ISOLATION AND CHARACTERIZATION OF MEMBRANE VESICLES SECRETED BY HUMAN RENAL CELLS

A thesis submitted for the Degree of Doctor of Philosophy

By

MAYANK SARASWAT (MSc)

The research work described in this thesis was carried out under the supervision of

PROF. HARRY HOLTHOFER

SCHOOL OF BIOTECHNOLOGY

&

CENTRE FOR BIOANALYTICAL SCIENCES (CBAS)

DUBLIN CITY UNIVERSITY

2013

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of **PhD** is entirely my own work, and that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: _____ (Candidate) ID No.: 59100931 Date: 20/12/2012

Acknowledgement

It is a pleasure to acknowledge and thank all the people who have made this thesis possible. I would start with Prof. Harry Holthofer who is my supervisor. I would like to thank him for giving me a chance to work with him on my PhD. He has been a tremendous source of support and encouragement throughout my stay in Dublin City University. He was quick to point me to right direction with simultaneously giving me freedom to perform this work. Group and personal meetings with him were always inspiring and motivating. Discussions with him about my immediate work and other related area were fun as well as informative. I possibly can't thank him enough for supporting me.

My mentor in the lab, Dr. Luca Musante, was always open for helping me out all these years. I have learnt a lot about science from him including the meticulous experiment design. He is very through in his approach and his untiring love of science is a great source of motivation. He always offered me encouragement about new ideas and came up with several of his own to help design and execute the experiments. He always insisted on strengthening fundamental knowledge and rarely have I come across someone with so deep knowledge of biochemistry and related disciplines. I would like to thanks him for teaching me so much in so little time. My training wouldn't be complete without his support. I thoroughly enjoyed our coffee sessions in the afternoon where he would quiz everybody about fundamental biochemical concepts related to our work. He still keeps sending interesting studies from our field and biographies of great scientists. He was very considerate of my problems and I would like to thank him for everything he did to make me keep going. I cannot thank him individually for all the things he did for me and those are many, but I am grateful for all his support and help. He is a mentor in true sense of the word.

Another great source of strength both personal and professional was Dr. Alessandra Ravida who has since moved on in her career to a different group. However, we never felt it because she knows very well how to keep in touch which I think I lack somewhere. But I think she understands me better than others. All the time she was there in the lab she was always there to offer help and encouragement. She has been a great source of strength whenever I have needed it. I believe the attachment is mutual and I thank her for all the training and help she has given me. She and other have made my stay in lab without hassles and it wouldn't have been the same without her.

I would also like to thank Brian Shortt for being so generous and helpful. He is a great person and being on the same stage of our studies we have exchanged kind words and support with each other. I admire him personally and professionally and his constant offer of help whenever needed. I also wish him all the success ahead of the degree and he deserves it.

I would also like to thank Dr. Alberto Benito Martin for all his support and help. Our discussions about new ideas and other's work in the field were thoroughly enjoyable. He is a good person who always helps others out. Short breaks in the university with him always kept me going. He is a friend and a scholar and I wish him well for his future.

I would also like to thank Dorota Tataruch for her support in the lab. She made it so easier to work and tolerated my scattered belonging in the lab. She is always eager to help whenever she can. She is a efficient worker and manages labs well. I thank her for all the common pre-prepared stock solutions I found in the shelf. I wish her well for her future endeavours.

I would also like to thank Dr. Barry Byrne who has also moved to newer positions in the university. Throughout his stay in the lab, he was always happy to help and teach all techniques he knew. He was quick to crack a joke or two and always kept the environment light and enjoyable. I liked his easy-going nature and we collaborated on some projects

during his stay. I thank him for performing SPR assays for my work. He is a great person and a good colleague.

I would like to thank Prof. Martin Clynes, Director, NICB for his generous support for mass spectrometry services. I would also like to thank Dr. Paula Meleady for helping me in arranging sample analysis by mass spectrometry. My special thanks go to Mr. Michael Henry who was untiring in analysing our samples and also otherwise always supportive and forthcoming.

I would also like to thank Dr. Elodie duriez and Prof. Bruno Domon, Luxembourg clinical proteomics centre for mass spectrometry sample analysis.

I would also like to thank Brian Freeland who managed the building and all the maintenance and monitoring. He never made me realise how things actually work in the building because he always took care of everything from gas supply to liaising with university management and maintenance. He always made sure things are in order in the building so that all of us could work un-interrupted. He was also a good person to talk to and always offered kind words.

I also thank Dong feng who is an exchange student in the lab. He is a good person and nice to talk to.

I would also like to thank everyone in the building who made it easier to work and exchanged kind words often. This includes Bianca, Karen, Sharon, Siobhan, Tanja, Michael, Moira, Harriet, Roisin and Mary. I meant everyone in the building and apologise if I have missed anyone. I didn't mean to.

I also thank Aoifa, Mary C, Joan and Mary F for helping on administrative things whenever I needed.

I also thank all my friends, Rohit, Ashwini, Tushar, Sanjay, Mukund, Shikhar, Yuvraj, Ankur, Udit, Bana, Ambuj, Viji and Mohit for all the support they have given me. A special thanks to Chandra and Priyanka for their undying support.

In the end I would like to thank my family including my wife Shruti. She has supported me so much during this time that I cannot be thankful enough. She has kept me going through all my insecurities and fears and helped me get over them. I love her and thank her for making this endeavour of mine possible. My parents have been so supportive sitting literally seven seas apart and still offering me support and love. My father has taught me values which I still use in everyday life and he always pushed me to do more, go that extra distance. My mother is my icon and her unconditional love and sacrifices for me has given me strength to keep doing it. I thank my sister and her family to offer me help and support at all times. This work wouldn't have been possible without all of them.

Table of Contents

LIST OF ABBREVIATIONS

17-ODYA	17-Octadecynoic acid
2DE	Two-dimensional electrophoresis
AAL	*Aleuria aurantia* lectin
ABE	Acyl-biotin exchange
ACN	Acetonitrile
AFM	Atomic force microscopy
AGE	Advanced glycation end products
AKI	Acute kidney injury
AM-AFM	Amplitude modulation-Atomic force microscopy
ANPEP	Aminopeptidase N
APO	Apolipoprotein
AQP	Aquaporin
Asn	Asparagine
ATPase	Adenosine triphosphatse
BOG	Beta-octyl glucoside
C	CHAPS
CD	Cluster of differentiation
CDGS	Carbohydrate Deficiency Glycoprotein Syndrome
CETP	Cholesteryl ester transfer protein
CFTR	Cystic fibrosis transmembrane conductance regulator
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonic
CHMP	Charged multivesicular body protein
CL	Cholesterol
CM5	Carboxymethylated dextran sensor chip
Con-A	Concanavalin-A

COX-2	Cyclooxygenase 2
CR-1	Complement receptor-1
CVN	Cyanovirin-N
D	DTT
DBA	*Dolichos biflorus*
DC	Dendritic cells
DEF	Diatomaceous earth filter
DMSO	Dimethylsulfoxide
DNA	Deoxyribose nucleic acid
DOC	deoxycholate
DPP IV	Dipeptidyl peptidase IV
DSA	*Datura stramonium* agglutinin
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EDC	1-ethyl-3-(3-dimethylaminopropyl) carbodiimide
EEA	*Euonymus europaeus* lectin
ESCRT	Endosomal Sorting Complex Required for Transport
FasL	Fas ligand
FC	Flow cytometer
FCP	Fibrocystin
FDR	False discovery rate
FESEM	Field emission scanning electron microscopy
FLISA	Fluorophore-Linked Immuno Sorbent Assay
FLLA	Flurophore-linked lectin assay
FSC	Forward light scattering
FT	Fourier transformation

Gal	Galactose
GalNAc	N-acetyl galactosamine
GlcNac	N-acetyl glucosamine
GNA	*Galanthus nivalis* lectin
GPI	Glycosylphosphatidylinositol
GRFT	Griffithsin
GSA	*Griffonia smplicifolia* agglutinin
GTPase	Guanidine triphosphatase
HA	Hydroxylamine
HBS	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid buffered saline
HDL	High density lipoprotein
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HHL	Hippeastrum hybrid lectin
HIV-1	Human immunodeficiency virus-1
HMW	High molecular weight
HPDP-biotin	N-[6-(Biotinamido)hexyl]-3'-(2'-pyridyldithio)propionamide biotin
HPLC	High-performance liquid chromatography
Hrs	Hepatocyte-growth-factor-regulated tyrosine-kinase substrate
HSA	Human serum albumin
Hsp	Heat shock protein
IAA	Iodoacetamide
IAC	Immuno-affinity chromatography
IgA	Immunoglobulin A
IgG	Immunoglobulin G
IL	Interleukin
ILV	Intraluminal vesicles

IP	Immuno-precipitation
IPA	Ingenuity pathway analysis
LAC	Lectin affinity chromatography
LBA	*Phaseolus lunatus* lectin
LC	Liquid chromatography
LCA	*Lens culinaris* agglutinin
LDL	Low density lipoprotein
LHX1	LIM homeobox 1
LMW	low molecular weight
LPS	Lipopolysaccharide
MAC	Membrane attack complex
MAL	*Maackia amurensis* lectin
Man	Mannose
MeOH	Methanol
MHC	Major histocompatibility complex
MFG-E8	Milk-fat-globule EGF factor-VIII
miRNA	Micro ribonucleic acid
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MudPIT	Multidimensional protein identification technology
MV	Microvesicles
MVB	Multivesicular body
MWCO	Molecular weight "cut-off"
N-Rh-PE	1,2-dipalmitoyl-sn-glycero-3-phophoethanolamine-N-[lissamine rhodamine B sulfonyl]
NCC	Thiazide-sensitive NaCl co-transporter

NEDD-4	Neural precursor cell expressed developmentally down-regulated protein 4
NEM	N-ethyl maleimide
NEP	Nephrilysin
NeuAc	N-acetylneuraminic acid
NGAL	Neutrophil gelatinase-associated lipocalin
NHS	N-Hydroxysuccinimide
NKCC	Na-K-Cl cotransporter
NPA	*Narcissus pseudonarcissus* lectin
NTA	Nanoparticle tracking analysis
OMIM	Online mendelian inheritance in man
P	Pellet
pAsp	phosphoaspartic acid
PBST	Phosphate buffer saline Tween-20
PC	Phosphatidylcholine
PC-1	Polycystin-1
PE	Phosphatidylethanolamine
PHA-E	*Phaseolus vulgaris* erythroagglutinin
PHA-L	*Phaseolus vulgaris* leukoagglutinin
pHis	phosphoHistidine
PI	Phosphatidylinositol
PIP2	Phosphatidylinositol 4,5-bisphosphate
PKC	Protein kinase C
PM	Plasma membrane
PMN	Polymorphonuclear leukocytes
PNA	Peanut agglutinin
PNGase F	Peptide N-glycosidase F
PON-1	Paraoxanase-1

PrP	Prion protein
pS	phosphoSerine
PS	Phosphatidylserine
PSA	*Pisum sativum* agglutinin
PSCA	Prostate stem cell antigen
pT	phosphoThreonine
PTL-II	*Psophocarpus tetragonolobus* lectin II
PTMs	Post-translational modifications
pY	phosphoTyrosine
RBC	Red blood cell
RCA	*Ricinus communis* agglutinin
RCF	Relative centrifugal force
RhCG	Rhesus blood group C glycoprotein
RNA	Ribonucleic acid
RT	Room temperature
RU	Response unit
SAA	Serum amyloid A
SBA	Soybean agglutinin
SD	Standard deviation
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
Ser	Serine
Sia	Sialic acid
SJA	*Sophora japonica* agglutinin
SLC12A1	Solute carrier family 12 (sodium/potassium/chloride transporters), member 1
SM	Sphingomyelin
SN	Supernatant

SNA	*Sambucus nigra* lectin
SNHS	Sulfo-N-Hydroxysuccinimide
SPR	Surface plasmon resonance
SSC	Side scattering
STA	*Solanum tuberosum* lectin
STAM	Signal transduction adapter molecule
SuccWGA	Succinylated Wheat Germ agglutinin
SUMO	Small ubiquitin-like modifiers
SVN	Scytovirin
TCEP	Tris(2-carboxyethyl)phosphine
TCTP	Translationally controlled tumour protein
TEM	Transmission electron microscopy
TGN	Trans-golgi network
THP	Tamm-Horsfall glycoprotein
Thr	Threonine
TNF	Tumour necrosis factor
Tris	Tris(hydroxymethyl)aminomethane
TSAP6	Tumour suppressor activated pathway 6
TSG101	Tumour susceptibility gene 101
UEA-I	*Ulex europaeus* agglutinin I
Ub	Ubiquitin
Ubl	Ubiquitin-like proteins
UIMs	Ubiquitin interaction motifs
VLDL	Very low density lipoprotein
VPS	Vacuolar protein-sorting

WB	Western blotting
WGA	Wheat Germ agglutinin
ZP	Zona pellucida

ABSTRACT

ISOLATION AND CHARACTERIZATION OF MEMBRANE VESICLES SECRETED BY HUMAN RENAL CELLS

Mayank Saraswat

Most cells release membrane vesicles for various purposes including, but not limited to, intercellular communication and disposal of membrane and soluble proteins. These vesicles are secreted into urine coming from the cells lining the urinary tract and bladder epithelium. These vesicles are a promising source of biomarkers for various cardiovascular and renal diseases. This thesis pursues twofold objectives, one being the development and improvement of an isolation method for urinary membrane vesicles and the second being proteomic characterization of the content of these vesicles. These objectives are important to realise the clinical potential of these vesicles. An alternative method for removal of contaminant high-abundant proteins was developed which preserves the activity of vesicular proteins. Moreover, lipid-affinity and lectin-affinity-based novel methods to enrich membrane vesicles from minimally processed urine were evaluated and developed. More than 600 proteins were identified in urinary membrane vesicles using shotgun proteomic analysis. Post-translational modification (PTM) proteomics was carried out to identify the PTM status of vesicular proteins. Many different PTMs like glycosylation, ubiquitination and palmitoylation were assessed. Surface glycan profiles of these vesicles were elucidated using fluorophore-linked lectin assay (FLLA) employing 18 different lectins. Lectin blotting, lectin-affinity chromatography using multiple lectins and hydrazide chemistry based enrichment of glycoproteins were carried out. As a result, 108 glycoproteins were identified. Immuno-affinity chromatography was used to enrich and identify ubiquitin-conjugated proteins present in urinary membrane vesicles. A number of potential palmitoylated proteins were identified as well. Computational prediction and validation methods were applied to these protein lists. In conclusion, novel methods to isolate urinary membrane vesicles were developed. In addition, a thorough proteomic characterisation of contents of urinary membrane vesicles was achieved. This work will serve as platform for further characterization of urinary membrane vesicles.

CHAPTER 1

REVIEW OF LITERATURE

1.1 Introduction

Urine is a combination of plasma filtrate and the secretion profile of cells lining the urino-genital tract. In healthy individuals, the origin of approximately 70% of the urinary proteome is kidney and the rest represents the plasma filtered by the glomerulus (Thongboonkerd & Malasit, 2005). Proteins present in urine are a collection of proteins secreted by a number of cell types and/or tissues so it is not possible to correlate the mRNA abundance of any particular tissue with the urine proteome, as is the case of other body fluids like plasma (Kawamoto *et al.*, 1996). This leaves proteomic analysis as the only option to reflect the disease state of an organism. Therefore, the urinary proteome might serve as a rich source of biomarkers for uro-genital and systemic diseases which have been reviewed previously (Pisitkun, Johnstone & Knepper, 2006). Moreover, urine collection is a non-invasive procedure which makes it an ideal source for biomarker discovery. Kidneys are located on each side of the vertebral column in the posterior part of the abdomen (Brenner, 1996). These organs maintain the water, electrolyte and pH balance of the human body (Knepper & Burg, 1983). Each kidney is enveloped by a fibrous capsule and contains a renal artery and a renal vein (Brenner, 1996).

