DOI 10.1002/pmic.200401054 *Proteomics* 2005, *5,* 1520–1532

REGULAR ARTICLE

The human brain mannose 6-phosphate glycoproteome: A complex mixture composed of multiple isoforms of many soluble lysosomal proteins

David E. Sleat[1, 2], Henry Lackland[1], Yanhong Wang[1], Istvan Sohar[1], Gang Xiao[1],** Hong Li[3] and Peter Lobel[1, 2]*

[1] Center for Advanced Biotechnology and Medicine, University of Medicine and Dentistry of New Jersey, Piscataway, NJ, USA
[2] Department of Pharmacology, University of Medicine and Dentistry of New Jersey, Piscataway, NJ, USA
[3] Department of Biochemistry and Molecular Biology, New Jersey Medical School, Newark, NJ, USA

The lysosome is a membrane delimited cytoplasmic organelle that contains at least 50 hydrolytic enzymes and associated cofactors. The biomedical importance of these enzymes is highlighted by the many lysosomal storage disorders that are associated with mutations in genes encoding lysosomal proteins, and there is also evidence that lysosomal activities may be involved in more widespread human diseases. The aim of this study was to characterize the human brain lysosomal proteome with the goal of establishing a reference map to investigate human diseases of unknown etiology and to gain insights into the cellular function of the lysosome. Proteins containing mannose 6-phosphate (Man6-P), a carbohydrate modification used for targeting resident soluble lysosomal proteins to the lysosome, were affinity-purified using immobilized Man6-P receptor. Fractionation by two-dimensional electrophoresis resolved a complex mixture comprising approximately 800 spots. Constituent proteins in each spot were identified using a combination of matrix-assisted laser desorption/ionization-time of flight mass spectrometry (both mass spectrometry (MS/MS) and tandem peptide mass fingerprinting) on in-gel tryptic digests and *N*-terminal sequencing. In a complementary analysis, we also analyzed a tryptic digest of the unfractionated mixture by liquid chromatography MS/MS. In total, 61 different proteins were identified. Seven were likely contaminants associated with true Man6-P glycoproteins. Forty-one were known lysosomal proteins of which 11 have not previously been reported to contain Man6-P. An additional nine proteins were either uncharacterized or proteins not previously reported to have lysosomal function. We found that the human brain Man6-P-containing lysosomal proteome is highly complex and contains more proteins with a much greater number of individual isoforms than found in previous studies of Man6-P glycoproteomes.

Received: July 27, 2004
Revised: September 1, 2004
Accepted: October 1, 2004

Keywords:
Brain / Human / Lysosome / Proteome

Correspondence: Dr. Peter Lobel, Center for Advanced Biotechnology and Medicine, 679 Hoes Lane, Piscataway, NJ 08854, USA
E-mail: lobel@cabm.rutgers.edu
Fax: +1-732-235-5289

Abbreviations: ACTH, adrenocorticotropic hormone; **LSD**, lysosomal storage disease; **Man6-P**, mannose 6-phosphate; **MPR**, mannose 6-phosphate receptor; **NCBI**, National Center for Biotechnology Information; **sCI-MPR**, soluble cation-independent mannose 6-phosphate receptor

1 Introduction

The lysosome is a cytoplasmic organelle containing over 50 hydrolytic enzymes and associated proteins that act in concert towards the cellular degradation of macromolecules

* **Additional corresponding author:** Dr. David Sleat:
E-mail: sleat@cabm.rutgers.edu
** Current address: Genomics Institute, University of Pennsylvania, Philadelphia, PA 10104, USA

including polypeptides, carbohydrates, lipids, and nucleic acids [1]. The importance of these proteins in cellular catabolism and molecular recycling is highlighted by the fact that mutations in the genes encoding ~75% of these proteins cause different hereditary lysosomal storage diseases (LSDs). The clinical expression of LSDs is varied, reflecting the wide range of biochemical defects, but there are common features. LSDs are typically recessive, progressive, and are often neurodegenerative in nature. The classical hallmark of LSDs is the lysosomal accumulation of unhydrolyzed substrate for the defective enzyme as well as storage of other material secondary to lysosomal dysfunction. Interestingly, there are a number of hereditary human disorders of unknown molecular etiology possessing some characteristics of LSDs that may result from defects in genes encoding lysosomal proteins yet to be discovered. Some potential candidates include Kufs' disease (OMIM # 204300), geleophysic dysplasia (231050), ethanolaminosis (227150), and Yunis-Varon syndrome (216340). In addition, there are also numerous individual cases of apparent LSDs of unknown cause.

While some LSDs are due to deficiencies in resident transmembrane proteins or in proteins involved in lysosomal biogenesis and vesicular trafficking, the majority of the characterized diseases involve defects in soluble lumenal lysosomal proteins that are targeted to the lysosome by the mannose 6-phosphate (Man6-P) pathway. Here, *N*-linked oligosaccharides on newly synthesized lysosomal proteins receive the Man6-P modification and bind to Man6-P receptors (MPRs) that redirect their subcellular trafficking from the secretory pathway to the lysosome. In most tissues, the Man6-P targeting signal is rapidly removed in the lysosome, but in some cell types such as neurons, the lysosomal proteins retain the Man6-P modification [2, 3].

We have developed methods that use a soluble form of the bovine cation-independent MPR (sCI-MPR) as a highly specific affinity reagent for the visualization and purification of lysosomal proteins containing the Man6-P targeting signal [2, 4]. The ability to purify Man6-P-containing glycoproteins with high specificity has enabled studies aimed at the identification of novel lysosomal proteins from many sources. Originally, we focused on rat brain as a rich source of Man6-P glycoproteins [2] but subsequent studies by ourselves and others have also examined human brain [5], urine [6], and induced secretions from cultured monocytes [7, 8] and breast cancer cells [8]. These studies have not only provided basic information about lysosomal function but have also provided important clues towards gene defects in human disease, leading to the discovery of the defective genes underlying two fatal neurodegenerative diseases: classical late infantile neuronal ceroid lipofuscinosis [5] and Niemann-Pick type C2 disease [9].

In this report, we provide a 2-D map of the human brain lysosomal proteome and report identification of 61 proteins representing known lysosomal Man6-P glycoproteins, potentially novel lysosomal Man6-P glycoproteins and proteins thought to be specifically associated with and copurifying with Man6-P glycoproteins. We find that the human brain proteome is highly complex and contains more lysosomal proteins with a much greater number of individual isoforms than found in other studies of Man6-P glycoproteomes. These results should provide a useful reference for comparative proteomic study of LSDs and should also prove useful in the investigation of more widespread conditions where lysosomal dysfunction has been implicated such as cancer, Alzheimer disease, arthritis, and aging.

2 Materials and methods

2.1 Materials and reagents

Frozen human autopsy brain was obtained from the National Disease Research Interchange (Philadelphia, PA, USA). Affigel 15 was from Bio-Rad (Hercules, CA, USA); all protease inhibitors were from Sigma (St. Louis, MO, USA) except Pefabloc, which was from Pentafarm (Basel, Switzerland). For 2-DE : carrier ampholytes and non-linear pH IPG strips were from Amersham Biosciences (Piscataway, NJ, USA); iodoacetamide, urea, DTT, and CHAPS were from Sigma; Surfact-Amps Triton X-100 was from Pierce (Rockford, IL, USA); thiourea was from Fluka (Hannover, Germany); ASB-14, Zwittergent, and TCEP were from Calbiochem (San Diego, CA, USA). Other reagents were routinely obtained from Sigma and Bio-Rad.

