, A. B., and Kornfeld, S. (1992) Lysosomal enzyme phosphorylation. I. Protein recognition determinants in both lobes of procathepsin D mediate its interaction with UDP-GlcNAc:lysosomal enzyme N-acetylglycosamine-1-phosphotransferase. J. Biol. Chem. 267, 23342–23348

65. Cuozzo, J. W., Tao, K., Cygler, M., Mort, J. S., and Sahagian, G. G. (1998) Lysine-based structure responsible for selective mannose phosphorylation of cathepsin D and cathepsin L defines a common structural motif for lysosomal enzyme targeting. J. Biol. Chem. 273, 21067–21076

66. Steet, R., Lee, W. S., and Kornfeld, S. (2005) Identification of the minimal lysosomal enzyme recognition domain in cathepsin D. J. Biol. Chem. 280, 33318–33323

67. Cuozzo, J. W., Tao, K., Wu, Q. L., Young, W., and Sahagian, G. G. (1995) Lysine-based structure in the proregion of procathepsin L is the recognition site for mannose phosphorylation. J. Biol. Chem. 270, 15611–15619

68. Tikkanen, R., Peltola, M., Ononen, C., Rouvinen, J., and Peltonen, L. (1997) Several cooperating binding sites mediate the interaction of a lysosomal enzyme with phosphotransferase. EMBO J. 16, 6684–6693

69. Baranski, T. J., Faust, P. L., and Kornfeld, S. (1990) Generation of a lysosomal enzyme targeting signal in the secretory protein pepsinogen. Cell 63, 281–291

70. Schiera, A., Dietz, F., Lange, H., Schestag, F., Parastar, A., and Gieselmann, V. (1999) Interaction of arylsulfatase A with UDP-N-acetylglucosamine:lysosomal enzyme-N-acetylglycosamine-1-phosphotransferase. J. Biol. Chem. 274, 3651–3658

71. Dittmer, F., and von Figura, K. (1999) Phosphorylation of arylsulfatase A occurs through multiple interactions with the UDP-N-acetylglucosamine-1-phosphotransferase proximal and distal to its retrieval site by the KDEL receptor. Biochem. J. 340, 729–736

72. Yaghoobfam, A., Schestag, F., Dierks, T., and Gieselmann, V. (2003) Recognition of arylsulfatase A and B by the UDP-N-acetylglycosamine: lysosomal enzyme N-acetylglycosamine-phosphotransferase. J. Biol. Chem. 278, 32653–32661

73. Cuozzo, J. W., and Sahagian, G. G. (1994) Lysine is a common determinant for mannose phosphorylation of lysosomal proteins. J. Biol. Chem. 269, 14490–14496

74. Chikh, K., Rodriguez, C., Vey, S., Vanier, M. T., and Millat, G. (2005) Niemann-Pick type C disease: subcellular location and functional characterization of NPC2 proteins with naturally occurring missense mutations. Hum. Mutat. 26, 20–28

75. Ludwig, T., Munier-Lehmann, H., Bauer, U., Hollinshead, M., Ottv, C., Lobei, P., and Hoffack, B. (1994) Differential sorting of lysosomal enzymes in mannose 6-phosphate receptor-deficient fibroblasts. EMBO J. 13, 3430–3437

76. Kasper, D., Dittmer, F., von Figura, K., and Pohlmann, R. (1996) Neither type of mannose 6-phosphate receptor is sufficient for targeting of lysosomal enzymes along intracellular routes. J. Cell Biol. 134, 615–623

77. Zhu, Y., Li, X., Kyazike, J., Zhou, Q., Thurbeg, B. L., Raben, N., Mattaliano, R. J., and Cheng, S. H. (2004) Conjugation of mannose 6-phosphate containing oligosaccharides to acid a-glucosidase improves the clearance of glycogen in Pompe mice. J. Biol. Chem. 279, 50336–50341

78. Zhu, Y., Li, X., McVie-Wylie, A., Jiang, C., Thurbeg, B. L., Raben, N., Mattaliano, R. J., and Cheng, S. H. (2005) Carbohydrate-remodelled acid alpha-glucosidase with higher affinity for the cation-independent mannose 6-phosphate receptor demonstrates improved delivery to muscles of Pompe mice. Biochem. J. 389, 619–628