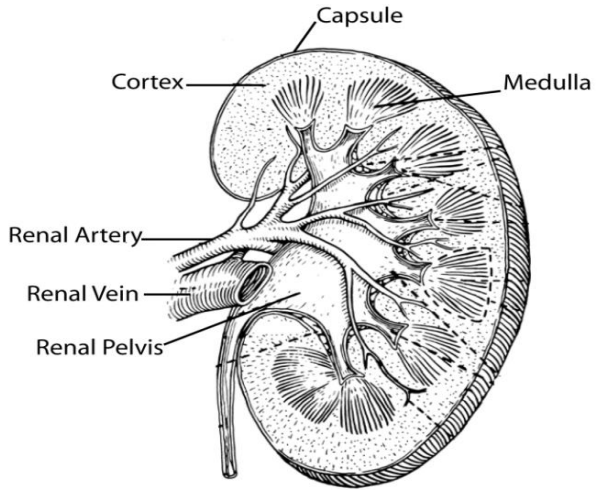

Figure 1.1: Schematic diagram of bisected kidney (Brenner, 1996)

Kidney can be divided into cortex and medulla as can be seen in Figure 1.1, showing the bisected kidney. Nephrons are the functional unit of kidney and each kidney consists of approximately 1 million nephrons. Figure 1.2 displays the schematic diagram of the nephron elements.

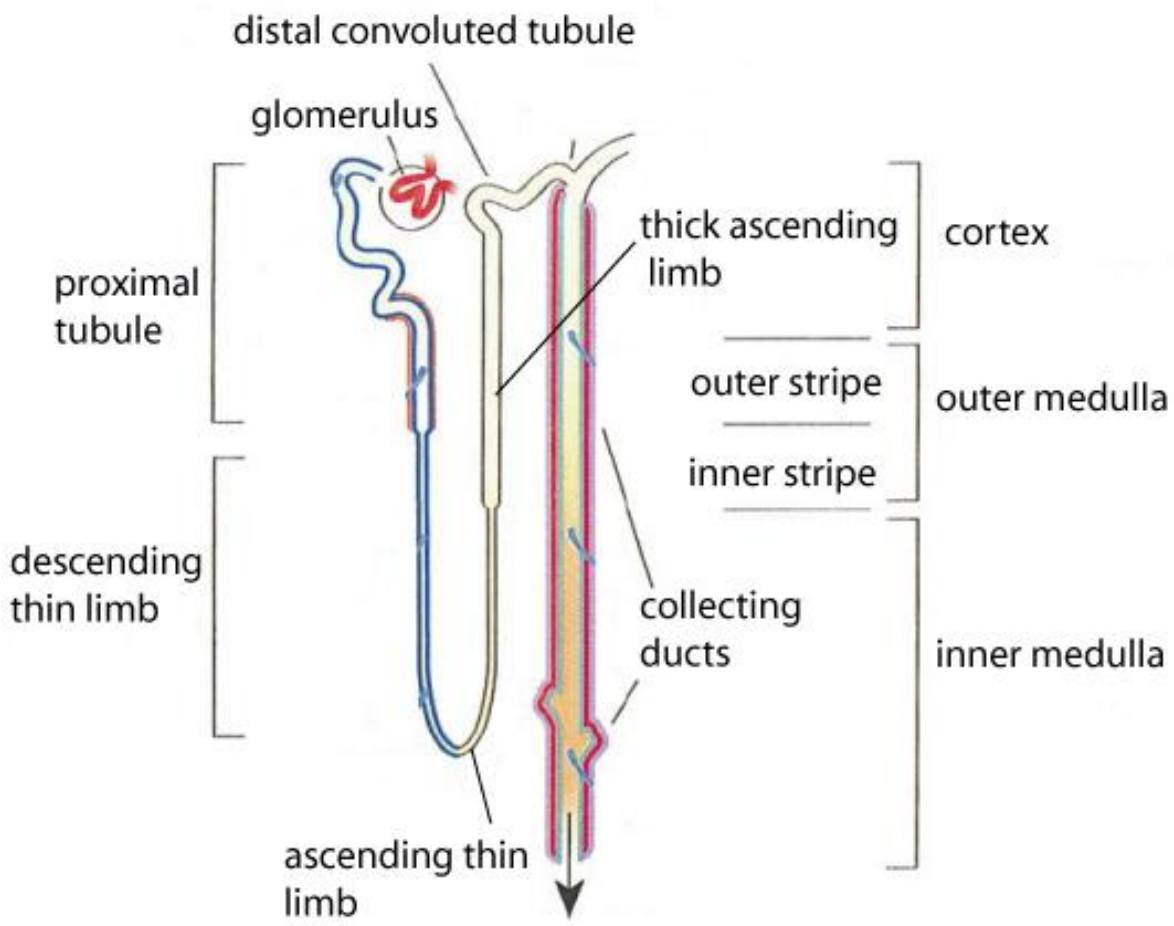

Figure 1.2: Schematic diagram of a nephron with its major segments (Nielsen *et al.*, 2002).

Glomeruli are found exclusively in the kidney cortex while tubular segements are found in the cortex, outer medulla and inner medulla, as can be seen in the Figure 1.2. The proximal tubule consists of segments of the convoluted and straight tubules and the loops of Henle, which is made up of a thin descending as well as an ascending limb and a thick ascending limb. This is followed by collecting ducts which consist of inner and outer medullar collecting ducts (Brenner, 1996). Kidney diseases are mostly due to genetic, inflammatory,

infectious or metabolic causes. Urinary proteins, interestingly, are derived from all segments of nephrons as well as the bladder epithelium and, thus, are expected to reflect respective changes in the characteristic urinary protein signature. Approximately 3% of urinary proteins are secreted encapsulated within membrane vesicles (Hoorn *et al.*, 2005).

1.2 Membrane vesicles

Extracellular spaces of multicellular organisms contain a special set of cell secretion products including ions, metabolites, proteins and complex carbohydrates. It has recently been established that they also contain membrane vesicles secreted by various cells. Current research endeavours mainly focus on two types of membrane vesicles: microvesicles and exosomes whose primary characteristics are listed in Table 1.1. There are two main research themes in the membrane vesicle field. One establishes the role of membrane vesicles in cellular processes like growth, proliferation, differentiation and immune responses. The other focuses on biomarker discovery and validation of the vesicular secretory pattern. These biomarkers are further used for various disease diagnoses, for monitoring and for therapeutic efficacy, using membrane vesicles as the source for biomarkers.

Acquisition of host major histocompatibility complex (MHC) antigens by donor T-lymphocytes in tissue transplantations was observed thirty years ago (Sharrow, Mathieson & Singer, 1981). Since then many studies have reconfirmed that membrane proteins can be acquired by either cell to cell contact or through secretion of membrane vesicles. Similar membrane-coated vesicles had been seen in cultured malignant cells and red cells while little was understood of the significance of this phenomenon. A popular view was that they are the result of a less important mechanism of membrane shedding or membrane blebbing (Lutz, Liu & Palek, 1977; Dumaswala & Greenwalt, 1984; Taylor *et al.*, 1988). They were even thought to be the result of cell lysis caused by small changes in the osmolarity of the medium.

Shortly thereafter, shedding of a major red cell protein, the transferrin receptor, by maturing reticulocytes was reported simultaneously in two publications (Harding, Heuser & Stahl, 1984; Pan *et al.*, 1985) which established them as being definitely physiological and not an artefact. Hence the name exosome was given to them owing to their obvious exocytic nature or derivation (Johnstone *et al.*, 1987). Since then these membrane-coated vesicles have been known to be secreted by a multitude of cells. The membrane vesicles appear spherical in shape and limited by a lipid bilayer. These vesicles contain soluble contents which are mainly derived from the cytosol of the respective secreting cell. Their membrane is in the same orientation (and not inside out) as that of the secreting cell. These membrane vesicles, depending on their intracellular site of origin may have different biochemical properties and composition. This suggests that they serve different functions in an organism's physiology. The nomenclature of such vesicles is a source of confusion for the scientific community because of the different names used by different groups and also because these vesicles share many common physico-chemical properties like size and density. In addition, they have been called microparticles, vesicles, microvesicles, nano vesicles, membrane particles, exosomes, dexosomes, argosomes and ectosomes among other names (Johnstone, 2006; Thery, Ostrowski & Segura, 2009).

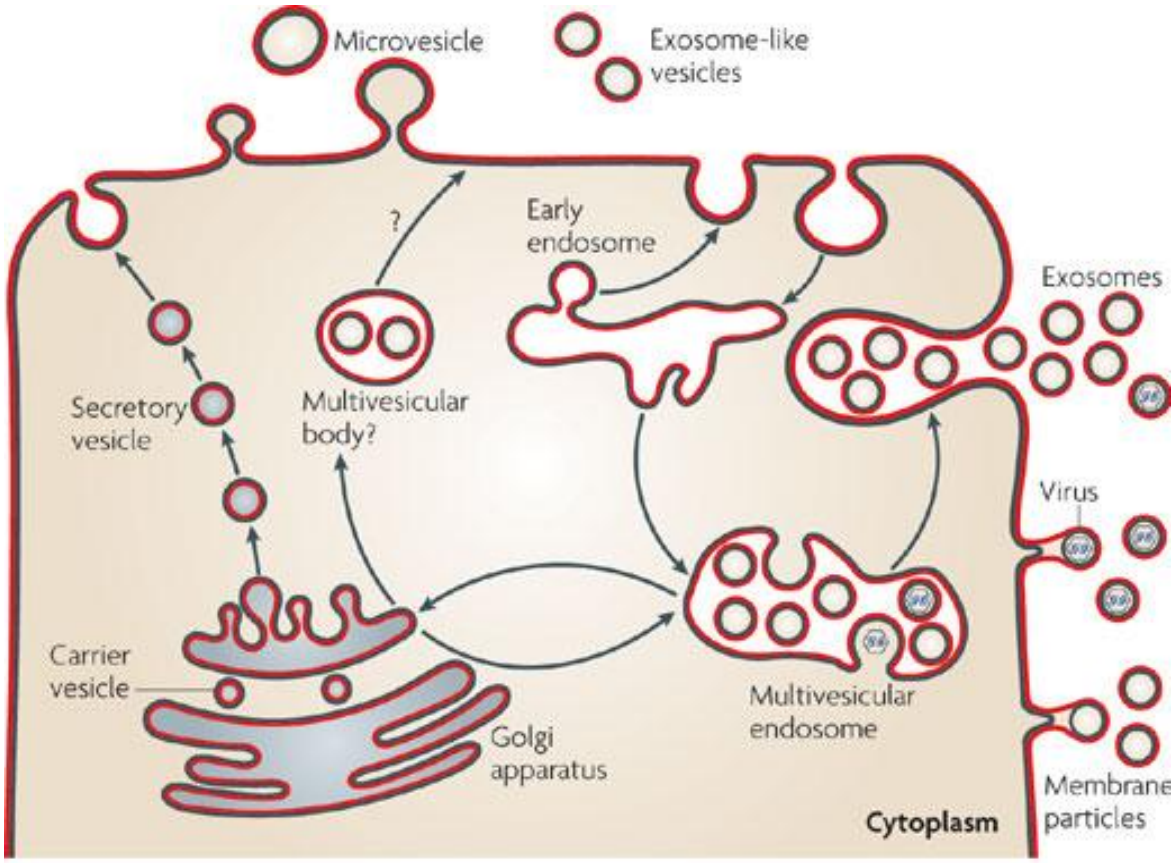

Figure 1.3: A schematic diagram showing the site and the process of secretion of various types of membrane vesicles. The figure was taken from (Thery, Ostrowski & Segura, 2009).

A schematic diagram of secretion of different types of membrane vesicles is shown in Figure 1.3. Major physicochemical characteristics of different types of defined vesicles known are summarised in Table 1.1.

Table 1.1 was taken from (Thery, Ostrowski & Segura, 2009). It summarises different type of membrane vesicles and their biophysical and other characteristics. ND: Not determined.

Feature*	Exosomes	Microvesicles	Ectosomes	Membrane particles	Exosome-like vesicles	Apoptotic vesicles
Size	50–100 nm	100–1,000 nm	50–200 nm	50–80 nm	20–50 nm	50–500 nm
Density in sucrose	1.13–1.19 g/ml	ND	ND	1.04–1.07 g/ml	1.1 g/ml	1.16–1.28 g/ml
Appearance by electron microscopy‡	Cup shape	Irregular shape and electron-dense	Bilamellar round structures	Round	Irregular shape	Heterogeneous
Sedimentation	100,000 g	10,000 g	160,000–200,000 g	100,000–200,000 g	175,000 g	1,200g , 10,000 g or 100,000 g
Lipid composition	Enriched in cholesterol, sphingomyelin and ceramide; contain lipid rafts; expose phosphatidylserine	Expose phosphatidylserine	Enriched in cholesterol and diacylglycerol; expose phosphatidylserine	ND	Do not contain lipid rafts	ND
Main protein markers	Tetraspanins (CD63, CD9), Alix and TSG101	Integrins, selectins and CD40 ligand	CR1 and proteolytic enzymes; no CD63	CD133; no CD63	TNFRI	Histones
Intracellular origin	Internal compartments (endosomes)	Plasma membrane	Plasma membrane	Plasma membrane	Internal compartments?	ND

*The listed features of vesicles secreted by live cells are based on observation of preparations of 100% pure vesicles. However, in practice, all vesicle preparations are heterogeneous, with different protocols allowing the enrichment of one type over another, and they can be classified according to the presence of several (but not necessarily all) of the listed features. ‡Appearance by electron microscopy is only an indication of vesicle type and should not be used to define vesicles, as their microscopic appearance can be influenced by the fixation and phase contrast techniques used. CR1, complement component receptor 1; ND, not determined; TNFRI, tumour necrosis factor receptor I; TSG101, tumour susceptibility gene 101.

This process of secretion of membrane vesicles may be evolutionarily conserved as suggested by release of such vesicles by certain types of bacteria (Mashburn & Whiteley, 2005). Only the eukaryotic vesicles are discussed here.

1.2.1 Ectosomes/Shed vesicles:

Large membrane vesicles or microvesicles (100-1000nm) are secreted by budding or shedding from the plasma membranes of many types of cells (Figure 1.4) including, but not limited to, platelets, neutrophils, dendritic cells and tumour cells (Poutsiaka *et al.*, 1985; Heijnen *et al.*, 1999; Hess *et al.*, 1999; Obregon *et al.*, 2006; Al-Nedawi *et al.*, 2008a).

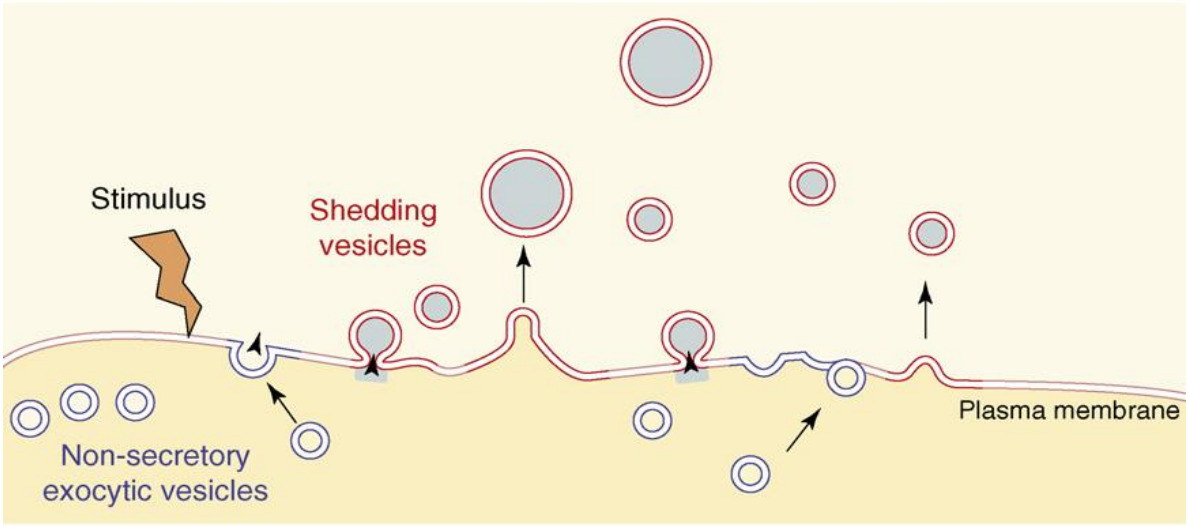

Figure 1.4: Release of shedding vesicles/microvesicles from the cell surface upon exogenous/endogenous stimulus. The Figure was taken from (Cocucci, Racchetti & Meldolesi, 2009).