2.2 Purification and immobilization of sCI-MPR

Soluble bovine cation-independent MPR (sCI-MPR) was purified from fetal calf sera by affinity chromatography on immobilized phosphomannan followed by gel filtration chromatography at both neutral and acid pH [4]. Gel electrophoresis of purified sCI-MPR revealed no contaminants other than a very minor amount of serum albumin [10]. Coupling of sCI-MPR to Affigel 15 was as described [2].

2.3 Affinity purification of human brain Man6-P glycoproteins

Human brain (100 g) was homogenized using a Polytron with a 20 mm generator (Brinkmann, Westbury, NY, USA) in 10 volumes w/v of homogenization buffer (20 mM sodium phosphate, pH 6.8, 150 mM NaCl, 5 mM β-glycerophosphate, 2.5 mM EDTA, 1 μg/mL pepstatin, and 1 μg/mL leupeptin) containing 1% Triton X-100 and 1 mM Pefabloc. The homogenate was centrifuged at 25 000 × *g* for 1 h at 4°C, then the supernatant was removed. The pellet was re-extracted by homogenization in five volumes of the same buffer and centrifugation repeated. Supernatants were pooled and then applied to an affinity column of immobilized sCI-MPR (~250 mg in 50 mL bed volume) that was equilibrated with homogenization buffer. The column was flow washed with five volumes of homogenization buffer containing 1% Triton

X-100, disassembled, and the resin resuspended and batch washed three times with two volumes homogenization buffer containing 1% Triton X-100 and then washed three times with two volumes homogenization buffer (without Triton X-100). The column was reassembled, flow-washed with 14 volumes homogenization buffer, and washed with 4 bed volumes of homogenization buffer containing 10 mM Man6-P to specifically elute Man6-P glycoproteins. Fractions of interest were identified by protein and lysosomal enzyme activity assays. Pooled fractions were concentrated and exchanged into 100 mM ammonium bicarbonate buffer using Centriprep 10 spin concentrators. Aliquots were stored at −80°C and volatile salts removed by repeated cycles of lyophilization and resuspension in water before use in downstream steps.

2.4 2-DE

Several protocols were used for 2-DE. The gels in Figs. 1–3 were run by the Swiss-2-D service at the Two-Dimensional Gel Electrophoresis Laboratory of Geneva, Switzerland, as described (http://us.expasy.org/ch2d/protocols/protocols.fm6.html). 2-DE was conducted "in-house" (Fig. 4 and data not shown) on pH 3–10 IEF gradients using either precast gels (Invitrogen, Carlsbad, CA, USA) or non-linear Immobiline DryStrip (Amersham Biosciences). Data was archived to the 2-D gels using Phoretix 2-D image analysis software (Nonlinear USA, Durham, NC, USA).

2.5 Protein identification by *N*-terminal microsequencing

Sample preparation and *N*-terminal Edman degradation sequencing was performed essentially as described earlier [2] using either an Applied Biosystems (Foster City, CA, USA) Model 477A protein/peptide sequencer with Model 120A analyzer or an Applied Biosystems 494 Procise sequencer.

2.6 Tryptic digestion of protein samples

Gel slices or spots excised from 2-D gels were digested with modified trypsin (Promega, Madison, WI, USA) according to the method used by the W.M. Keck Biotechnology Resource Laboratory (http://keck.med.yale.edu/prochem/geldig3.htm). Samples were incubated at 37°C for 24 h then stored at 4°C.

2.7 Protein identification by MALDI-TOF

Recrystallized CHCA was dissolved at a final concentration of 26 mM in 50% ACN containing 0.1% TFA and 10 mM monobasic ammonium phosphate [11]. Matrix and tryptic digests were mixed in a ratio of 1:2 and 1 μL spotted onto 100- or 192-well stainless steel or 400-well MALDI plates with hydrophobic coating. Each spot included 20 fmol bradykinin (monoisotopic mass 1060.5692 Da) and 120 fmol 18–39 clip of adrenocorticotropic hormone (ACTH, monoisotopic mass 2465.1989 Da) added as internal calibrants to the matrix. Automated data acquisition was achieved using an Applied Biosystems Voyager-DE Pro MALDI-TOF mass spectrometer operated in positive-ion reflectron mode or using an Applied Biosystems 4700 Proteomics Analyzer mass spectrometer operated in positive-ion reflectron mode for MS and 1 kV positive-ion mode without collision gas for MS/MS.

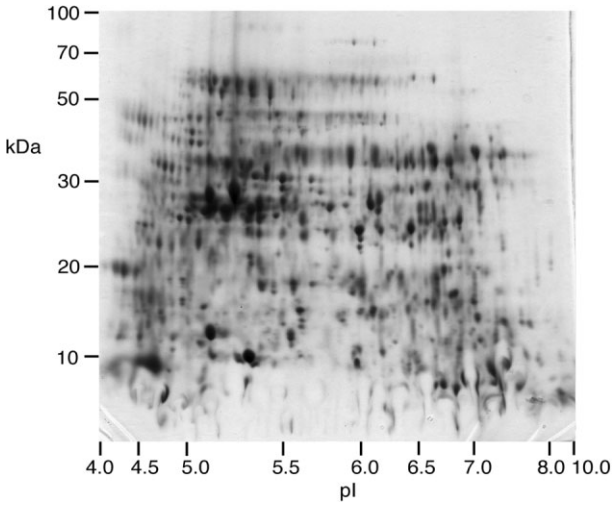

Figure 1. Coomassie-Blue staining of 1 mg of human brain Man6-P glycoproteins fractionated by 2-D PAGE using a pH 3.5–10 non-linear IPG gradient. Approximate positions of MW and p*I* markers were extrapolated from equivalent gels. This reference gel is illustrative of Man6-P glycoprotein preparations from brain autopsy specimens from over 10 different human subjects.

2.8 Protein identification by LC-MS/MS

Human brain Man6-P glycoproteins were denatured in guanidine HCl, reduced with DTT, and alkylated with iodoacetamide using the procedure outlined in Glazer *et al.* [12]. The mixture was buffer exchanged into 20% ACN in aqueous 0.1% TFA by gel filtration chromatography on a Superdex Peptide column (Amersham Biosciences). Fractions containing protein were pooled, dried, resuspended to 3 mg/mL using 50 mM ammonium bicarbonate pH 8.0, and digested at 30°C for 16 h using 0.2 μg trypsin (Promega) *per* 100 μg Man6-P glycoproteins. For analysis using a Finnigan LTQ, data were collected from a single experiment where a 6 μg sample was fractionated, absorbed onto a 300 μm × 5 mm C18 trap column (Agilent, Palo Alto, CA, USA) , washed, and fractionated on an RP column (75 μm × 10 cm, 15 μm tip; New Objective, Woburn, MA, USA) using a 90-min linear gradient of 0–65% ACN in 0.1% aqueous formic acid at a flow rate of 200 nL/min. MS spectra were acquired using dynamic fill time and were followed by MS/MS analysis of the three most intense peaks. Redundant data acquisition was reduced using dynamic exclusion for 40 s with a repeat count of two. A similar analysis was conducted using an