However, some studies have reported their size is not 100-1000nm as is widely believed. These studies claim, that, if stringent criteria are used to distinguish the interference of cell debris and organelles from disrupted/apoptotic/necrotic cells, then these vesicles are no more than 200 nm in diameter (Scholz *et al.,* 2002; Cocucci *et al.,* 2007; Eken *et al.,* 2008). The term, microvesicle is taken from the study for vesicles proven to be distinct from exosome (Heijnen *et al.,* 1999) and it is not to be confused with context where the term microvesicles was applied to refer to a mixture of heterogenous vesicles including ectosomes and exosomes. These membrane shedded or budded vesicles are also called ectosome and the process is called ectocytosis. We use the term microvesicles for membrane-shed vesicles and it does not include exosomes. Apart from the specific sorting of proteins to these vesicles, specific enrichment of cholesterol and diacylglycerol suggest specific lipid sorting (Stein & Luzio, 1991). Membrane asymmetry at the cell surface upon activation was observed (Frasch *et al.,* 2004) but it has not been conclusively proven that only these sites support ectocytosis and shedding of ectosomes. However, the exposure of phosphatidylserine (PS) on their

surface is a characteristic of microvesicles while one study disputes that PS exposure is always the case with these vesicles (Connor *et al.*, 2010). Another study has characterised platelet-derived microvesicles after separating them into four fractions using chromatography (Dean *et al.*, 2009). These four fractions differ in their average size of vesicles and also have different functions as well as different protein and lipid compositions as judged by different localisation of these proteins among the respective fractions (Figure 1.5).

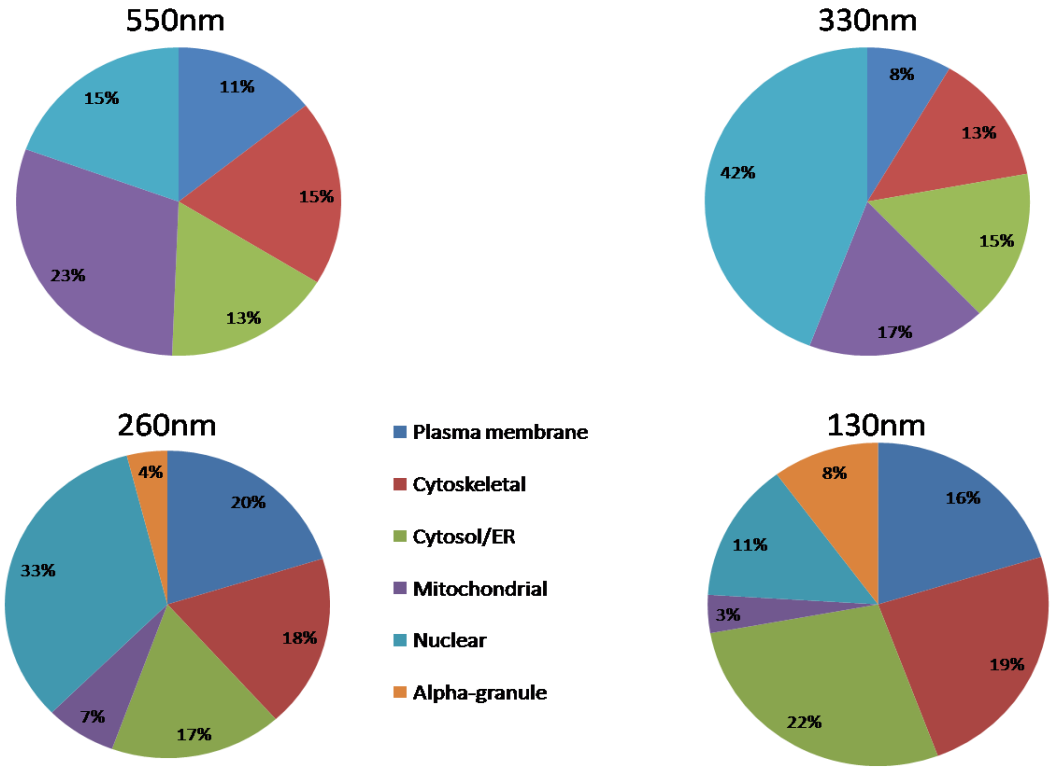

Figure 1.5: The subcellular localization of proteins identified in microvesicles fractionated according to size. The data for generating this figure was taken from (Dean *et al.*, 2009). nm is short form for nanometer and size given is the diameter of vesicles.

Together, these data highlight that there is a great heterogeneity among the vesicles, which may not only be cell-specific, but may also be stimulation and secretion pathway-specific when derived from the same type of cells. Although secretion of such vesicles is stimulus-

regulated, a constitutive level of their secretion is thought to occur and this is supported by the finding of such vesicles in normal plasma (Freyssinet 2003a) and urine (Pascual *et al.,* 1994; Lescuyer *et al.,* 2008).

1.2.1.1 Ectosome biogenesis

Small cytoplasmic protrusions bud out from the cell at specific locations and then detach from the cell by fission of their stalk (Dolo *et al.,* 2000; Cocucci *et al.,* 2007). Complement attack was known to induce secretion of ectosomes by endothelial and circulating blood cells (Pilzer *et al.,* 2005) while bacterial cell wall components such as lipopolysaccharide (LPS) induce their release from monocytes (Satta *et al.,* 1994). Platelets have been shown to release ectosomes upon activation by thrombin (Freyssinet, 2003b) while many cancerous cells have an activated phenotype with high levels of ectocytosis without any stimuli (Dolo *et al.,* 1998; Johnstone, 2006; Al-Nedawi *et al.,* 2008b). Specific protein sorting was observed in the buds with inclusion of some membrane proteins while exclusion of others (Cocucci *et al.,* 2007; Moskovich & Fishelson 2007). The mechanism behind the specific sorting of cargo to these vesicles remains less well defined but many types of stimuli known to induce the secretion were reported. Use of inducers and inhibitors of specific pathways has shed light on the critical factors involved in release of such vesicles. The use of inhibitors of cholesterol synthesis has implicated the involvement of lipid rafts (Del Conde *et al.,* 2005) but these special membrane domains are also implicated in exosome biogenesis (Lakkaraju & Rodriguez-Boulan 2008). Stimulation of cells, like dendritic cells and microglia, by calcium results in induction of ectosome release which can be recorded by confocal time-lapse microscopy (Bianco *et al.,* 2005; Pizzirani *et al.,* 2007). Induction by calcium is cell-type specific as induction of PC-12 cells by Ca^{++}-ionophore fails to release shed vesicles while induction by phorbol ester involving activation of protein kinase C (PKC), is effective in these cells (Cocucci *et al.,* 2007). Increase in intracellular calcium occurs before release of

vesicles upon activation of cell surface receptor or apoptosis (Baroni *et al.,* 2007; Kahner, Dorsam & Kunapuli 2008). Activation of the purinergic receptor-channel P_2X_7 increases the release of shed vesicles in dendritic cells, macrophages and microglia while activation of P_2Y coupled with Gq protein is important in other cells like platelets and PC12 (MacKenzie *et al.,* 2001; Wilson *et al.,* 2004; Bianco *et al.,* 2005; Pizzirani *et al.,* 2007). Regardless of the type of stimulation the generation of shed vesicles is delayed by tens of seconds to 2 minutes depending on stimuli and type of cell (Lee *et al.,* 1993; MacKenzie *et al.,* 2001; Pilzer & Fishelson, 2005; Cocucci *et al.,* 2007; Moskovich & Fishelson 2007; Pilzer *et al.,* 2005; Pizzirani *et al.,* 2007;). This implies that a specific sorting of proteins lipids and metabolites occurs upon stimulation, and consequently, the shed vesicles are released. The mechanism of cargo sorting and biogenesis of these vesicles remains to be understood in detail.

1.2.1.2 Ectosome/Shed vesicle functions

One of the first discovered physiological roles mediated by shed vesicles or microvesicles was that they supply the membrane surface necessary for assembling of procoagulant enzyme complexes (Sims *et al.,* 1988). This is mediated by a tissue factor displayed by these vesicles on their surface (Spek, 2004). These vesicles bind to macrophages, neutrophils and other platelets activating them in the process (Polgar, Matuskova & Wagner, 2005; Pluskota *et al.,* 2008). Vesicles released by neutrophils are typically enriched in the adhesion molecule, Mac-1, which also activate platelets (Andrews & Berndt 2004). Shed vesicles can also regulate the blood flow as shown in a study where, in a vesicular preparation from the T cell line, CEM, mediated the modulation of dilation and relaxation of mouse aortic rings and mesenteric arteries by decreasing the expression of nitric oxide synthase and consequently increasing the expression of caveolin-1 (Martin *et al.,* 2004). Moreover, platelet-derived microvesicles can induce the expression of cyclooxygenase 2 (COX-2) in endothelial cells thereby inducing the vasodilation by prostagandin production (Barry *et al.,* 1997; Barry *et al.,* 1998). Complement

attack can also trigger the release of ectosomes, containing the membrane attack complex, by neutrophils, oligodendrocytes, platelets, glomerular epithelial cells, and the tumor cell lines Ehrlich, U937, and K562 thereby protecting them from cell death (Carney, Hammer & Shin, 1986; Morgan *et al.*, 1986; Morgan, 1989; Scolding *et al.*, 1989). Human monocyte-derived macrophages demonstrate inflammatory responses to Zymosan A (inducer of experimental sterile inflammation) and LPS (inducer of immune response) which can be blocked by ectosomes derived from polymorphonuclear leukocytes (PMN) by inhibiting the release of tumour necrosis factor-α (TNF-α) and reducing the release of interleukin (IL)-8 and IL-10 (Gasser & Schifferli 2004). PMN-derived ectosomes also interfere with maturation of monocyte-derived dendritic cells (Eken *et al.,* 2008). However, in a later stage of inflammation these vesicles can become pro-inflammatory as suggested by the complement component C1q binding to microvesicles released from apoptotic Jurkat cells, and subsequent deposition of complement components C3 and C4 (Nauta *et al.*, 2002). Classical complement pathway activation could thus trigger the proinflammatory effect of complement. Interestingly, platelet-derived microvesicles can deliver arachidonic acid directly to target cells. Arachidonic acid from microvesicles can also increase adhesion of monocytes to endothelial cells (Barry *et al.,* 1998). Furthermore, arachidonic acid derived from microvesicles also increases expression of COX-2 in endothelial and monocyte cells stimulating production of prostaglandins (Barry *et al.*, 1997; Barry *et al.*, 1999), which may be involved in regulation of inflammation. Extracellular ATP-based induction of P_2X_7 receptor in THP-1 monocytic cells stimulates shedding of microvesicles containing IL-1β (MacKenzie *et al.,* 2001) which is a mediator of the inflammatory response and is distinctly involved in cell proliferation, differentiation and apoptosis. These contrasting anti- and pro-inflammatory effects of microvesicles might be cell type and physiological state-dependent and probably the type of stimulation might determine their function post secretion.

1.2.2 Exosomes

Although any membrane vesicles released by cells in extracellular space contain cell type-specific proteins and lipids, distinct features of a sub-population of vesicles define them and they can be referred to as exosomes according to the following characteristics. Exosomes are 40-100nm in diameter, and appear as "cup-shaped" or round in morphology in transmission electron microscopy or cryo-electron microscopy (Conde-Vancells *et al.,* 2008). They typically float on a sucrose density gradient to a density that ranges from 1.13 to 1.19 g/ml (Simons & Raposo, 2009). Elements of cellular compartments, named multivesicular bodies (MVBs), containing intraluminal vesicles, can fuse with plasma membranes to release these vesicles, termed exosomes, to the extracellular space (Lakkaraju & Rodriguez-Boulan, 2008). Nearly all exosomes, regardless of the cell of origin, contain specific proteins because of their endosomal origin. These include proteins involved in membrane fusion and transport (Rab GTPases, Annexins, flotillin), MVB biogenesis (Alix, TSG101), heat shock proteins (hsp70 and 90), integrins and tetraspanins (CD63, CD9, CD81 and CD82). Some of these proteins merely reflect the cellular abundance of these proteins in secreting cells while others can be termed exosomal markers (Alix, TSG101, Flotillin and CD63). Similarly, some lipids are enriched in exosomes such as lipids localised to rafts e.g. cholesterol, sphingomyelin, ceramide and glycerophospholipids with long and saturated fatty-acyl chains (Wubbolts *et al.,* 2003; Subra *et al.,* 2007; Trajkovic *et al.,* 2008). Although some vesicles shed directly from membranes can also be in the exosomal size range (50-100nm) (Booth *et al.,* 2006), endosomal origin and specific enrichment of exosomal marker proteins and lipid as well as morphology of vesicles are used to define exosomes and differentiate them from such vesicles.

1.2.2.1 Exosome biogenesis

Exosomes are intraluminal vesicles (ILVs) of MVBs and their biogenesis is thought to occur in MVB compartments (Figure 1.6).

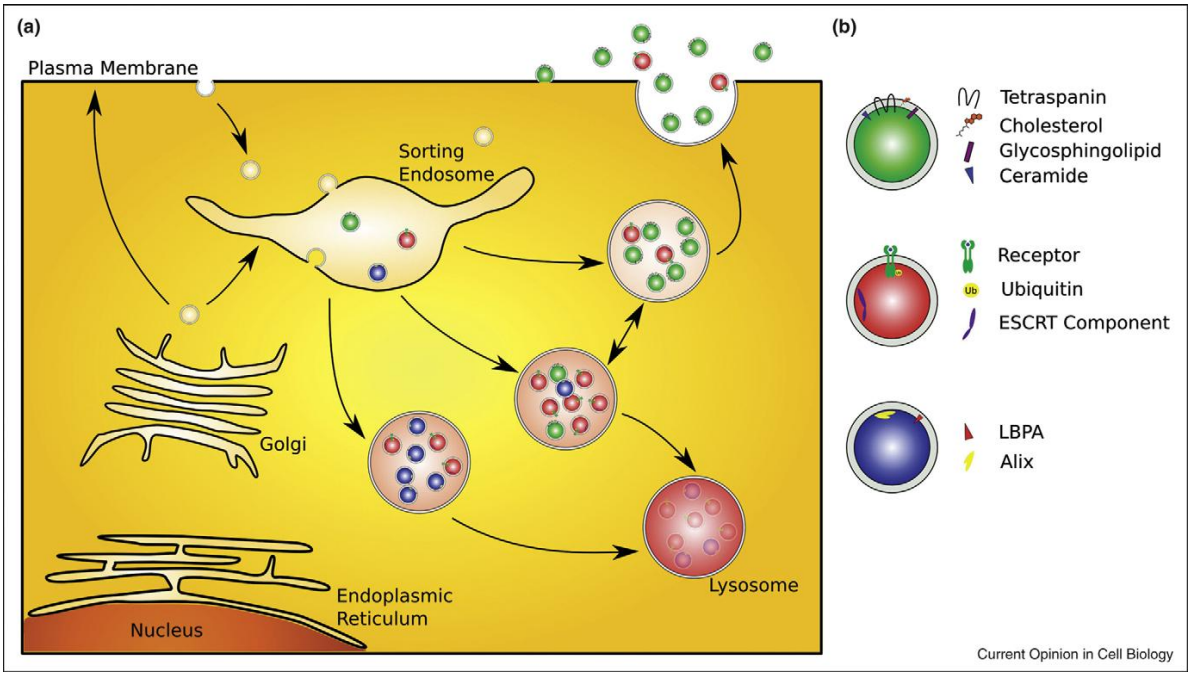

Figure 1.6: Formation of multivesicular body (MVB) and subsequent release of exosomes upon its fusion with plasma membrane. This figure was taken from (Simons & Raposo 2009)

Strong evidence that ILVs are indeed released as exosomes comes from a study in which a specific label was allowed to be internalised by cells and was found to be released with exosomes coming from endosomal-like domains (Raposo *et al.*, 1996). The generation of ILVs involves lateral segregation of cargo within the limiting membrane of endosomes followed by inward budding of ILV and subsequent release of ILV into the lumen of endosome. Although Endosomal Sorting Complex Required for Transport (ESCRT) machinery is thought to function in cargo sorting, its role is not clear. Exosomes derived from various cell types contain enriched ESCRT components and ubiquitinated protein (Buschow *et al.*, 2005; Liu *et al.*, 2009) lending support to this theory. Figure 1.7 shows a schematic diagram of ESCRT components mediating sorting of cargo into multivesicular body lumen.

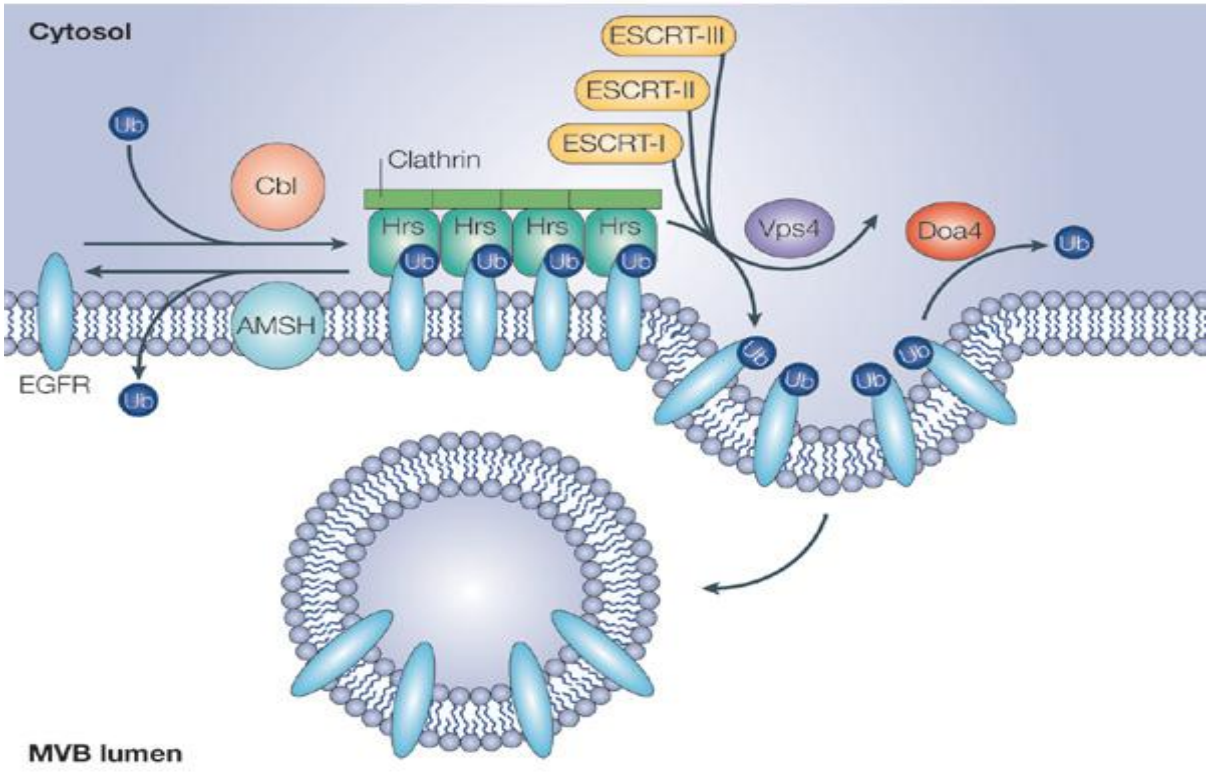

Figure 1.7: Ubiquitination of membrane proteins and their subsequent ESCRT-mediated incorporation into intralumenal vesicles of multivesicular bodies. This figure is taken from (Welchman, Gordon & Mayer, 2005)

More support comes from other observations that Nedd4 family interacting protein-1 expression is associated with increased ubiquitination in exosomes (Putz *et al.*, 2008) and Alix, which associates with ESCRT machinery, is required for sorting of transferrin receptor into exosomes (Geminard *et al.*, 2004). Moreover, treatment of TS/A tumour cell (metastasizing mouse cell line, originated from a mammary adenocarcinoma) with curcumin (an active component of spice turmeric) enhances ubiquitination of exosomal protein and their accumulation in tumour cell-derived exosomes (Zhang *et al.*, 2007). Exosomes contain a number of ubiquitinated proteins (Buschow *et al.*, 2005) leading to the conclusion that ubiquitination of proteins is involved in sorting of the cargo to exosomes and affects exosome functions.

However, the situation is more complex than it seems and some proteins are known to be sorted to exosomes independent of ESCRT involvement (Fang *et al.,* 2007; Trajkovic *et al.,* 2008). Ceramide was shown to be involved in one of these ESCRT-independent pathways (Trajkovic *et al.,* 2008; Bianco *et al.,* 2009; Zhang *et al.,* 2009). This is not surprising as ceramide is known to induce phase-separation and domain formation in model membranes (Goni & Alonso 2006). Protein such as tetraspanins may then partition into these domains. Tetraspanins are being suggested because they form tetraspanin webs by forming oligomers by interacting among themselves and also with other transmembrane and cytosolic proteins (Zoeller, 2009). This is further supported by observations that antibody-induced clustering of proteins, such as transferrin receptor, MHC-II and CD43, increases their secretion with exosomes in diverse cell types (Vidal, Mangeat & Hoekstra, 1997; Fang *et al.,* 2007; Muntasell, Berger & Roche 2007). Addition of multiple homo-oligomerization domains to an acylated reporter protein enhances its exosomal secretion (Fang *et al.,* 2007). Taken together, these data suggest that oligomerization and clustering of exosomal cargo alone, may lead to recruitment of these proteins into, and stabilisation of, exosomal membrane. Exosomal enrichment of lipid-rafts may lead to spontaneous budding of vesicles, driven by tension between liquid-ordered and disordered boundaries. This kind of mechanism would not need involvement of ESCRT machinery. Another example of oligomerization-mediated sorting of proteins into intralumenal vesicles is glycoprotein Pmel17. Pmel17 forms fibers with its luminal domains which induce its sorting into intralumenal vesicles of multivesiclular endosome in an ESCRT-independent manner (Theos *et al.,* 2006). However, exosomes, in addition to proteins and lipids, also contain mRNA and miRNA and only a subset of cells' mRNA and miRNA are found in exosomes suggesting involvement of an active sorting process. No studies have tried to uncover sorting of these RNA species to exosomes. In addition to MVBs, vesicles with exosomal characteristics have also been found to accumulate

into specific domains at plasma membranes and these domains have been termed, 'endosome-like' domains (Booth *et al.,* 2006). These domains are enriched in exosomal and endosomal proteins like CD63 and CD81. Moreover, a lipid probe, N-Rh-PE (1,2-dipalmitoyl-sn-glycero-3-phophoethanolamine-N-[lissamine rhodamine B sulfonyl]) which normally internalises, and releases with exosomes, accumulates in these domains. Electron microscopy detected the budding profile of these domains proving that these domains serve as the sites for exosomal biogenesis. Moreover, addition of acyl chains to a reporter protein targeted it to these 'endosome-like' domains and increased its secretion with exosomes. Higher order oligomerization can target plasma membrane proteins to exosomes (Fang *et al.,* 2007). Various plasma membrane anchors, like the myristoylation tag and PIP2 binding domain, can target oligomeric proteins to these vesicles (Shen *et al.,* 2011). This suggests that oliogmerization and plasma membrane association is one of the several signals which affect protein sorting to exosomes. Another important factor in specifying exosome secretion may be the interaction with a protein with known association to exosomes. Accordingly, CD43 fusion protein is secreted in association with exosomes although to a lesser extent than other plasma membrane anchors (Shen *et al.,* 2011). Tumour suppressor activated pathway 6 (TSAP6), a p53-inducible 5-6 transmembrane protein, is associated with exosomes and it promotes the exosomal secretion of translationally controlled tumour protein (TCTP) (Amzallag *et al.,* 2004). More studies are needed to clarify the role of these different pathways in exosome biogenesis which might be cell type-specific.

1.2.2.2 Exosomes function

One of the earliest functions attributed to exosomes was elimination of proteins from cells needed for their maturation (Harding, Heuser & Stahl 1984; Pan *et al.,* 1985). Transferrin receptor is lost from RBCs in all known species but sheep cells retain glucose transporter while losing nucleoside transporters and the reverse is true of pig cells (Jarvis *et al.,* 1980;

Jarvis & Young 1982; Zeidler & Kim 1982; Johnstone *et al.,* 1987) In addition to transmembrane proteins, GPI-anchored proteins like acetylcholine esterase and prions are secreted out from the cells using exosomes (Johnstone, Bianchini & Teng 1989; Fevrier *et al.,* 2004; Porto-Carreiro *et al.,* 2005). Some of the p53-regulated extracellular proteins are secreted via exosomes and this is accomplished by upregulation in the expression of the protein TSAP6 by p53 (Yu, Harris & Levine, 2006). Some cytokine receptors like Tumour Necrosis Factor (TNF) receptor 1 are released through exosomes (Zhang *et al.,* 2006). These exosomes may compete for ligand binding and the local concentrations of such exosomes may regulate the effect of cytokines on target cells. Also, transfer of active cytokine receptors to target cells may increase the target cell response to such cytokines. In another example of their function, exosomes released by the epididymis, called epididysomes, are known to transfer proteins to sperm cells which are necessary for their maturation and for egg binding (Sullivan *et al.,* 2005). Exosomes also have a well-recognised role in activation of the immune system. Thus, exosomes from mature dendritic cells (DC) are two orders of magnitude more effective in antigen-specific T cell activation compared to those from immature DC (Segura, Amigorena & Thery, 2005). This also suggests that exosome functions can also be regulated by differentiation/maturation status of the cell secreting them. Exosome secreted from tumour-antigen-pulsed DCs induce anti-tumour immunity (Quah & O'Neill, 2005) although this immunity is not specific to tumour type suggesting that mechanisms deeper than only antigen presentation are at work. However, this immunity is T cell dependent so caution has to be exercised while making inferences. Exosomes also function in antigen cross-presentation as they can transfer the antigens from tumour cells to DCs (Wolfers *et al.,* 2001). Apart from antigen presentation, exosomes also function as immunosuppressive agents. Injection of donor-haplotype exosomes from bone marrow DCs before transplantation leads to prolonged heart allograft survival in congenic MHC-

mismatched rats (Peche *et al.,* 2003). The same study also shows significant decrease in CD4+ T cells upon exosome treatment suggesting a role played by them in immuno-tolerance. Exosomes from IL-4 and IL-10-treated immature DCs reduce the severity of collagen-induced arthritis (Kim *et al.,* 2005; Kim *et al.,* 2007). Exosomes from T cells, melanoma cells and ovarian cancer cells are known to carry the Fas ligand (FasL) on their surface which can induce T cell apoptosis (Karlsson *et al.,* 2001; Van Niel *et al.,* 2003; Mallegol *et al.,* 2007). Moreover, exosomes released from DCs, virally transduced to produce FasL, exhibit anti-inflammatory activity (Kim *et al.,* 2006). Based on all the qualities outlined above, exosomes have huge potential clinical applications, particularly in treatment of autoimmune diseases like rheumatoid arthritis and inflammatory diseases. Among their roles in immune functions, it has been discovered that exosomes released from basolateral enterocytes appear to function to induce tolerance for food antigens possibly by transferring antigens from the intestine to local dendritic cells (Lin, Almqvist & Telemo, 2005; Mallegol, van Niel & Heyman, 2005). Therefore, intestinal exosomes may help to maintain tolerogenicity and immuno-suppressive environment of the intestine. Exosome-mediated transfer of mRNA and miRNA has been proposed to confer completely new functions to the target cells (Valadi *et al.,* 2007).

1.2.3 Analysis methods for exosomes and shed vesicles and associated challenges

As the clinical interest in exosomes and shed vesicles increases, new detection and analysis methods are urgently needed. All methods have their own advantages and drawbacks which are discussed below. The following are the commonly used and novel methods used for this purpose.

1.2.3.1 Electron microscopy

Transmission electron microscopy (TEM) uses electrons to generate an image and the wavelength of electrons is three orders of magnitude shorter than visible light. Therefore

resolution down to 1nm can be achieved. TEM is currently among the best method to determine the morphology and size of both exosomes and shed vesicles (Pisitkun, Shen & Knepper, 2004). A known source of artefacts is the fixation and dehydration steps which may affect size and morphology of vesicles. Purified vesicles are used for TEM analysis. Therefore, pre-analytical factors, such as enrichment and aggregation of vesicles, can bias the counting making the determination of concentration inaccurate. The measurement time required for TEM is several hours and, additionally, fixation and negative staining have to be performed, further increasing the time required. Use of immunogold TEM can also provide information on biochemical composition (Pisitkun, Shen & Knepper, 2004) of the vesicles by detecting specific proteins although quantitation is not very accurate (Griffiths & Hoppeler, 1986) and sensitivity depends on retention of antigen conformation and antibody binding during fixation and staining. Field emission scanning electron microscopy (FESEM) has also been applied to image exosomes (Sharma *et al.,* 2010) and the resulting images are three-dimensional which seem better than TEM in determining morphology and size (Figure 1.8).

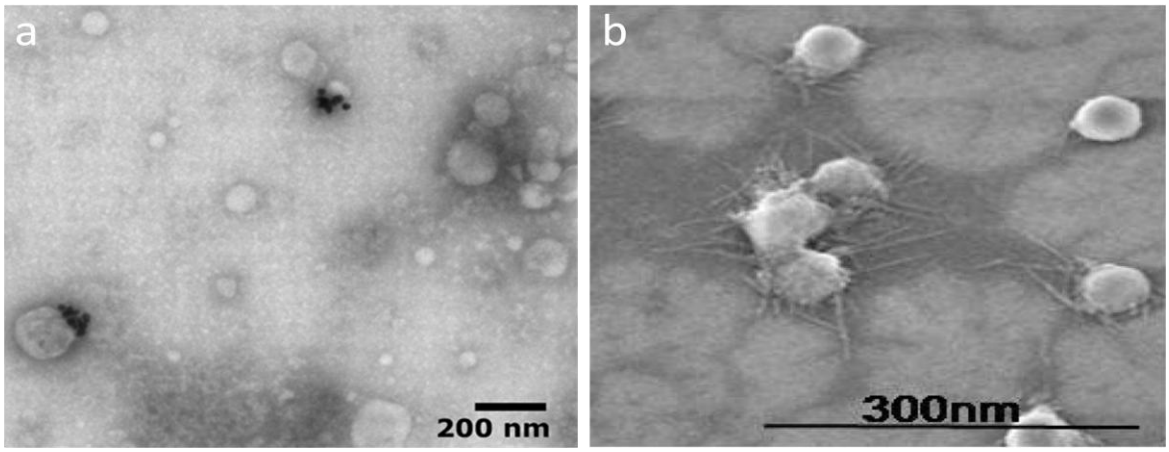

Figure 1.8: TEM (a) is compared with FESEM (b) for determination of morphology, shape and size of exosomes. Both parts of image are for exosomes derived from saliva. The TEM image in part a) is taken from (Berckmans *et al.,* 2011) and the FESEM image in part b) is taken from (Sharma *et al.,* 2010)

A distinct advantage of FESEM is minimal sample preparation and elimination of the need for negative counterstaining of the exosome vesicles. Moreover, FESEM images are obtained at much lower beam energy compared to TEM which is expected to reduce the direct damage to vesicles during multiple imaging.