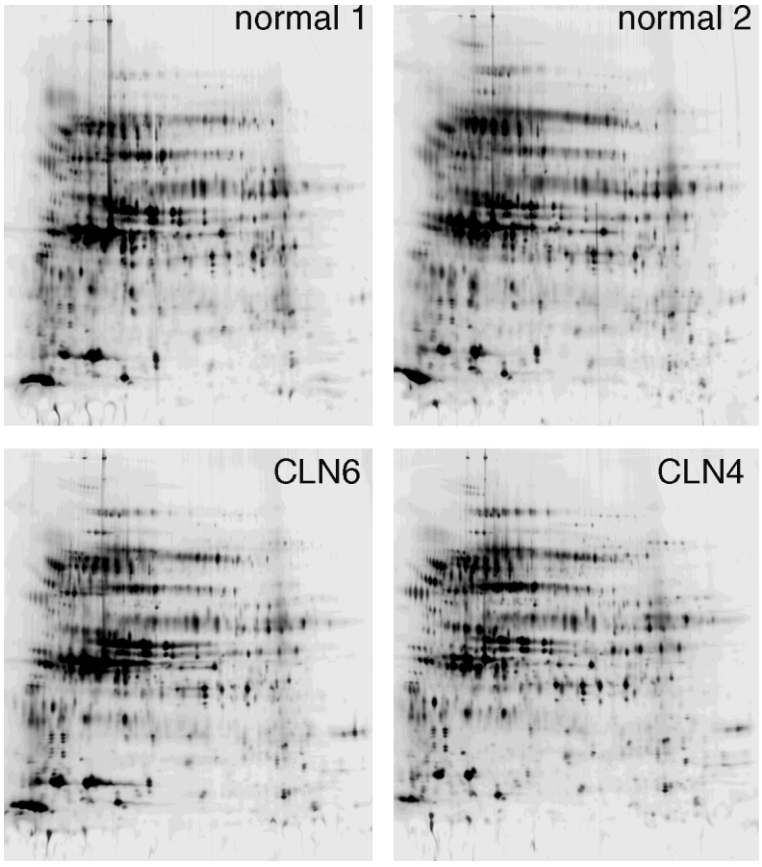

Figure 2. Silver-staining of 30 μg human brain Man6-P glycoproteins derived from normal controls or patients with variant late-infantile (CLN6) or adult neuronal ceroid lipofuscinosis (CLN4).

Applied Biosystems QSTAR Pulsar, except data were collected and merged from two experiments. Briefly, samples (0.9 μg) were absorbed onto a 300 μm × 1 mm C18 trap column, washed, and fractionated on a 75 μm × 15 cm column (LC Packings Sunnyvale, CA, USA) and a New Objective emitter using either a 1 or 2 h gradient of 5–40% ACN in aqueous 0.1% formic acid. Data collection cycles consisted of 1 s acquisition of MS data followed by MS/MS on the four most intense peaks (2 s *per* MS/MS) with a dynamic exclusion of 40 s.

2.9 Data analysis

MALDI-TOF MS spectra acquired using the Voyager-DE PRO were analyzed manually using the MS-Fit algorithm of the Protein Prospector suite [13] after processing and calibration with Data Voyager version 4.0 (Applied Biosystems). Spectra were also analyzed on an automated basis using the Profound PMF search engine of the Knexus software suite (Genomic Solutions, Ann Arbor, MI, USA). Both manual and automated searches used the NCBI non-redundant database restricted to *Homo sapiens* with up to one missed cleavage site allowed and cysteine carbamidomethylation and methionine oxidation as fixed and variable modifications, respectively. Calibration was achieved using internal standards (bradykinin and 18–39 clip of ACTH) and, if detected, the 842.5100 Da fragment of trypsin. In general, criteria for identification were the matching of at least four peptides with a mass accuracy of 20 ppm or less. However, given the limitations of PMF in identifying multiple proteins within a mixture, and the fact that many of the spots in this study contained more than one species (Table 1), analysis of PMF data was to some extent knowledge-based, with higher scoring "hits" sometimes overlooked in favor of more likely hits (*e.g.*, known soluble lysosomal proteins), and the subsequent researching of unmatched peaks to identify different components. Combined MALDI-TOF MS and -MS/MS data obtained using the ABI 4700 Proteomics Analyzer were used to search the NCBI non-redundant *H. sapiens* database using the MASCOT search engine [14] of the GPS Explorer software (Applied Biosystems). Proteins whose highest confidence identification achieved both protein and ion confidence scores of 100% were considered unambiguous assignments by MALDI-TOF MS/MS. MS/MS data obtained using the QSTAR and the LTQ mass spectrometers were searched against the NCBI 34d (ENSEMBL) human proteome database using Tandem, which is part of the Global Proteome Machine, an open source interface for analyzing tandem mass spectra against eukaryote genomes (www.thegpm.org). Threshold values for significance were assigned as the highest score achieved for a random identification ($\log(E) = -2.7$ for both the LTQ and QSTAR) when searching a database consisting of the reversed human NCBI 34d proteome.

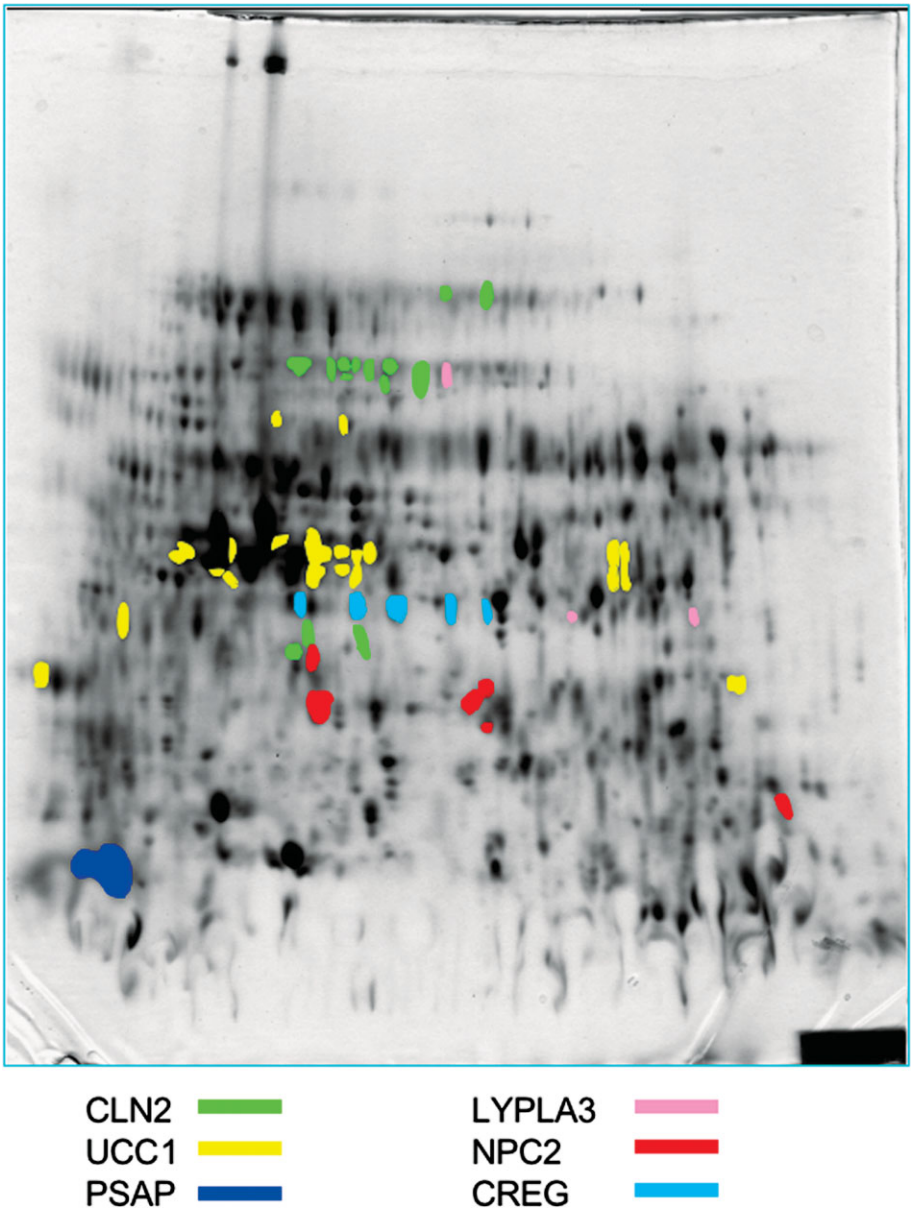

Figure 3. Identification of select Man6-P glycoproteins after fractionation by 2-D PAGE.

Table 1. Number of proteins identified *per* spot

Proteins/spot[a)]	Number of spots	Number of proteins
1	314	314
2	166	332
3	84	252
4	16	64
5	1	5
Unidentified	117	
Keratin	34	
No spectra obtained	44	
Total	776	967

a) Excluding contaminants unless stated otherwise.

We consider protein identifications to be unambiguous when made with more than one method (LC-MS using the LTQ, LC-MS using the QSTAR, MALDI-TOF PMF, MALDI-TOF PMF + MS/MS, or chemical sequencing).

3 Results and discussion

3.1 A 2-D map of the human brain lysosomal proteome

Coomassie-Blue staining of human brain Man6-P glycoproteins fractionated by 2-DE revealed a rich pattern of polypeptides, with ~1000 spots representing proteins ranging in

mass from ~10 to ~100 kDa distributed across the IEF gradient albeit with a bias towards the more acidic pHs (Fig. 1). Evaluation of four individual brain autopsy samples indicates that this complex pattern of isoforms appears to be highly reproducible (Fig. 2).

The complexity of the human brain proteome is considerably greater than that of Man6-P proteomes from either monocytic or breast cancer cells, from which about 50 individual polypeptides were resolved by 2-DE [7, 8]. This reflects our choice of human brain as the source for lysosomal protein purification as the source most likely to yield the greatest coverage of the human lysosomal proteome, given that this tissue contains more (in both a quantitative and qualitative sense) Man6-P glycoproteins than other tissues. In addition, some previous studies have also focused on Man6-P glycoproteins secreted from cultured cells [7, 8]; such proteins are not proteolytically processed to the extent that they would be within the lysosome and this also contributes to the relative lack of complexity of these proteomes. One disadvantage of choosing such a complex mixture here for analysis is that spot series resulting from multiple isoforms of individual lysosomal proteins tend to overlap. In practical terms, this complexity ruled out identification strategies that reduce the number of spots analyzed by pooling all members of a spot series or analyzing only the most abundant member. In addition, the fact that many spots contained more than one species (out of 776 analyzed spots, 267 contained more than one protein that met our identification criteria; many others probably contained additional, unidentified species judging from the number of unassigned peaks; Table 1) meant that a combination of protein identification approaches was required for unambiguous assignments: chemical sequencing, MALDI-TOF MS, and MALDI-TOF MS/MS analysis of proteins fractionated by 2-DE and LC-MS/MS of the unfractionated mixture.

3.2 Approaches to the identification of human brain Man6-P glycoproteins

In total, 61 different proteins excluding obvious contaminants (see below) were unambiguously identified in the Man6-P glycoprotein mixture purified from human brain (Table 2). Of these, 46 were identified by 2-DE while all but one were identified by LC-MS. These results suggest that LC-MS/MS is a better method to identify peptide components in a complex mixture compared to 2-D gels combined with MALDI-TOF MS/MS albeit information regarding protein processing is lost.

These observations illustrate some of the pitfalls present in a proteomic analysis approach incorporating 2-DE but one advantage of this approach is that it provides useful information regarding the distribution and frequency of different isoforms. Of the 776 2-D gel spots analyzed by MALDI-TOF, 581 yielded clear identifications of one or more proteins (Tables 1 and 2). In total, 46 different proteins were identified yielding a total of 967 different isoforms. The distribution of select proteins is illustrated in Fig. 3 and 2-D distribution maps for all proteins are presented in online supplementary material Figs. S1–S46. Only seven proteins were identified in single spots and at the other extreme, one protein (acid ceramidase) was identified as a component in 124 individual spots. The heterogeneity of purified Man6-P glycoproteins can be largely attributed to variations in both proteolytic and carbohydrate processing and is discussed for each individual protein in online supplementary material, Section 1. Approximately half of the spots contained a single identifiable noncontaminant protein while the rest contained between two and five identifiable species (Table 1).

Coverage for each Man6-P glycoprotein by PMF is presented in online supporting material Table 1 and ranged from 6.0 to 34.4% (Voyager DE Pro) and 2.9 to 41.4% (4700 Proteomics Analyzer) of primary translation products. Lysosomal proteins are extensively proteolytically processed and different spot series tend to represent different proteolytic fragments or chains, thus coverage of the entire primary translation product is not necessarily a good indicator of confidence of assignment. Relative PMF coverage from the *N*-terminus of the most *N*-terminal peptide to the *C*-terminus of the most *C*-terminal peptide was higher, ranging from 6.8 to 75.8% (Voyager DE Pro) and 8.3 to 100.0% (4700 Proteomics Analyzer). When determined using Edman degradation, the *N*-terminal sequence of individual isoforms is presented in online supporting material Table 2.

3.3 Human Man6-P glycoproteins

Individual distribution maps and details of processing for each identified protein are presented in online supplementary material. Based on the source, relative abundance, localization, presence of potential *N*-linked glycosylation sites, and biological properties of the proteins identified, we have classified them into a number of broad categories (Table 2). In some cases, this classification should be regarded as tentative in the absence of more detailed studies.

3.3.1 Contaminants

Several proteins were identified that we considered obvious contaminants. Some represent small amounts of highly abundant, well-characterized non-lysosomal proteins that nonspecifically bind to and leach from the column (*e.g.*, tubulin, actin, and serum albumin). Other contaminants include bovine sCI-MPR leached from the affinity column, and others (*e.g.*, keratins and porcine trypsin) which are introduced during the handling and processing of samples. These proteins are not listed in Table 2 and are not considered further.

We identified ferritin heavy chain and this may well represent a contaminant. However, we have included it in Table 2 as there is evidence that it may be glycosylated [15] and deficiencies in ferritin immunoreactivity have been observed

in Niemann-Pick C1 and C2 diseases [16]; thus this protein could conceivably represent a true ligand for the MPR.

A number of proteins were identified that probably represent contaminants that are specifically eluted from the column with Man6-P by interaction with true Man6-P glycoproteins rather than with the immobilized MPR (Table 2, class 1 proteins). We include in this category, non-lysosomal protease inhibitors cystatins B and C, which are most likely copurified with a lysosomal cysteine protease (*e.g.*, cathepsin B or H). Angiotensinogen may also have been purified by virtue of its properties as a serine protease inhibitor, although it does contain *N*-linked carbohydrates [17] and could be a true Man6-P glycoprotein.

We have also included in this category of contaminants, several lectins that probably bind to and copurify with true Man6-P glycoproteins. These include galactoside-binding galectin-1 and F-box only 2 (FBXO2). Both contain a potential *N*-linked glycosylation site but neither is predicted to have a signal sequence. FBXO2 is one of a group of proteins that together comprise one of the subunits of the ubiquitin protein ligase complex and is thought to specifically promote ubiquitinylation of *N*-glycosylated proteins by binding to high mannose oligosaccharides [18]. S-phase kinase-associated protein 1A, which contains neither a predicted signal sequence nor potential *N*-linked glycosylation sites, is thought to interact with F-box proteins and may have purified in a complex with FBXO2 [19]. We also classified another lectin, myelin-associated glycoprotein (MAG), as a class 1 contaminant although it is glycosylated so it could represent a true Man6-P glycoprotein.