1.2.3.2 Atomic force microscopy

An atomic force microscope (AFM) consists of a cantilever with a tip at the end which scans the surface (Mica) without physical contact (tapping mode) and through recording the accurate movements of the tip, a three-dimensional image is created. AFM can be used to obtain information about size, morphology, shape and density of a vesicle population (Palanisamy *et al.,* 2010). Figure 9 compares the images of exosomes obtained using tapping mode and amplitude modulation-AFM (AM-AFM) (Sharma *et al.*, 2010).

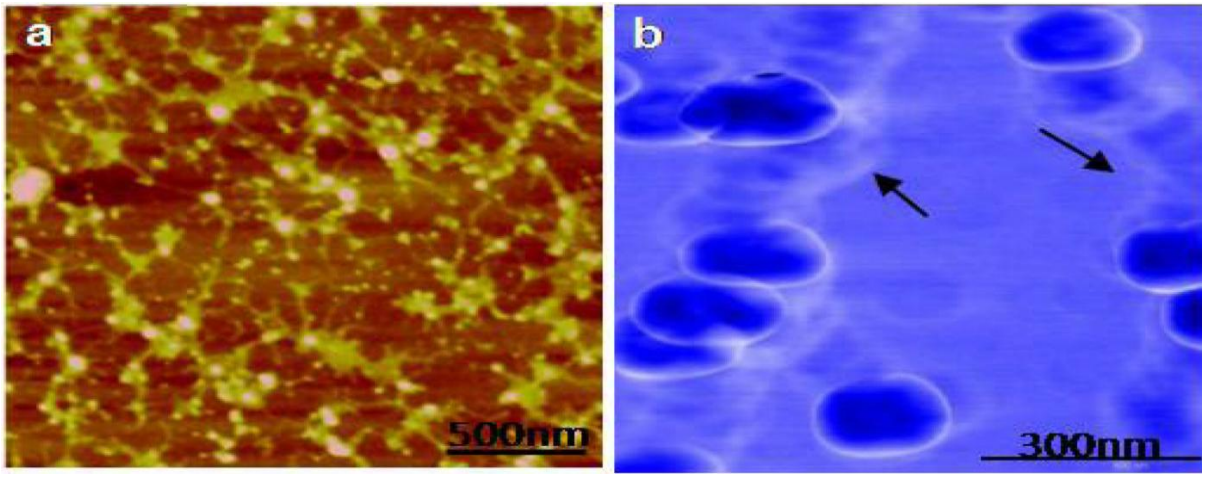

Figure 1.9: Salivary exosomal image obtained using the tapping mode (a) and the image obtained using amplitude modulation mode (b). In part (b), aggregation of vesicles can be more easily observed and tube like connections between individual vesicles are marked by arrows. This figure is taken from (Sharma *et al.*, 2010)

However, AFM analysis entails immobilization of vesicles on a mica surface which can be physical adsorption or antibody-mediated immobilisation. Due to the variable efficiency of

these processes, the concentration of vesicles cannot be determined. Moreover, shape can be altered upon physical immobilization and the force applied by the AFM tip (Sharma *et al.,* 2010) may lead to generation of artefacts. A rise in concentration of the vesicles during isolation and purification can also produce artefacts like aggregation of exosomes. Density of protein receptors on the surface of individual vesicles can be obtained using antibody-functionalised AFM tips or applying gold-conjugated antibody before AFM imaging (Sharma *et al.,* 2010). The measurement time involved is in the order of many hours.

1.2.3.3 Nanoparticle tracking analysis

Nanoparticle tracking analysis (NTA) measures the absolute size distribution of particle in a fluid ranging from 50 nm to 1 μm. Particles in a fluid are illuminated by a laser light and scattered light is collected by an optical microscope. The movement of individual particles is followed by a video sequence and velocity is calculated which gives information about absolute size distribution of the particles after calibrating with beads of known size and concentration (Filipe, Hawe & Jiskoot, 2010). However, owing to the detection limit of the microscope, particles smaller than 50 nm cannot be detected. Fluorescence-based NTA methods are developing at a fast pace and they are ideally suited for exosomes (Sokolova *et al.,* 2011) as the limit of detection is in range of the size of exosomes. Bigger vesicles (>50nm) can be readily analysed using NTA although the only information obtained is size distribution and biochemical composition and cellular origin remain undetermined.

1.2.3.4 Flow cytometry

In a flow cytometer (FC) particles are guided in a hydro-dynamically focused fluid stream using a laser beam. There are two detectors one measuring the forward light scattering (FSC) which is in line with laser and the other being perpendicular to the beam measuring side scattering (SSC). The lower detection limit of commercial flow cytometers for polystyrene beads is 300-500nm (Steen, 2004; Perez-Pujol, Marker & Key, 2007; Robert *et al.,* 2009).

Only the particles differing in size by at least 280nm can be resolved by FC (Perez-Pujol, Marker & Key, 2007; Robert *et al.,* 2009). Both these limitations mean that only certain sub-populations can be detected by scattering FC and exosomes are not detected at all. Quantitative information about size is obtained by comparing the light scattering intensity of vesicles with beads of known size. However, scattering intensity is dependent not only on size but also on shape, refractive index and light absorption making inferences about size difficult. The concentration can be determined if the flow rate is known, provided that all the vesicles are detected and intensity of scattered light is above the detection limit. No information on morphology of vesicles from scattering FC is obtained. Information about rough biochemical composition and cellular origin can be obtained by correlating FSC with SSC. For example, FSC correlates with the volume of the particle and SSC correlates with inner structures such as shape of the nucleus. However, light scattered from surface of large vesicles masks the light scattered from structures inside the vesicles making it difficult to distinguish the vesicles with different cellular origins having different inner structures. However this analysis can be improved by analysing the polarization of light scattered sideward (Degrooth *et al.,* 1987). Fluorescence based FC (FFC) is more sensitive than scattering-based FC as the intensity of the fluorescence is higher than scattered light. Modern FFC can detect a single fluorophore by minimizing interfering background fluorescence. However, enumeration of the vesicles depends on signal-to-background noise ratio which sometimes can give false results. Also the efficient labelling of all the vesicles and low background depends on efficiency of antibodies used and the density of the respective antigens on individual vesicle's surface. Lipid-based probes have been proposed but when working on a whole fluid or non-purified samples then staining of cells and particularly, interference by membrane debris may pose problems. Calcein-AM has also been tested and brings advantages as only the intact vesicles are stained and not the debris (Kendall &

Macdonald, 1982). FC remains an active area of research for developing size evaluation and counting methods for exosomes and other defined shed vesicles. However, exosomes pose significant challenge owing to their small size.

1.2.4 Urinary Exosomes and microvesicles

Urine contains exosomes secreted by epithelial cells along all nephron segments (Pisitkun, Shen & Knepper, 2004) as well as other type of vesicles like podocyte microvilli-derived vesicles (Hara *et al.,* 2010). Exosomes are released into urine by fusion of MVB membrane with the apical plasma membrane while podocyte-derived vesicles are generated by a process called tip vesiculation (Hara *et al.,* 2010). Other type of vesicles, like ectosomes have not been studied from the urine although there is no reason to believe they are not present. Vesicles have been harvested from urine using the same methodology in multiple studies. However, different authors have given them different names leading to a high degree of confusion. For example Pistikun *et al.,* describe these vesicles as exosomes (Pisitkun, Shen & Knepper, 2004) while Smalley *et al.,* call them microparticles (Slrnalley *et al.,* 2008). Exosomes from the urine contain the typical exosomal markers including CD63 and Tsg101 while podocyte-derived vesicles contain podocalyxin and complement receptor 1 (CR1 or CD35) although there is an overlap of some of the marker proteins. Biogenesis of exosomes and the pathways and signals involved in protein sorting to exoxomes have already been discussed in previous sections (section 1.2.2.1). Urinary exosomes typically contain cytoplasmic protein entrapped at the time of inward budding into endosomal membrane. It also contains a snapshot of endocytotic proteome of the apical plasma membrane. Moreover, exosomes isolated from urine are in fact a mixture of exosomes secreted by all types of renal epithelial cells. Therefore, they contain proteins from glomerular podocytes, epithelial cells of proximal tubules, thick and thin ascending limbs of Henle, the distal convoluted tubule, the collecting duct, the renal papilla, the renal pelvis, the ureter as well as the transitional

epithelia of urinary bladder and urethra (see Figure 1.2). These characteristics make urinary exosomes an ideal source and starting material for biomarker discovery for diagnosis and monitoring of diseases affecting different parts of the nephron. Also, the isolation of urinary exosomes enriches for low abundant proteins by removing the higher abundant proteins normally present in urine and reduces the complexity of the urinary proteome. Together these facts make exosomes ideally suited for biomarker discovery. In a large-scale urinary exosome proteomics study, 177 proteins, that are associated with distinct diseases, were identified and 34 of them were involved in specific kidney pathologies (Gonzales *et al.*, 2009). Some of the proteins found in exosomes and those associated with specific renal pathologies or blood pressure regulation are shown in Table 1.2.

Table 1.2: Selected **p**roteins identified in urinary vesicles that are associated with kidney diseases or hypertension. The table has been taken from (Pisitkun, Shen & Knepper, 2004)

Kidney diseases or hypertension	**Identified proteins**
Autosomal dominant and autosomal recessive nephrogenic diabetes Insipidus	Aquaporin-2
Antenatal Bartter syndrome type 1	Sodium potassium chloride cotransporter-2
Gitelman's syndrome	Thiazide-sensitive Na-Cl cotransporter
Autosomal recessive pseudohypoaldosteronism type 1	Epithelial sodium channel α, β, γ
Liddle syndrome	Epithelial sodium channel β, γ
Familial renal hypomagnesemia	FXYD domain-containing ion transport regulator-2

Autosomal recessive syndrome of osteopetrosis with renal tubular acidosis	Carbonic anhydrase II
Proximal renal tubular acidosis	Carbonic anhydrase IV
Autosomal dominant polycystic kidney disease type 1	Polycystin-1
Medullary cystic kidney disease 2 and familial juvenile hyperuricemicnephropathy	Uromodulin
Autosomal recessive steroid–resistant nephrotic syndrome	Podocin
Fechtner syndrome and Epstein syndrome	Nonmuscle myosin heavy chain IIA
2,8-Dihydroxyadenine urolithiasis	Adenine phosphoribosyltransferase
Hypertension	Angiotensin I converting enzyme isoform-1 Aminopeptidase A Aminopeptidase N Aminopeptidase P Neprilysin Hydroxyprostaglandin dehydrogenase 15- (NAD) Dimethylarginine dimethylaminohydrolase-1

1.2.4.1 Exosomes as biomarkers

An early example of clinical utility of urinary exosomes was demonstrated by immunoblotting analysis of urinary exosomes from two patients with Bartter syndrome type I. The patient with this syndrome present with mutations in SLC12A1 gene which codes for sodium-potassium-chloride co-transporter protein (NKCC2) and accordingly NKCC2 bands were absent in the urine of patients compared with appropriate controls (Gonzales *et al.,* 2009). Aquaporin-2, an important channel protein in water homeostasis, is delivered into urine via exosome secretion and is a well established biomarker of several water-balance disorders, such as diabetes insipidus (Kanno *et al.,* 1995; Elliot *et al.,* 1996; Valenti *et al.,* 2000; Ishikawa & Schrier, 2003). The Na+/H+ exchanger protein isoform 3 is a candidate biomarker of kidney tubular damage and could be useful in differential diagnosis between

acute tubular necrosis and other causes of acute kidney injury (AKI) (Du Cheyron *et al.,* 2003). Another study has reported urinary exosomal fetuin-A to be a potential biomarker of AKI in a rat model of the disease further verified in immunoblot analysis of human patient urine samples (Zhou *et al.,* 2006a). A number of transcription factors were identified in urinary exosomes which could reflect a new class of biomarkers of AKI and chronic kidney diseases (Hogan *et al.,* 2009) more related to activation of specific genetic pathways. Moreover, analysis of urinary exosomes in patients with malignancies of urinary drainage tract would be very helpful in diagnosis of the disease and monitoring of therapy. Accordingly the proteomic study of urinary microvesicles from individuals with bladder cancer revealed eight over-expressed proteins and one downregulated protein compared to healthy individuals (Slrnalley *et al.,* 2008). Urinary exosomes also contain mRNA (Valadi *et al.,* 2007) and identification of genetic mutations in these would be a lucrative non-invasive diagnostic test for genetic diseases.

1.2.4.2 Isolation methods for urinary exosomes

The traditional method for urinary exosome isolation has been differential centrifugation where a low speed centrifugation (17,000g) is adopted to remove whole cells, tubular casts and membrane fragments. This step is followed by ultracentrifugation (200,000g) to pellet down urinary exosomes and other similar sized vesicles (Pisitkun, Shen & Knepper, 2004). However, it has to be appreciated that direct 17,000g centrifugation pellets down not only cells but also larger groups of vesicles like shed vesicles. Furthermore, the 200,000g pellet often consists of vesicles other than exosomes such as similar size membrane particles and ectosomes. The stepwise centrifugation method yields a heterogenous mixture of vesicles and cannot be considered to contain pure exosomes. One additional problem associated with this method is contamination of 200,000g pellet with soluble or aggregated high-abundant proteins like Tamm-Horsfall glycoprotein (THP) and albumin. Thus, this pellet at best can be

considered a crude pellet of exosomes. Furthermore, it was found that entrapped in THP polymers, exosomes were also pelleting down at low speed (P18,000g) (Fernandez-Llama *et al.*, 2010). To remove contamination of aggregated proteins one solution has been proposed which includes treatment of low speed and high speed pellets with dithiothreitol (DTT) followed by a second centrifugation at low and high speed.

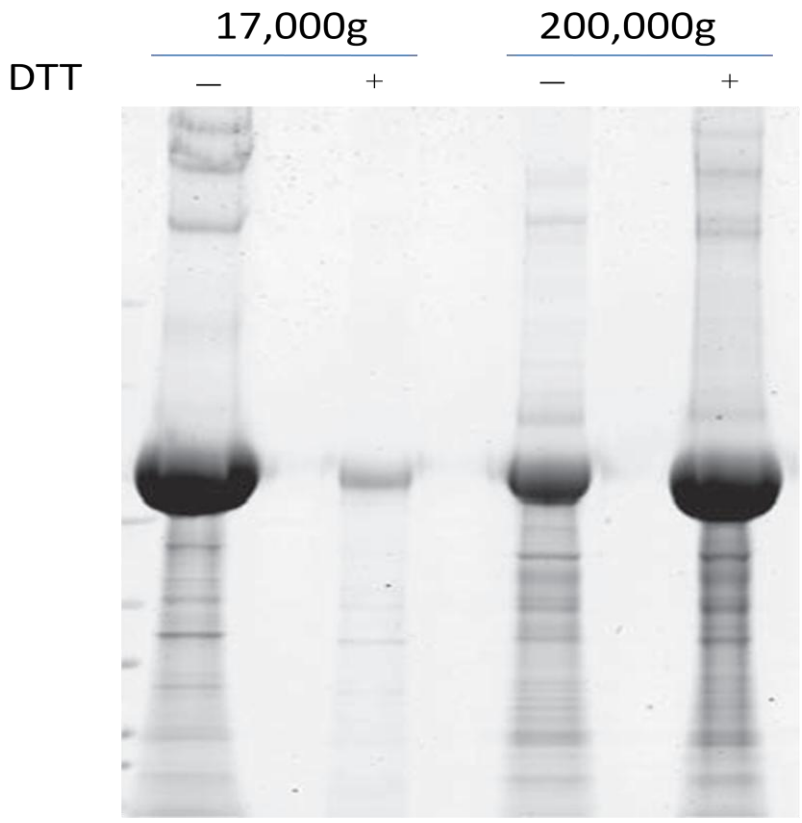

Figure 1.10: Effect of DTT treatment on THP behaviour of low (17,000g) and high speed (200,000g) pellets of urine. The figure has been taken from (Fernandez-Llama *et al.*, 2010)

DTT treatment largely removes the THP and exosomal markers from low speed pellet which now precipitate with the high speed pellet. This procedure increases the exosomal yield without any doubt but it creates a new problem as THP is now present in very high amounts in the high speed pellet (Figure 1.10). This fact often leads to interference in further analysis of the exosomal pellet by masking low abundance proteins. Another method which has been

employed is sucrose density gradient centrifugation. This was used to isolate exosomes from immune cells and cell culture and also on urinary exosome-like vesicles (Hogan *et al.*, 2009). This method is presently the most sensitive one available and yields the purest forms of exosomes with minimal contamination with high abundant proteins. It has to be empahsized that both these ultracentrifugation methods are very labour intensive and require extensive instrumentation and long processing times making them unsuitable for clinical diagnostics. Moreover, this would make them impractical where a large number of samples have to be handled. Alternatives have been proposed like nanomembrane concentration (pore size 13nm and Molecular Weight 'cut-off' of 100kDa) to enrich the exosomal fraction from urine (Cheruvanky *et al.*, 2007). It was possible in this study to recover membrane vesicles with exosomal markers as tested with western blotting (WB) and the yield was found to be similar to ultracentrifugation. There was minor loss of vesicles on the filter membrane surface which were recovered by 2X pre-heated Laemmli buffer which solubilised those vesicles. Advantages of the method include the short 'processing time' and the use of inexpensive instrumentation, e.g. the table top centrifuge. However, recovery of vesicles from filter surface with Laemmli buffer would yield solubilised exosomal proteins which can only be analysed by SDS-PAGE or WB and FC or TEM analysis would be rendered impossible. Other types of analysis like MS analysis would require additional purification steps. The 'cut-off' at 100 kDa means that high molecular weight proteins and protein complexes not part of exosomes will be enriched as well and will interfere with further analysis by masking low abundance proteins in exosomes. As many proteins, like THP and albumin, remain in polymeric forms they will be enriched with exosomes as well. It has already been shown that nanomembrane ultrafiltration enriches soluble proteins like albumin and α-1-antitrypsin along with exosomes when applied to nephrotic urine (Rood *et al.*, 2010) and limits the detection of microvesicular proteins. Another alternative method showed that loading the crude

ultracentrifugation pellet onto a size exclusion chromatography column yielded three fractions (HMW, LMW and >10kDa). The high molecular weight fraction (HMW) showed the presence of exosomal markers while the low molecular weight (LMW) fraction had only high abundant proteins. The LMW fraction was present only in nephritic urine and absent in normal urine suggesting that large amounts of high abundance proteins interfere with the isolation and subsequent analysis of exosomes in nephrotic urine. The impact of storage conditions and preservation was assessed on yield and stability of the exosomal fraction and associated proteins (Zhou *et al.*, 2006b). The study showed that some of the exosomal proteins can be susceptible to degradation if protease inhibitors are not added at the time of sample collection. However, this was done on two different samples and not on the same sample which had been divided into two parts thus it may be subject to sample-to-sample variations. Freezing urine samples at -80°C was better for recovery of exosomes by ultracentrifugation compared to freezing at –20°C and extensive vortexing improves the recovery after freezing. Finally, first and second morning urine is comparable in terms of content and recovery of exosomes and associated proteins. Adjustment of pH before freezing the samples did not affect protein stability (Miki & Sudo, 1998).

1.2.4.3 Proteomic analysis of urinary exosomes and other vesicles

Proteomic analysis is an important tool which helps to annotate and understand the function of exosomes as well as presenting avenues for biomarker discovery by enabling differential protein analysis between normal and pathologic patient samples. Exosomal composition varies according to the cell type of origin but some components appear to be present in all exosomes which reflects their secretion pathway and common origin. The presence of tetraspanins (CD63, CD9, CD81 and CD82) in exosomes, regardless of their cell of origin, is such a typical characteristic. Exosomes also contain heat shock proteins Hsp70 and Hsp90 which assist loading of peptide onto the major histocompatibility complexes MHCI and

MHCII, among other functions. Exosomes are also enriched in proteins involved in vesicle formation and trafficking, like Alix and Tsg101 among others. Proteins involved in membrane fusion and exosome docking, like Rabs and annexins including I, II, V, VI are also present (Mears *et al.*, 2004; Futter & White, 2007) common to most exosomes from various cells. A variety of adhesion molecules, like CD146, CD9, CD18, CD11a, b and c, milk-fat-globule EGF factor-VIII (MFG-E8), intercellular adhesion molecule-1 and CD58, have been found in exosomal fractions (Thery *et al.,* 2001; Mears *et al.,* 2004). Proteins involved in apoptosis such as 14-3-3, galectin-3 and thioredoxin peroxidise have also been found. Metabolic enzymes including enolase-1, pyruvate and lipid kinases and peroxidases are also commonly found (Hegmans *et al.,* 2004). Figure 1.11 is a schematic showing protein composition found in exosomes from various sources along with their function for some of them.

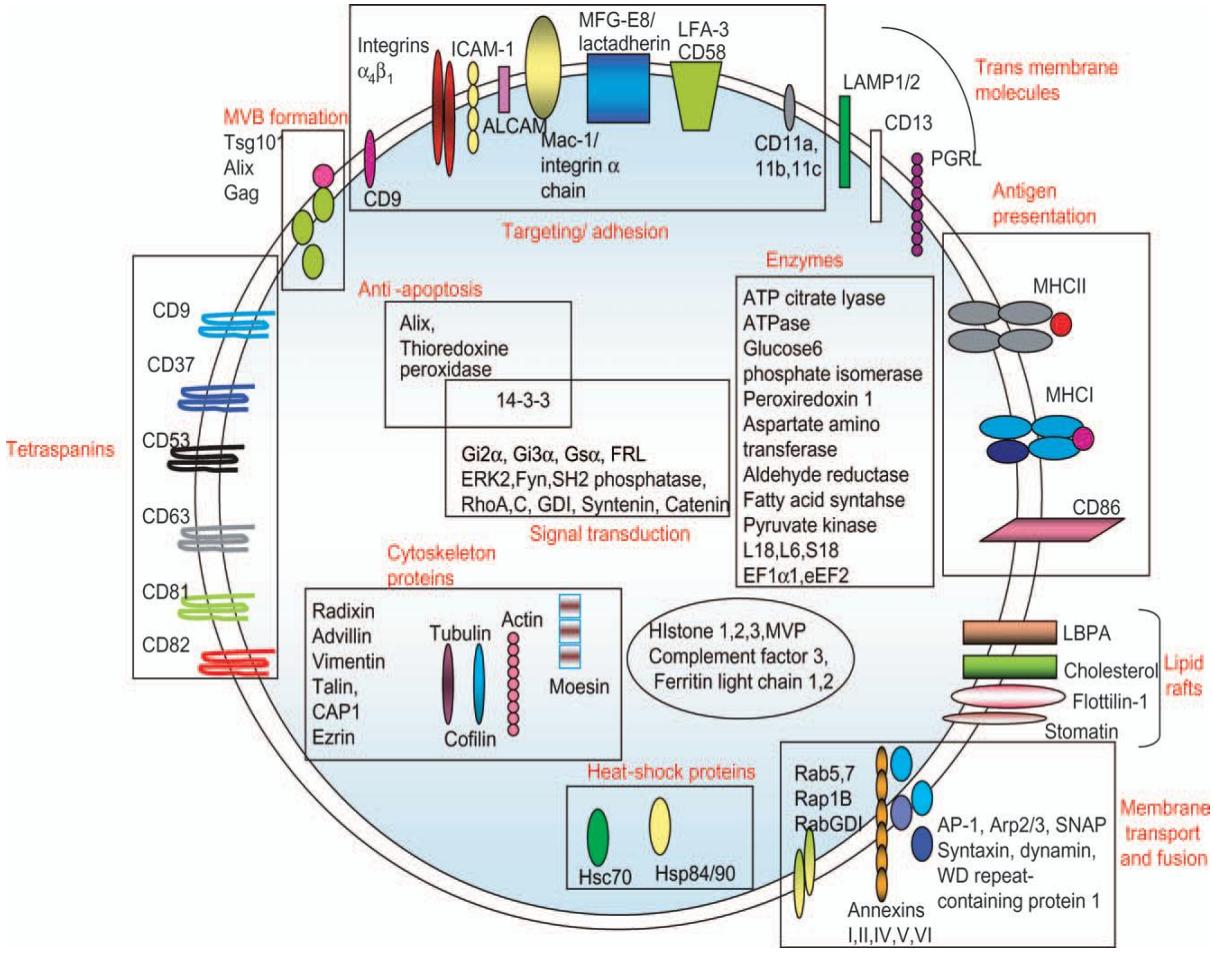

Figure 1.11: Schematic representation of typical proteome content of exosomes derived from various types of cells. This figure has been taken from (Schorey & Bhatnagar, 2008)

There have been a number of proteomic studies of different types of urine vesicles such as membrane particles and urinary exosomes. One of the earliest was by Pistikun *et al.*, (Pisitkun, Shen & Knepper, 2004) who isolated low density membrane vesicles by differential centrifugation followed by DTT treatment to remove interference of THP and they termed these vesicles "exosomes". They identified 295 unique proteins, from which 73 are known to be involved in membrane trafficking including endosomal traffic. Forty-eight proteins were integral membrane proteins, eight were GPI-anchored and twenty three were peripheral membrane proteins. All the integral proteins identified were from the apical side of plasma membrane and none were from the baso-lateral side suggesting that exocytosis of

vesicles found in urine occurs at the apical side of cells. Additionally a number of high-molecular weight ubiquitin bands were observed in the immunoblot (Figure 1.12) suggesting these membranes consist of ubiquitinated proteins.

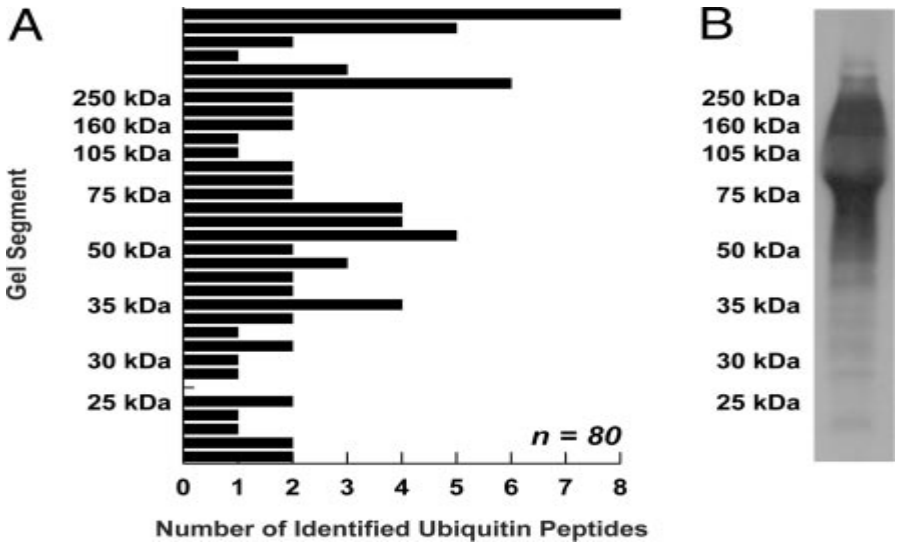

Figure 1.12: Instances of identification of ubiquitin and immunoblot using anti-ubiquitin antibodies. Part A shows the numbers of different ubiquitin peptides were identified at a given molecular weight; x-axis is the number of ubiquitin peptides identified and the y-axis is the molecular weight corresponding to the gel segment from where the band was cut. Part B is the western blot of high-speed pellet (P200,000g) using anti-ubiquitin antibodies. The figure is taken from (Pisitkun, Shen & Knepper, 2004)

A number of solute and water transporters were also identified. Readers are directed to the original study for complete list of the proteins. Polycystin-1, which codes for protein product of the gene responsible for autosomal dominant polycystic kidney disease, was also identified. Shortly thereafter the same group published a large-scale proteomics and phosphoproteomics study on urinary exosomes isolated by the same method as their previous

study (Gonzales *et al.*, 2009). The list of phosphoproteins identified and the phosphorylation site is provided in the following section about post-translational modifications. They identified 1132 proteins unambiguously including 205 proteins from the previous study and 927 new proteins. A large number of integral membrane proteins, mainly apical solute and water transporters were identified from every kidney tubule segment including the proximal tubule (sodium-hydrogen exchanger 3, sodium-glucose co-transporter 1 and 2, and aquaporin-1 [AQP1]), the thick ascending limb (sodium-potassium-chloride co-transporter 2 [NKCC2]), the distal convoluted tubule (thiazide-sensitive Na-Cl co-transporter [NCC]), and connecting tubule/collecting duct (AQP2, rhesus blood group C glycoprotein [RhCG, an ammonia channel], B1 subunit of vacuolar H^+-ATPase, and pendrin). Twenty two proteins, part of the ESCRT-0, ESCRT-I, ESCRT-II, and ESCRT-III complexes involved in multivesicular body formation, were also identified. These proteins cover 75% of total ESCRT machinery. This original study was followed by another study published by Hogan *et al.*, (Hogan *et al.*, 2009), who isolated a urinary exosomal fraction by a traditional centrifugation method and this fraction was subfractionated by density gradient ultracentrifugation (5-30% sucrose gradient). This resulted in three fractions representing markers aquaporin-2, polycystin-1 (PC-1) and podocin. These PKD vesicles were analysed by proteomic methodology and 552 proteins were identified including 232 proteins not reported in urinary proteomic databases. They also identified polycystin-2 (PC-2) and fibrocystin (FCP) in their PKD vesicle preparations. They identified the *in vivo* cleavage products of PC-1 and FCP which confirmed the previous *in vitro* studies (Ponting, Hofmann & Bork 1999; Qian *et al.*, 2002; Kaimori *et al.*, 2007). They also demonstrated the interaction of PKD vesicles with primary cilia of kidney and biliary epithelial cells. In addition to these studies, another study has shown isolation of podocyte membrane vesicles from normal and pathological (several types of nephropathies) urine by immunoadsorption onto anti-

complement receptor-1 (CR-1) antibody-bound magnetic beads (Lescuyer *et al.,* 2008). They identified 76 proteins out of which 37 were identified only in pathological urine. Interestingly, 55 proteins of the 76 were plasma proteins and one could speculate about proteinuric states in these nephropathic samples. Paraoxanase-1 (PON-1) was identified in human urine for the first time from a patient with severe lupus nephritis and renal insufficiency.

A comparison of proteins identified in all these proteomic studies of urinary vesicles with the most comprehensive urine protoemics study to date (Marimuthu *et al.,* 2011) and with each other is presented in Figure 1.13. CR-1 immunopurified vesicles proteomics is marked as ectosomes and all other studies are marked with the first author of the studies.