3.3.2 Known lysosomal proteins containing Man6-P

Thirty-two of the proteins identified during the course of this analysis were well-characterized lysosomal proteins that are known to contain Man6-P. These are indicated in Table 2 and properties reported in the literature are detailed in the online supplementary material.

Table 2. Proteins identified in the affinity-purified mixture from human brain

Protein description	Spots (2-D map)	Gene code	Localization	M6-P	Potential signal sequence	NXS/T	Class
Aspartylglucosaminidase	4 (S1)	AGA	Lysosomal	Yes	Yes	2	2
Angiotensinogen		AGT	Extracellular		Yes	4	1
Arylsulfatase A	30 (S2)	ARSA	Lysosomal	Yes	Yes	3	2
Arylsulfatase B	2 (S3)	ARSB	Lysosomal	Yes	Yes	6	2
Acid ceramidase	124 (S4)	ASAH	Lysosomal	Yes	Yes	6	2
Palmitoyl-protein thioesterase 1	32 (S5)	CLN1	Lysosomal	Yes	Yes	3	2
Tripeptidyl peptidase I	14 (S6)	CLN2	Lysosomal	Yes	Yes	5	2
CLN5		CLN5	Lysosomal		Yes[a)]	8	3
Clusterin		CLU	?		Yes	7	4
Cellular repressor of E1a-stimulated genes	5 (S7)	CREG	Extracellular?		Yes	3	4
Sialic acid specific 9-O-acetylesterase	69 (S8)	CSE-C	Not known		Yes	6	4
Cystatin C	3 (S9)	CST3	Secreted	No	Yes	0	1
Cystatin B	2 (S10)	CSTB	Cytoplasmic/nuclear	No	No	0	1
Di-*N*-acetyl-chitobiase	1 (S11)	CTBS	Lysosomal	No	Yes	4	3
Cathepsin C	2 (S12)	CTSC	Lysosomal	Yes	Yes	4	2
Cathepsin D	89 (S13)	CTSD	Lysosomal	Yes	Yes	2	2
Cathepsin F		CTSF	Lysosomal		Yes	5	3
Cathepsin H	1 (S14)	CTSH	Lysosomal	Yes	Yes	3	2
Cathepsin L	1 (S15)	CTSL	Lysosomal	Yes	Yes	2	2
Cathepsin Z	10 (S16)	CTSZ	Lysosomal		Yes	2	3
Deoxyribonuclease II	7 (S17)	DNASE2	Lysosomal		Yes	4	3
Dipeptidylpeptidase 7	78 (S18)	DPP7	Lysosomal		Yes	6	3
F-box only protein 2	1 (S19)	FBXO2	Cytoplasmic		No	1	1
Fc fragment of IgG binding protein		FCGBP	?		Yes	33	1
FLJ22662		FLJ22662	?		Yes	7	4
Ferritin, heavy polypeptide 1		FTH1	Cytoplasmic		No	1	1
Alpha-l-fucosidase	12 (S20)	FUCA1	Lysosomal	Yes	Yes	4	2
Alpha-glucosidase	44 (S21)	GAA	Lysosomal	Yes	Yes	8	2
N-acetyl-6-galactosamine sulfatase	13 (S22)	GALNS	Lysosomal	Yes	Yes	2	2
Gamma-glutamyl hydrolase	39 (S23)	GGH	Lysosomal		Yes	4	3
Alpha-galactosidase		GLA	Lysosomal	Yes	Yes	4	2

Table 2. Continued

Protein description	Spots (2-D map)	Gene code	Localization	M6-P	Potential signal sequence	NXS/T	Class
Beta-galactosidase	8 (S24)	GLB1	Lysosomal	Yes	Yes	7	2
N-acetyl-glucosamine-6-sulfatase	32 (S25)	GNS	Lysosomal	Yes	Yes	13	2
Beta-glucuronidase	8 (S26)	GUSB	Lysosomal	Yes	Yes	4	2
Hexosaminidase A	40 (S27)	HEXA	Lysosomal	Yes	Yes	4	2
Hexosaminidase B	47 (S28)	HEXB	Lysosomal	Yes	Yes	6	2
Iduronate 2-sulfatase	6 (S29)	IDS	Lysosomal	Yes	Yes	8	2
Alpha-l-iduronidase	3 (S30)	IDUA	Lysosomal	Yes	Yes	6	2
Galectin-1	1 (S31)	LGALS1	Nuclear		No	1	1
Legumain		LGMN	Lysosomal		Yes	5	3
Acid lipase A	18 (S32)	LIPA	Lysosomal	Yes	Yes	6	2
LOC196463	22 (S33)	LOC196463	Not known		Yes	6	4
Lysophospholipase 3	3 (S34)	LYPLA3	Lysosomal		Yes	4	3
Myelin associated glycoprotein		MAG	Plasma membrane		Yes	10	1
Alpha-mannosidase	13 (S35)	MAN2B1	Lysosomal	Yes	No	11	2
Epididymis specific alpha mannosidase	15 (S36)	MAN2B2	Extracellular?		Yes	12	4
Beta-mannosidase	1 (S37)	MANBA	Lysosomal	Yes	Yes	8	2
Myeloperoxidase		MPO	Lysosomal	Yes	Yes	6	2
Alpha-*N*-acetylgalactosaminidase	8 (S38)	NAGA	Lysosomal	Yes	Yes	6	2
Alpha-*N*-acetylglucosaminidase	5 (S39)	NAGLU	Lysosomal	Yes	Yes	7	2
Niemann-Pick disease, type C2	6 (S40)	NPC2	Lysosomal	Yes	Yes	3	2
Plasma glutamate carboxypeptidase	40 (S41)	PGCP	Lysosomal		Yes	5	3
Protective protein for beta-galactosidase	10 (S42)	PPGB	Lysosomal	Yes	Yes	2	2
Prolylcarboxypeptidase	69 (S43)	PRCP	Lysosomal		Yes	7	3
Prosaposin	4 (S44)	PSAP	Lysosomal	Yes	Yes	5	2
Ribonuclease T2		RNASET2	Extracellular		Yes	3	4
Serine carboxypeptidase 1		SCPEP1	?		Yes	3	4
Sulfamidase	1 (S45)	SGSH	Lysosomal	Yes	Yes	5	2
S-phase kinase-associated protein 1A		SKP1A	?	No	No	0	1
Acid sphingomyelinase		SMPD1	Lysosomal	Yes	No	6	2
Upregulated in colorectal cancer gene 1	25 (S46)	UCC1	Extracellular?		Yes	4	4

Proteins identified are classified as: 1, probable contaminants; 2, known lysosomal proteins that are known to contain Man6-P; 3, known lysosomal proteins that have not been reported to contain Man6-P; 4, previously uncharacterized proteins or known proteins not previously thought to have lysosomal function. 2-D distribution maps refer to online supplementary material Figures S1–S46. Potential signal sequences and *N*-linked glycosylation sites were identified using the NetNGlyc 1.0 Server (www.cbs.dtu.dk/services/NetNGlyc/).

a) A signal sequence is predicted within the human CLN5 protein when an initiator methionine is chosen that corresponds to that of the rat and mouse orthologs. Translated human CLN5 database entries contain an *N*-terminal extension relative to other species that is derived from an upstream in-frame methionine and which is likely not to be biologically relevant.

3.3.3 Characterized lysosomal proteins whose Man6-phophorylation status is not known

There are a number of proteins that are either known or suggested to be lysosomal whose status in terms of Man6-phosphorylation is not known (to our knowledge). The identification of the following 11 of such proteins here strongly suggests that each contains Man6-P.