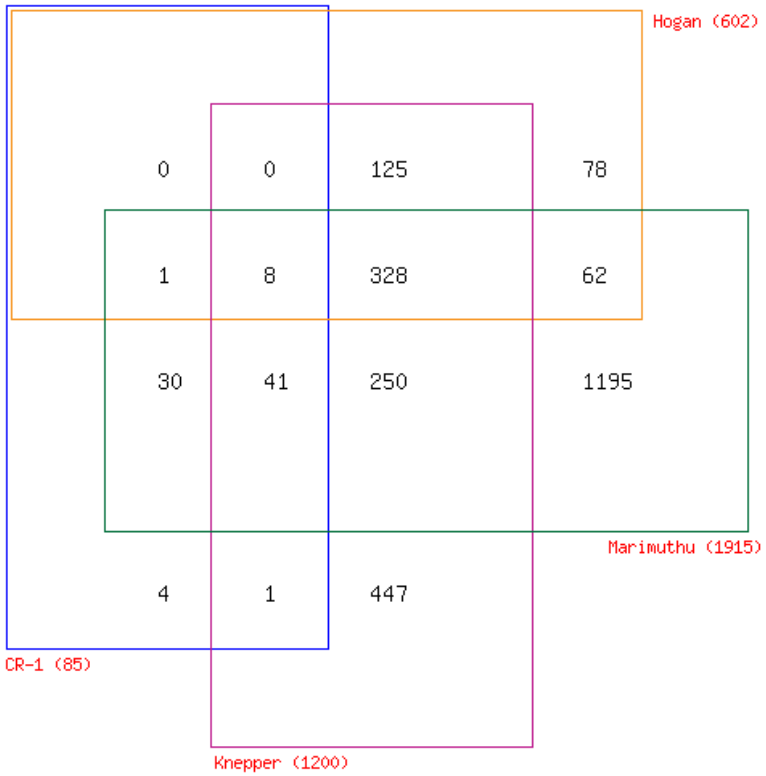

Figure 1.13: exosomal proteomics studies (Combined (Pisitkun, Shen & Knepper, 2004) and (Gonzales *et al.,* 2009)), exosome-like vesicles (Hogan *et al.,* 2009), CR-1 immuno-isolated vesicles (ectosomes (Lescuyer *et al.,* 2008)) and whole urine proteomic study (Marimuthu *et al.,* 2011) are compared to each other.

Pistikun and Gonzalez *et al.,* have worked with crude vesicle preparations. Therefore, their exosomal fraction is probably a mixture of many different types of vesicles like exosomes, ectosome/shed vesicles and membrane particles. Hence, they are likely to have overlaps with all other studies which are illustrated in figure 1.13. On the other hand Hogan *et al.,* purified vesicles which have a significant fraction of proteins not covered by any of the earlier studies. Apart from the highly dynamic nature of exosomes, this suggests that theirs is a novel protein set with possible different functions than other exosomal and ectosomal vesicles.

1.3 Post-translational modification and proteomic analysis

The function of a cell is dependent on a multitude of biochemical reaction cycles, going on simultaneously in the cell, at various locations. A very high degree of diversity in proteins involved in these reactions is needed to temporally and spatially regulate the respective activities at different sub-cellular microdomains. Altered and sometimes many different localisation of a same protein may be needed for different pathways. The same protein might be required, for example, in cytoplasm or plasma membrane to execute the signals propagated by different stimuli. This extremely high level of diversity is not possible with only approximately 25,000 genes coding for proteins in humans (U.S. Human Genome Project, http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml). One of the major ways to generate such diversity is through covalent post-translational modifications (PTMs) of the proteins at one or more sites (Walsh, Garneau-Tsodikova & Gatto, 2005). Because α-carboxy and amino group of an amino acid are involved in bonding to make peptide bond in protein backbone, these modifications mainly occur at amino acid side chains and/or at the amino- or carboxy-terminus of the proteins. Table 1.3 lists some of the major PTMs of the different amino acid side chains.

Table 1.3: Common post-translational modifications of amino acid side chains. This table has been adapted from (Walsh, Garneau-Tsodikova & Gatto, 2005)

Residues	Reaction	Example
Multiple residues	Proteolysis	Signal peptide cleavage, activation of proteases
Asp	Phosphorylation	Protein tyrosine phosphotases, response regulators in two-component systems
	Isomerisation to isoAsp	
Glu	Methylation	Chemotaxis receptor proteins
	Carboxylation	Gla residues in blood coagulation
	Polyglycination	Tubulin
	Polyglutamylation	Tubulin
Ser	Phosphorylation	Protein serine kinases and phosphotases

	O-glycosylation	Notch O-glycosylation
	Phosphopantetheinylation	Fatty acid synthase
	Autocleavages	Pyruvamidyl enzyme formation
Thr	Phosphorylation	Protein threonine kinases/phosphatses
	O-glycosylation	
Tyr	Phosphorylation	Tyrosine kinases/phosphatses
	Sulfation	CCR5 receptor maturation
	Ortho-nitration	Inflammatory responses
	TOPA quinine	Amine oxidase maturation
His	Phosphorylation	Sesnsor protein kinases in two-cmoponent regulatory systems
	Aminocarboxypropylation	Dipthamide formation
	N-methylation	Methyl CoM reductase
Lys	N-methylation	
	N-acylation by acetyl, biotinyl, lipoyl, ubiquityl groups	Histone acetylation, swinging-arm prosthetic gropus, ubiquitin, SUMO tagging of proteins
	C-hydroxylation	Collagen maturation
Cys	S-hydroxylation (S-OH)	Sufenate intermediates
	Disulfide bond formation	Protein in oxidising environments
	Phosphorylation	PTPases
	S-acylation	Ras
	S-prenylation	Ras
	Protein splicing	Intein excisions
Met	Oxidation to sulfoxide	Met sulfoxide reductase
Arg	N-methylation	Histones
	N-ADP ribosylation	Gsα
Asn	N-glycosylation	N-glycoproteins
	N-ADP ribosylation	eEF-2
	Protein splicing	Intein excision step
Gln	Transglutamination	Protein cross-linking
Trp	C-mannosylation	Plasma membrane proteins
Pro	C-hydroxylation	Collagen, HIF-1α
	Cis-trans isomerisation	Protein prolyl isomerases
Gly	C-hydroxylation	C-terminal amide formation

These PTMs can potentially change the activity and/or localisation of the proteins acting as regulators of various biological processes in the cell. Exosomes/microvesicles originate from specialised compartments of the cells involved in regulatory and communication processes of the various target cells. It is, therefore, extremely important to characterise PTMs of proteins found in exosomes/microvesicles because they will reflect upon the functions of these vesicles and also provide information about classes of proteins sorted to these vesicles.

Protein modifications are not homogenous and a single gene can give rise to many sub-populations of a protein by alternative splicing of the gene and a single or multiple modifications of the protein residues. Usually only a fraction of a protein is modified by PTMs. Therefore, the amount of modified protein is often not sufficient for complete characterization by MS. As a result, recombinant proteins are often used as a source for such studies. However *in vitro*-modified proteins can be significantly different from the *in vivo* modifications. Two-dimensional electrophoresis (2DE) is often used to separate different modification states of a protein population followed by their characterization by MS. For example, phosphorylated proteins on 2DE leave a horizontal trail owing to charge differences among modified and non-modified forms. N- and C-terminal processing of enolase was studied with a combination of 2DE and MS and more than 10 different forms were defined (Larsen *et al.,* 2001). Antibody precipitation and chromatographic methods are also used to enrich the modified proteins followed by their identification or characterization by MS. The isolation method selection mainly depends upon the type of modification in question. Anti-phosphotyrosine antibody can enrich phosphorylated tyrosine-containing proteins (Pandey *et al.,* 2000) while an anti-phosphoamino acid antibody can be used for detection of phosphoserine or phosphothreonine (Gronborg *et al.,* 2002). However, glycosylphosphatidylinositol (GPI)-anchored proteins are usually released by specific enzymes from the cell surfaces and, after removing the cells by centrifugation, proteins in the supernatant can be identified. This type of analysis will yield a global identification of GPI-anchored proteins. Transfection of cells with a tagged version of ubiquitin or SUMO proteins and subsequent tag-based purification yields all proteins covalently attached to these modifiers (Vertegaal *et al.,* 2006) which can be subsequently identified by MS. If the enrichment protocol is designed carefully it eliminates the need for further complex characterization by MS and only identification of enriched proteins is often sufficient. There

have been very few studies characterising PTM of the exosomal/microvesicle proteins. These are discussed according to the type of PTM involved.

1.3.1 Glycosylation

One of the most common and important type of PTM is glycosylation in which proteins contain one or more heterosaccharides covalently bound to the polypepide backbone. These conjugated proteins are called glycoproteins. These proteins are present in all forms of life and almost every type of tissues and cells as well as in cell secretion products. These heterosachharides conjugated to proteins are called glycans and they make upto 10-60% of the weight of glycoproteins. Two major types of glycosylations are N-linked and O-linked glycosylation. The N-linked refers to the bond of glycan with a nitrogen atom in the amino acid asparagine while O-linked refers to bonding of glycans to oxygen atom in serine or threonine. N-linked glycoproteins contains β-glycosidic linkage between a GlcNAc residue and the δ amide N of an asparagine (Asn) side chain. The common feature between different types of N-linked glycans is a pentasaccharide (Man_3 $GlcNAc_2$) core which reflects common precursors to all forms. N-linked glycans fall into three main subclasses as shown in Figure 1.14.

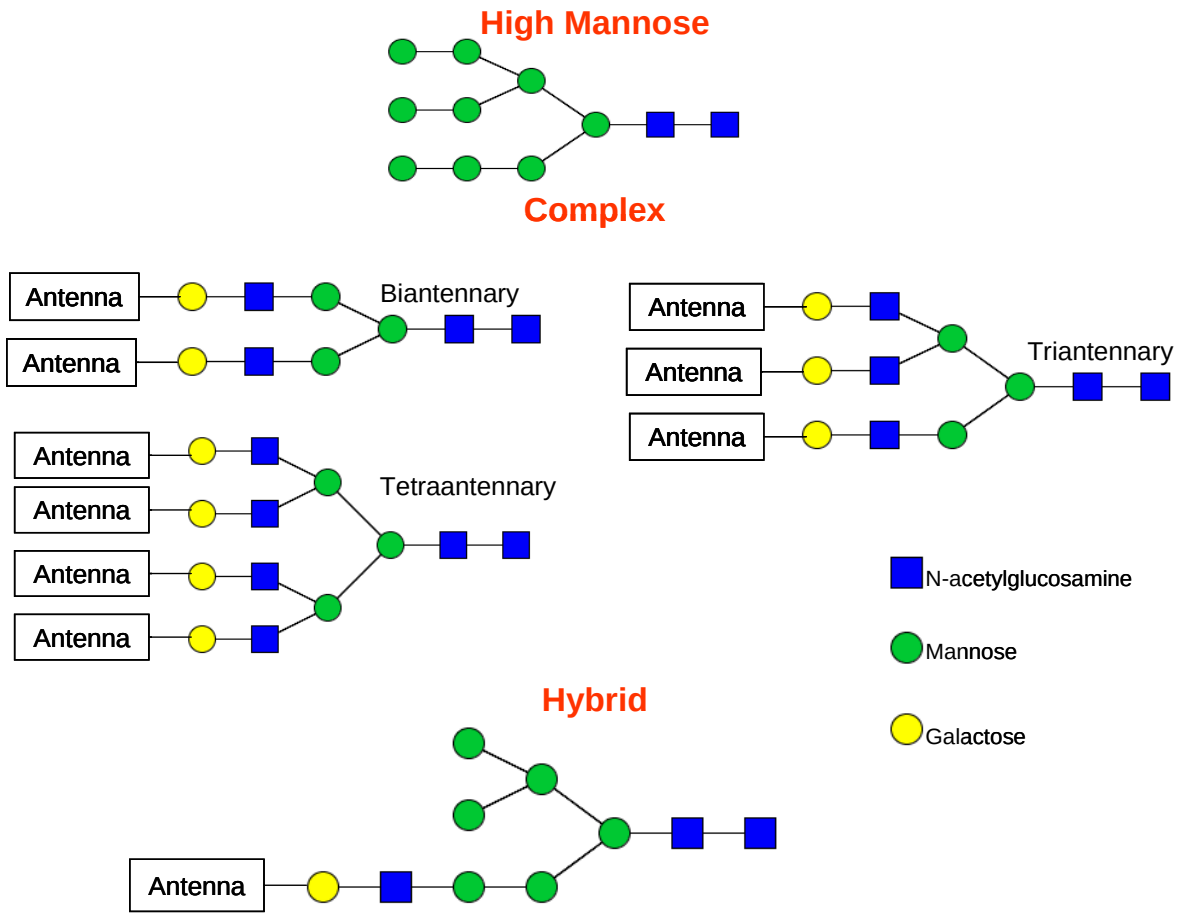

Figure 1.14: Structural varieties in N-glycosylation.

Polylactosamine extensions are composed of number of (Galβ1-4GlcNAcβ1-)n disaccharide units within one of the antenna. Alternatively, GlcNAc residue directly attached to the Asn residue or on the antennae could be derivatised with the monosaccharide fucose. Despite the structural composition of each antenna, all N-linked glycans may be capped with an N-acetylneuraminic acid- galactose (NeuAc-Gal) disaccharide. The variety of sialic acids, containing NeuAc residue, is either α2-3 or α2-6 linked to the penultimate galactose. A summary of the various potential modifications that can occur are shown in Figure 1.15.

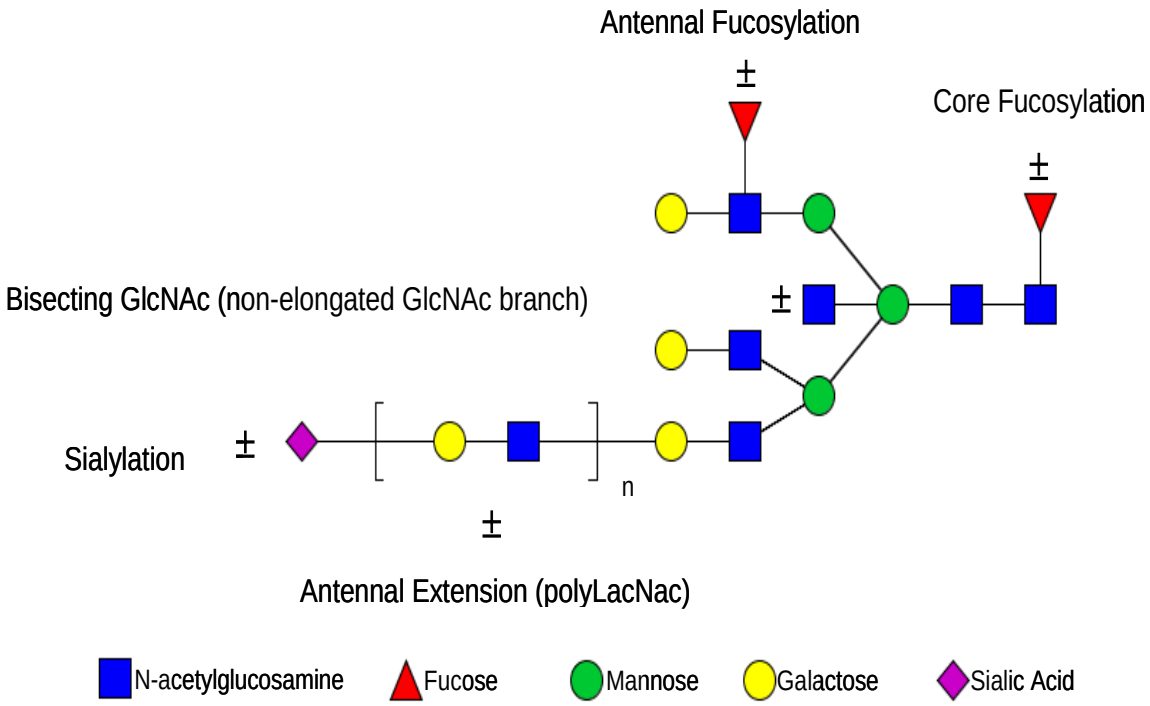

Figure 1.15: Possible modifications of the N-glycosylation glycan structure.