3.3.3.1 CLN5

Like CLN2 also [5], CLN5 is defective in a form of Batten disease [20], a fatal neurodegenerative disease that typically affects children. Originally proposed to be a transmembrane protein [20], more recent reports suggest that CLN5 is actually a glycosylated soluble lysosomal protein [21, 22] and this is consistent with its identification here as a potential Man6-P glycoprotein. This report also indicated that CLN5 expressed in transfected BHK cells contained endoglycosidase H-sensitive oligosaccharides, consistent with Man6-phosphorylation which blocks further processing to endoglycosidase H-resistant structures.

3.3.3.2 CTBS (di-*N*-acetyl-chitobiase)

CTBS is known to be lysosomal based upon its pH optima and subcellular fractionation [23].

3.3.3.3 CTSF (cathepsin F)

Based on its short half life at cytosolic pH, CTSF was suggested to be lysosomal [24].

3.3.3.4 CTSZ (cathepsin Z; cathepsin P; cathepsin X)

Based upon homology to other cathepsins, CTSZ was proposed to be lysosomal [25].

3.3.3.5 DNASE2 (deoxyribonuclease II)

A ~30 kDa isoform of DNASE2, a known lysosomal protein, was identified here. While the status of DNASE2 in terms of Man6-phosphorylation is unknown, we found several spots containing this protein alone that are specifically detected by radiolabeled MPR (Fig. 4), strongly suggesting that this protein is a true Man6-P glycoprotein.

3.3.3.6 DPP7 (dipeptidyl peptidase 7; quiescent cell dipeptidase)

DPP7 is known to be a glycoprotein [26] that localizes to the lysosome.

3.3.3.7 GGH (gamma glutamyl hydrolase)

GGH is a known lysosomal protein [27].

3.3.3.8 LGMN (legumain; asparaginyl endopeptidase)

LGMN was shown to be lysosomal by subcellular fractionation [28].

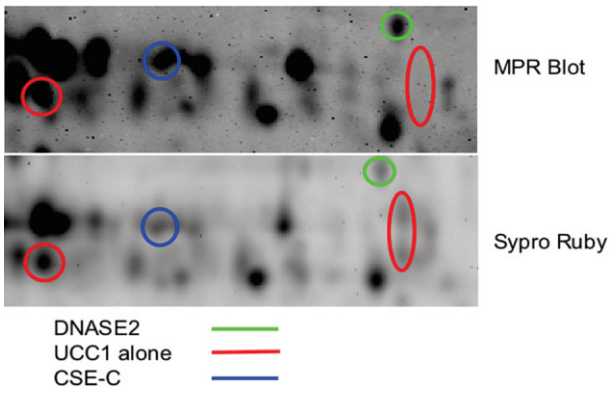

Figure 4. Sypro Ruby staining of a gel region with spots corresponding to DNASE2, UCC1, and CSE-C alone and a radiolabeled MPR blot of the corresponding region. Both spots containing DNASE2 and CSE-C are detected by radiolabeled MPR suggesting the presence of Man6-P. Not all spots containing UCC1 were detected by MPR blotting suggesting that not all isoforms contain Man6-P.

3.3.3.9 LYPLA3 (LCAT-like lysophospholipase, phospholipase A2)

LYPLA3 is glycosylated and was shown to be lysosomal by subcellular fractionation [29].

3.3.3.10 PGCP (plasma glutamate carboxypeptidase, lysosomal dipeptidase)

PGCP is the same as lysosomal dipeptidase [30]. Despite failure to detect PGCP in brain by Western blotting [31], we find PGCP to be one of the most abundant brain Man6-P glycoproteins, present in 40 individual spots. In placenta, it has been reported that PGCP forms a complex with protective protein/cathepsin A [31] (PPGB); thus the possibility exists that it does not actually contain Man6-P but is copurified with PPGB. However, this is not the case as: (i) PGCP is more abundant in brain than PPGB, and; (ii) we find spots corresponding to PGCP that are detected with a radiolabeled MPR probe in a blotting assay. PGCP is abundant in blood. However, it is also present within cells and the presence of Man6-P and its localization to a punctate cytoplasmic compartment together make it highly likely that it actually represents a lysosomal protease that, like many other lysosomal proteins, can be detected in serum and plasma.

3.3.3.11 PRCP (prolylcarboxypeptidase, angiotensinase C, pro-X carboxypeptidase)

PRCP has been shown to be lysosomal by subcellular fractionation [32, 33].

3.3.4 Potential new lysosomal proteins

A number of previously characterized and novel proteins were identified that are not known to have lysosomal function to our knowledge. In the absence of evidence that these may represent contaminants (see Section 3.1), we have classified the following as potentially lysosomal proteins based on the following criteria: evidence for glycosylation or presence of potential *N*-linked glycosylation sites; evidence for, or predicted, signal peptide; and function consistent with lysosomal function (*e.g.*, hydrolytic enzyme).

3.3.4.1 CLU (clusterin, complement lysis inhibitor, SP-40,40, sulfated glycoprotein 2, testosterone-repressed prostate message 2, apolipoprotein J)

CLU is an acidic sulfated glycoprotein thought to be involved in a number of cellular processes including complement activation and apoptosis. Its identification here as a likely Man6-P glycoprotein in brain, together with an earlier observation that cultured thyrocytes secrete a glycosylated form that binds the MPR [34], suggest that CLU may also have a lysosomal function although no enzymatic activity has been described for this protein.

3.3.4.2 CREG (cellular repressor of E1a-stimulated genes)

Previous studies have indicated that CREG contains Man6-P modified *N*-linked oligosaccharides [7, 35] although whether this protein was lysosomal is not clear [35]. No enzymatic activity has been described for this protein.

3.3.4.3 CSE-C (sialic acid-specific *o*-acetylesterase)

It was not previously known whether CSE-C contains Man6-P; however, we find here (Fig. 4) that it is detected by radiolabeled MPR suggesting that this is the case. The enzymatic activity of this protein is consistent with lysosomal function.

3.3.4.4 NM_024829 (hypothetical protein FLJ22662)

Hypothetical protein FLJ22662 was identified by LC-MS/MS alone and is 32% identical and 48% similar to hypothetical protein LOC196463 (see below). FLJ22662 is predicted to be 506 amino acids in length and like LOC196463, is in the laminin A family and has significant similarity (33% identical, 53% similar) to *D. discoideum* phospholipase B. FLJ22662 contains both a predicted signal peptide and potential *N*-linked glycosylation sites.

3.3.4.5 NP_775813 (hypothetical protein LOC196463)

Using a combination of approaches (chemical sequencing, PMF, tandem MALDI-TOF MS, and LC-MS) to search a conceptual translation database of the human genome, we identified a protein found in 22 different spots of 30–42 kDa to be a hypothetical protein, LOC196463. This protein appears to be evolutionarily conserved, with orthologs in mouse (83% identical, 91% similar), rat, *C. elegans, D. discoideum,* and there appears to be a predicted human paralog (FLJ22663; see above). The 589 amino acid protein is predicted to contain a 42-residue signal peptide. Based on *N*-terminal sequencing and MS coverage, there appears to be a cleavage site between amino acids 290 and 291. Assuming no additional cleavages except the signal cleavage, two chains with theoretical masses of 28 and 33 kDa, not including carbohydrates, are predicted. There are six potential *N*-linked glycosylation sites, three in each predicted chain. Significant resemblance to *D. discoideum* phospholipase B (36% identical, 52% similar) is suggestive of an enzymatic function and LOC196463 is considered a member of the laminin A protein family which is named for a type member (lama) expressed in *Drosophila* lamina glia (not be confused with the mammalian laminin family of matrix proteins). The function of LOC196463 and its paralog FLJ22662 and their lysosomal localization therefore require further investigation.

3.3.4.6 MAN2B2 (epididymis-specific alpha-mannosidase)

Previous studies [8] have noted that MAN2B2 contains Man6-P modified *N*-linked oligosaccharides but the lysosomal localization of MAN2B2 remains to be verified. The enzymatic activity of MAN2B2 is consistent with lysosomal function and it contains a predicted signal peptide and 12 potential *N*-linked glycosylation sites.

3.3.4.7 RNASET2 (ribonuclease 6)

RNASET2 was identified only by LC-MS/MS. Based upon its amino acid sequence [36] and similarities to other ribonucleases, RNASET2 is predicted to have a mass of ~25 kDa and to be a secreted glycoprotein with three potential *N*-linked glycosylation sites. Our data suggest that it may be a lysosomal enzyme.

3.3.4.8 SCPEP1 (serine carboxypeptidase 1, retinoid-inducible serine carboxypeptidase)

SCPEP1 has a predicted signal sequence and five potential *N*-linked glycosylation sites [37]. As a protease, SCPEP1 represents another promising candidate for a soluble lysosomal protein.

3.3.4.9 UCC1 (mammalian ependymin related protein)

UCC1 was identified in two spot series of ~43 and 27 kDa. UCC1 was first described as an unknown protein in rat brain (band 3 [2]) that allowed identification of human and mouse EST cDNAs and deduction of the amino acid sequence of the encoded protein (data not shown). UCC1 was subsequently identified as a transcript up-regulated in colon cancer [38] and down-regulated in hematopoetic cells [39] and appears to be related at the sequence level to a glycoprotein expressed in the CNS of fish, ependymin. UCC1 is predicted to have a signal sequence and has four potential *N*-linked glycosylation sites. The function of ependymin is not clear, but it has been implicated in memory consolidation and has been suggested to be a cellular adhesion molecule. In order to determine whether UCC1 is a true Man6-P glycoprotein or a contaminant, we examined the region of the 2-D gel containing this protein by MPR blotting. This analysis suggested that some isoforms of UCC1 contain Man6-P (Fig. 4) and thus it is possible that UCC1 has a novel lysosomal function although no enzymatic function has been described to date.

3.4 Post translational processing of human brain Man6-P glycoproteins

One outcome of this study has been the identification of a number of novel proteolytic isoforms of many more of the lysosomal proteins than have previously been characterized

(individual details for each protein identified are given in the online supplementary material). There are a number of possible explanations for this observation. First, we purify Man6-P containing proteins from brain, which contains a variety of cell types that could contribute to the complexity. For instance, in neurons, the Man6-P marker appears to be long-lived [2] in contrast to other cell types where the Man6-P modification is rapidly hydrolyzed and is mainly found on newly synthesized lysosomal proteins in transit to the lysosome. Given the diversity of cell types, it is possible that the sample contains mature, partially processed and unprocessed forms of lysosomal proteins, contributing to the many isoforms observed. Second, it is possible that some of these isoforms result from proteolysis during purification although measures are taken to prevent this occurring (*e.g.*, inclusion of protease inhibitors and EDTA, and conducting the purification at neutral pH at 4°C to inhibit lysosomal activities). Finally, proteolysis may have accompanied cell death and lysis during the period between the death of the donor and removal and freezing of the brain, a process not readily controllable when dealing with human autopsy samples. However, this seems unlikely given that the pattern of human brain Man6-P glycoproteins is highly reproducible from sample to sample (>10 individual brain autopsy specimens have been examined of which four are shown in Fig. 2); if this complexity did arise from partial sample degradation, some variability might be expected reflecting differences in sample handling (*e.g.*, time from patient death to autopsy and freezing of sample). Thus, while the underlying mechanisms for the complex pattern of lysosomal isoforms in brain remains to be clarified, this is a highly reproducible observation which is of significance given that one of the aims of this study was to generate a 2-D reference map for the investigation of lysosomal disease.

3.5 Clinical application of the human brain Man6-P glycoproteome

To date, the molecular bases for about 45 LSDs have been identified and the vast majority is associated with deficiencies in soluble, Man6-P containing lysosomal proteins. However, diagnosis of these diseases is often difficult as mutations of variable severity can result in a wide range of clinical manifestations and there are numerous cases of apparent LSD in which no defect can be detected by biochemical or molecular-genetic means. One aim of this study was therefore to provide useful tools for the diagnosis of difficult LSDs and in the discovery of defective genes in diseases of currently unknown etiology. We envisage that our results might find clinical application in two approaches. First, generation of a detailed 2-D map of the human brain lysosomal proteome should provide a useful resource for comparative proteomics using brain autopsy samples to investigate LSDs (*e.g.*, as in the identification of the defective gene in LINCL [5]) but also to investigate global lysosomal changes in more widespread disorders. Second, the identification of novel lysosomal proteins can provide candidates for unknown LSD gene (*e.g.*, as in the identification of the defective gene in Niemann-Pick C2 disease [9]).

4 Concluding remarks

One goal of our subcellular proteomic research is to identify all soluble Man6-P containing components of the human lysosome. In order to achieve this, it has been necessary to consider and optimize: (i) sources and protocols for Man6-P glycoprotein purification; (ii) stringent methods for data analysis that maximize filtering out of false positives, and; (iii) methods for data interpretation that filter out true positives that most likely do not represent lysosomal proteins. Our approach to each step is discussed below.

In terms of a source for an initial survey, we chose to analyze brain because of the quantitative and qualitative abundance of Man6-P glycoproteins. Even so, in a survey of different rat tissues [2] as well as our ongoing efforts in characterizing Man6-P glycoproteomes from bovine epididymis and human plasma, it is clear that the complement of Man6-P glycoproteins, while overlapping, varies from tissue to tissue. Complete analyses of Man6-P glycoproteomes from a number of different sources will be required to catalog the entire lysosomal proteome and in addition, investigation of individual cell types containing more specialized lysosome-like compartments such as the secretory granule of neutrophils and the acrosome of sperm may also extend the repertoire of known human Man6-P glycoproteins.

Identification of false positives is a significant problem in proteomic surveys and to minimize such occurrences, we have analyzed our protein sample set using several complementary approaches and instruments and have analyzed the resulting data using a number of different software tools. In terms of identification methods, we have used *N*-terminal chemical sequencing, PMF, PMF combined with tandem MALDI-TOF MS, and LC-MS/MS on both a QStar hybrid quadrupole-TOF mass spectrometer and an LTQ linear IT. PMF data was analyzed manually using MS-Fit (Protein Prospector) and on an automated basis using ProFound (Knexus). LC-MS/MS data from both instruments was analyzed using Tandem (Global Proteome Machine). The aim of this multiple approach was to focus upon proteins identified by more than one method.

Another significant problem is the differentiation between contaminants and true positives. Contaminants may simply represent abundant species that are not completely removed by washing the affinity column, exogenous proteins such as bovine sCI-MPR or porcine trypsin, or proteins that are specifically associated with lysosomal proteins (class 1). The latter can be very difficult to distinguish from the new lysosomal Man6-P glycoproteins whose identification was the aim of this study. We have attempted to do this by considering the known biological properties (function, cellular localization, presence of oligosaccharides, or potential

N-linked glycosylation sites, *etc.*) of each potential candidate. However, the unambiguous assignment of such proteins as lysosomal Man6-P glycoproteins requires additional experimental verification of both the Man6-P modification and lysosomal localization.

In some cases, we have been able to strengthen our conclusion that a given protein contains Man6-P by detection by radiolabeled sCI-MPR after 2-DE and transfer to NC (Fig. 4). However, this approach, which is likely to be valuable for less complex proteomes containing well separated spots by 2-DE, has several serious drawbacks when attempting to verify the presence of Man6-P in the highly complex mixture observed here. First, many of the spots contained more than one protein species and even spots that apparently comprised a single protein could possibly contain other species below the threshold of detection that might contribute to a signal observed with radiolabeled MPR blotting. Second, the transfer to NC results in a significant loss of gel resolution thus individual spots frequently become difficult to distinguish in complex areas of the gel. Third, many of the proteins of interest were detected by LC-MS of the unfractionated mixture but not by MALDI analysis of the 2-D gels. Fourth, absence of detection by radiolabeled MPR does not necessarily indicate a lack of Man6-P as such spots may represent non-phosphorylated or non-glycosylated subunits of a homo- or hetero-multimer.

Bearing these considerations in mind, we are currently developing mass spectrometric-based approaches to directly identify Man6-P containing glycopeptides to directly determine the sites of *N*-linked glycosylation. Quantitative MS approaches will also be useful to determine the subcellular localization of select candidate lysosomal proteins using the principles outlined by de Duve [40] for analytical cell fractionation, as have recently been applied to centrosomal proteins [41]. Such approaches will greatly simplify the validation of potential novel lysosomal proteins and represent an important next step in the interpretation of data arising from studies focusing on the purification and identification of Man6-P glycoproteins.

This study was supported by NIH grant DK054317 (P.L.). We thank Dr. Reiko Kiyonami and Dr. Terry Zhang (Thermo Electron, San Jose, CA) and Dr. Srinivasan Krishnan (Applied Biosystems, Framingham, MA) for performing LC-MS experiments.

5 References

[1] Holtzman, E., *Lysosomes*, Plenum Press, New York 1989.

[2] Sleat, D. E., Sohar, I., Lackland, H., Majercak, J., Lobel, P., *J. Biol. Chem.* 1996, *271*, 19191–19198.

[3] Jadot, M., Lin, L., Sleat, D. E., Sohar, I. *et al.*, *J. Biol. Chem.* 1999, *274*, 21104–21113.

[4] Valenzano, K. J., Kallay, L. M., Lobel, P., *Anal. Biochem.* 1993, *209*, 156–162.

[5] Sleat, D. E., Donnelly, R. J., Lackland, H., Liu, C. G. *et al.*, *Science* 1997, *277*, 1802–1805.

[6] Sleat, D. E., Kraus, S. R., Sohar, I., Lackland, H., Lobel, P., *Biochem. J.* 1997, *324*, 33–39.

[7] Journet, A., Chapet, A., Kieffer, S., Louwagie, M. *et al.*, *Electrophoresis* 2000, *21*, 3411–3419.

[8] Journet, A., Chapel, A., Kieffer, S., Roux, F., Garin, J., *Proteomics* 2002, *2*, 1026–1040.

[9] Naureckiene, S., Sleat, D. E., Lackland, H., Fensom, A. *et al.*, *Science* 2000, *290*, 2298–2301.

[10] Valenzano, K. J., Remmler, J., Lobel, P., *J. Biol. Chem.* 1995, *270*, 16441–16448.

[11] Zhu, X., Papayannopoulos, I. A., *J. Biomol. Tech.* 2003, *14*, 298–307.

[12] Glazer, A. N., Delange, R. J., Sigman, D. S., *Chemical Modification of Proteins*, Elsevier, New York 1975, p103.

[13] Clauser, K. R., Baker, P., Burlingame, A. L., *Anal. Chem.* 1999, *71*, 2871–2882.

[14] Perkins, D. N., Pappin, D. J., Creasy, D. M., Cottrell, J. S., *Electrophoresis* 1999, *20*, 3551–3567.

[15] Ghosh, S., Hevi, S., Chuck, S. L., *Blood* 2004, *103*, 2369–2376.

[16] Christomanou, H., Vanier, M. T., Santambrogio, P., Arosio, P. *et al.*, *Mol. Genet. Metab.* 2000, *70*, 196–202.

[17] Gimenez-Roqueplo, A. P., Celerier, J., Lucarelli, G., Corvol, P., Jeunemaitre, X., *J. Biol. Chem.* 1998, *273*, 21232–21238.

[18] Yoshida, Y., Chiba, T., Tokunaga, F., Kawasaki, H. *et al.*, *Nature* 2002, *418*, 438–442.

[19] Dias, D. C., Dolios, G., Wang, R., Pan, Z. Q., *Proc. Natl. Acad. Sci. USA* 2002, *99*, 16601–16606.

[20] Savukoski, M., Klockars, T., Homberg, V., Santavvori, P. *et al.*, *Nat. Genet.* 1998, *19*, 286–288.

[21] Isosomppi, J., Vesa, J., Jalanko, A., Peltonen, L., *Hum. Mol. Genet.* 2002, *11*, 885–891.

[22] Isosomppi, J., Vesa, J., Jalanko, A., Peltonen, L., *Neurobiol. Dis.* 2004, *16*, 29–40.

[23] Kuranda, M. J., Aronson, N. N., Jr., *J. Biol. Chem.* 1986, *261*, 5803–5809.

[24] Wang, B., Shi, G. P., Yao, P. M., Li, Z. *et al.*, *J. Biol. Chem.* 1998, *273*, 32000–32008.

[25] Santamaria, I., Velasco, G., Pendas, A. M., Fueyo, A., Lopez-Otin, C., *J. Biol. Chem.* 1998, *273*, 16816–16823.

[26] Chiravuri, M., Agarraberes, F., Mathieu, S. L., Lee, H., Huber, B. T., *J. Immunol.* 2000, *165*, 5695–5702.

[27] Wang, T. T., Chandler, C. J., Halsted, C. H., *J. Biol. Chem.* 1986, *261*, 13551–13555.

[28] Chen, J. M., Dando, P. M., Rawlings, N. D., Brown, M. A. *et al.*, *J. Biol. Chem.* 1997, *272*, 8090–8098.

[29] Hiraoka, M., Abe, A., Shayman, J. A., *J. Biol. Chem.* 2002, *277*, 10090–10099.

[30] Dolenc, I., Mihelic, M., *Biol. Chem.* 2003, *384*, 317–320.

[31] Gringas, R., Richard, C., EL-Alfy, M., Morales, C. R. *et al.*, *J. Biol. Chem.* 1999, *274*, 11742–11750.

[32] Matsunaga, M., Kira, J., Saito, N., Ogino, K., Takayasu, M., *Jpn. Circ. J.* 1968, *32*, 137–143.

[33] Yang, H. Y., Erdos, E. G., Chiang, T. S., *Nature* 1968, *218*, 1224–1226.

[34] Lemansky, P., Brix, K., Herzog, V., *Exp. Cell. Res.* 1999, *251*, 147–155.

[35] Di Bacco, A., Gill, G., *Oncogene* 2003, *22*, 5436–5445.

[36] Trubia, M., Sessa, L., Taramelli, R., *Genomics* 1997, *42*, 342–344.

[37] Chen, J., Streb, J. W., Maltby, K. M., Kitchen, C. M., Miano, J. M., *J. Biol. Chem.* 2001, *276*, 34175–34181.

[38] Nimmrich, I., Erdmann, S., Melchers, U., Chtarbova, S. *et al.*, *Cancer Lett.* 2001, *165*, 71–79.

[39] Gregorio-King, C. C., McLeod, J. L., Collier, F. M., Collier, G. R. *et al.*, *Gene* 2002, *286*, 249–257.

[40] De Duve, C., *J. Cell. Biol.* 1971, *50*, 20d–55d.

[41] Anderson, J. S., Wilkinson, C. J. Mayor, T., Mortensen, P. *et al.*, *Nature* 2003, *426*, 570–574.