O-linked glycans are formed by an α-glycosidic linkage between a GalNAc and the hydroxyl group of a Ser or Thr residue. Following attachment of the GalNAc residue to Ser/Thr, a Galβ1-3 attachment to the GalNac forms what is termed a core 1 O-linked glycan. This is either disialyated to complete the core 1 structure or further extended to produce a core 2 O-linked glycan. Core 1 and 2 structures may be seen in Figure 1.16.

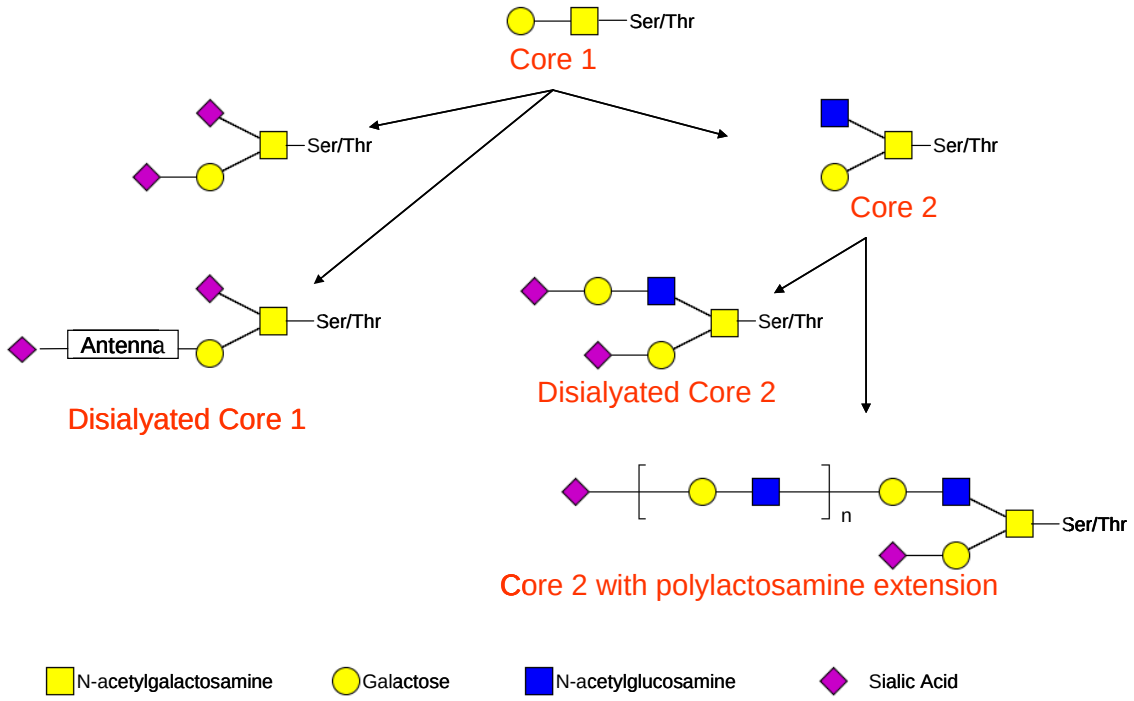

Figure 1.16: various core structures of O-glycosylation.

Unlike N-glycosylation, there is no preformed core structure that is formed before the final glycan structure is established. O-linked glycosylation unlike N-linked glycosylation takes place completely in golgi apparatus. O-linked glycans contain non-reducing α-linked sugars (sialic acid and fucose) while mannose and glucose are absent.

Glycosylation is a dynamic PTM and it is believed that it changes upon alterations in physiological state of an organism are reflected in altered expression of glycans or a change in their type. Several hereditary diseases show alterations in glycosylation associated with mutant phenotype. For example, persons with Carbohydrate Deficiency Glycoprotein Syndrome (CDGS) exhibit abnormal glycosylation of proteins like α-1 antitrypsin and transferrin (Jaeken & Carchon, 1993). For other diseases, increase in sialylation and branching of glycans expressed by cancer cells has been widely observed and it correlates directly with metastatic capacity of these cells (Troy, 1992). The glycans on Fc region of

immunoglobulins (IgG) lacking in galactose have been observed to be increased in rheumatoid arthritis (Youings *et al.,* 1996). Incomplete N-glycosylation in cystic fibrosis transmembrane conductance regulator (CFTR) protein leads to ER retention of the protein and premature degradation (Jilling & Kirk, 1997) and this contributes to symptoms of cystic fibrotic. In case of cardiovascular diseases, abnormal glycosylation of fibrinogen in hepatoma can lead to dysfibrinogenemia (Gralnick, Givelber & Abrams, 1978) and also impaired fibrin polymerization (Maekawa *et al.,* 1992; Ridgway *et al.,* 1997). In case of IgA nephropathy it has been found that there is overrepresentation of IgA1 O-glycoforms in serum which are poorly galactosylated. It is suggested that these glycoforms either act as autoantigens driving the formation of glycan specific antibodies or as antigens for cross reacting antimicrobial antibodies and subsequent formation of immune complexes which take part in pathogenesis of the disease (Barrett & Feehally, 2011). Efficient and appropriate glycosylation of several kidney proteins is essential for their function. For example, altered glycosylation of rat thiazide-sensitive Na-Cl cotransporter (rNCC) affects its normal function and cell surface expression in rat renal distal convoluted tubule (Hoover *et al.,* 2003). Glycosylation is an important PTM to be studied as it can lead to novel biomarkers of various diseases as well as enhanced understanding of the pathogenesis.

1.3.1.1 Glycomic approaches to study glycosylation

There is an increasing amount of interest in characterizing glycans of various glycoproteins coupled with identification of these proteins. There are LC and MS approaches to study of glycans either in purified proteins or from a complex mixture like cell lysates or body fluids. This, however, requires liberation of glycans from their corresponding proteins followed by reducing terminal labelling for detection and separation. Inability of these techniques to assign whole cell or body fluid glycans to their corresponding proteins is one major drawback. If one protein at a time is analysed, it reduces the throughput and requires purified

proteins in sufficient amount because glycans usually are a small but distinctive portion of a whole protein's mass. Moreover, this procedure is not helpful in the case of O-linked glycans as protocols for their liberation are based on chemical removal and the reproducibility of such techniques is low. In addition, glycans of glycoproteins show high degree of heterogeneity.

Another approach to glycomic analysis is the use of carbohydrate binding proteins (anti-carbohydrate antibodies and lectins) to fractionate the proteins prior to identification with MS. Antibodies are not very useful for this prupose as most of the glycans and components thereof are evolutionarily conserved and evoke minimal immunogenicity in antibody-producing host species. A notable exception is the natural antibodies to blood group determinants different than the host blood group. Lectins are sugar-binding proteins having distinct specificity to one type of sugar conjugate over others. This specificity varies among different lectins. Use of lectins also has limitations in identifying the composition of glycans attached to target proteins. These limitations include the inability to infer complete glycan composition from lectin binding. Lectin microarray is one rapidly emerging technology for profiling of total glycans of complex mixtures like cell lysates and differences among pathological states compared to the normal physiological state, present an avenue for development of highly sensitive glyco-biomarkers. A definite advantage is the possibility for high-throughput analysis of glycans by the lectin microarrays. The respective biomarkers discovered have an advantage over protein biomarkers as changes in glycan profile e.g. in response to a disease course are rapid and early detection of anomalous structures is possible. However, there are many challenges associated with technical and biological variabilities. Typically lectins are immobilised or printed (using techniques like that of ink-jet printing) on a substrate surface or glass slide and pre-labelled glycoprotein, or a mixture thereof, is hybridised to them (Hirabayashi, Kuno & Tateno, 2011). Fluorescence intensity can be measured to obtain quantitative information about type of glycans present on a purified

protein or the whole glycans of a mixture like a cell lysate. Controls can be compared to a test population of samples to evaluate differences and similarities among them. Figure 1.17 presents various schemes that can be implemented in lectin microarray.

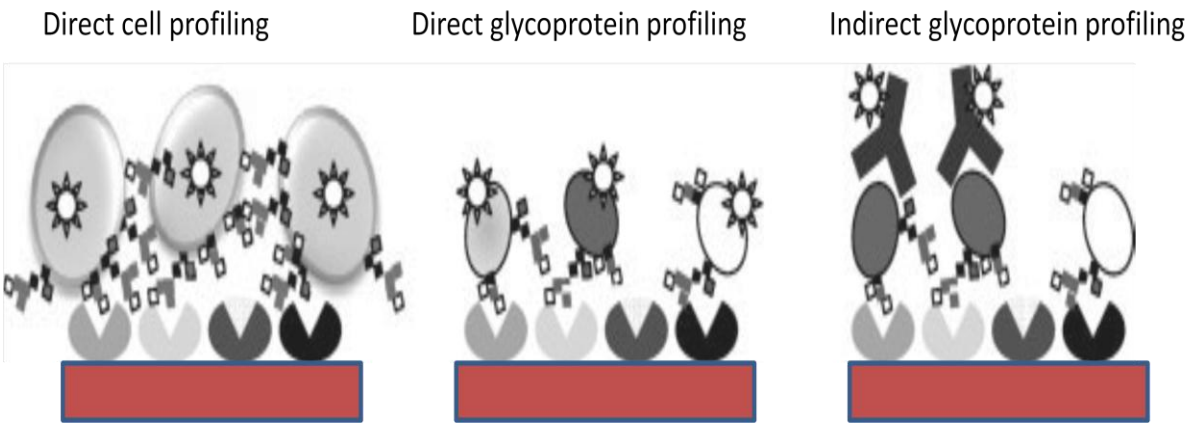

Figure 1.17: Various schemes of lectin microarray including direct cell glycan profiling and direct and indirect glycoprotein profiling. Lectins are immobilised on a substrate and labelled cells (direct cell profiling) or glycoprotein (s) (direct glycoprotein profiling) are added. In another approach glycoproteins captured by lectins can be probed further with a specific antibody (indirect glycoprotein profiling). The figure has been taken from (Hirabayashi, Kuno & Tateno, 2011)

Lectin affinity chromatography coupled to MS can also be used to determine the nature of glycans along with corresponding glycoproteins. This is important because, once native glycosylation of a protein is known; it becomes much easier to develop antibody and lectin-based hybrid microtitre plate assays to determine rapidly whether glycosylation of target protein is changing in a given pathological state. Lectins have affinity for distinct glycan epitopes (Sharon & Lis, 1989) and this property can be used to isolate, fractionate and analyse complex mixtures of glycoproteins using e.g. lectin-agarose or magnetic beads with immobilised lectins (Wiener & vanHoek, 1996; Bundy & Fenselau, 2001). Lectins only enrich a selective fraction of glycoproteins and, therefore, broad specificity lectins are

generally used (Mechref *et al.,* 2000; Yamamoto, Tsuji & Osawa, 1998). However, to enable complete analysis of most glycoproteins present in a complex biological sample like biofluids, multi-lectin columns are needed (Yang & Hancock, 2004; Wang, Wu & Hancock, 2006). Another method based on hydrazide chemistry can also be used which oxidises the carbohydrates attached to glycoproteins and enables their subsequent attachment to a hydrazide resin. The N-linked proteins or peptides can subsequently be released by peptide N-glycosidase F (PNGase F) (Zhang *et al.,* 2003; Tian *et al.,* 2007). Lectin and hydrazide chemistry-based enrichment both have their advantages and disadvantages. For example, in a comparative study to analyse rat liver membrane proteins, it was found that lectins mostly enriched high molecular weight glycoproteins while the hydrazide method enriched mostly low molecular weight proteins (Lee *et al.,* 2009). Therefore, both these methods should be used in tandem if possible for optimal results. Reaction with boronic acid can also be used to enrich glycoproteins and it was applied to enrich low-abundance glycoproteins from human blood samples (Sparbier, Wenzel & Kostrzewa, 2006). Coupling of high-resolution MS instruments to these enrichment methods can be used for qualitative and/or quantitative analysis of glycoproteins (Wang, Wu & Hancock, 2006).

1.3.1.2 Exosome/Microvesicle glycosylation

The glycan signature of the exosome/microvesicle reflects the cell secreting them and differences among and between cells and between cells and the vesicles, respectively. SKOV3 ovarian carcinoma cell-derived exosomes are remarkably enriched in high-mannose or NeuAcα2, 3/6-containing proteins compared to parent cells (Escrevente *et al.,* 2011) although the full identity of these proteins is not yet known. Exosome uptake by dendritic cells can be specifically inhibited by mannose and N-acetylglucosamine where interaction between cell and exosome is mediated by a C-type lectin (Hao *et al.,* 2007). Galectin-5 mediates the uptake of exosomes by macrophages and this process can be inhibited by

supplying lactose, the nominal glycan specificity of this lectin, to the medium (Barres *et al.,* 2010). This suggests that glycosylation affects the uptake of exosome/microvesicles in a cell type-dependent manner. Glycosylation is also an important "address code" signal for microvesicle cargo. This is supported by the fact that distinct glycoforms of prion protein (PrP) are incorporated preferentially over other forms into exosomes (Vella *et al.,* 2007). Lectin microarrays were employed to reveal the glycan expression pattern of microvesicles, cell membranes and HIV-1 particles from matched cells. Notably, it was found that HIV-1 particles and microvesicles (MV) share similar glycosylation, different from the parent cells (Krishnamoorthy *et al.,* 2009). Microvesicles and HIV-1 particles were enriched in high-mannose epitopes (*Galanthus nivalis* lectin {GNA, *Narcissus pseudonarcissus* lectin {NPA, Hippeastrum hybrid lectin {HHL, Concanavalin-A {Con-A, *Pisum sativum* agglutinin {PSA, cyanovirin-N {CVN, scytovirin {SVN} and griffithsin {GRFT) compared to parent H9 cell membrane. MV and HIV were also enriched in complex N-linked glycans (*Phaseolus vulgaris* erythroagglutinin {PHA-E} and leukoagglutinin {PHA-L), N-acetyllactosamine (*Datura stramonium* agglutinin {DSA, *Maackia amurensis* lectin {MAL-I, *Ricinus communis* agglutinin {RCA, *Solanum tuberosum* lectin {STA, Wheat Germ agglutinin {WGA), sialic acid (*Sambucus nigra* lectin {SNA, MAA, MAL-II), and fucosylated (*Ulex europaeus* agglutinin I {UEA-I, *Psophocarpus tetragonolobus* lectin II {PTL-II, *Aleuria aurantia* lectin {AAL) epitopes. Moreover blood antigens A/B (*Euonymus europaeus* lectin {EEA, *Phaseolus lunatus* lectin {LBA) were absent in microvesicles and HIV-1 while being present on the cell membrane. These glycan epitopes enriched in MV and HIV-1 colocalized to specific membrane microdomains suggesting that MV arise from specific domains on PM and HIV-1 hijacks the glycosylation machinery of MV for its advantage. The immune system would see HIV-1 glycans as MV and would not attack the viral particles. The glycan signatures among different cell lines (all T-cell lines) were similar in MV but some

differences were also present while in the parent cell membrane glycosylation was different among cells. It suggests that sorting of glycoproteins to these vesicles is at least in some ways glycosylation-specific and may serve as an important mediator of communication between cells including that mediated by exosome/microvesicle uptake. Another study by the same group, this time using a wide variety of cell line and human breast milk as source of MV, reinforced these results showing that MV have a conserved glycan pattern which is different from the parent cell membrane (Batista *et al.,* 2011). No other studies have been done on glycosignature of exosome/microvesicles or glycoproteomics. Notably these studies did not identify the glycoprotein constituents of the vesicles.

1.3.2 Phosphorylation

Mammalian phosphoproteome (all the proteins modified by phosphorylation of at least one or more amino acids) typically contains phosphoSerine (pS), phosphoThreonine (pT) and phosphoTyrosine (pY) residues with the majority being pS and pT (90:10 pS,pT/pY) while bacteria and fungi also have phosphoHistidine (pHis) and phosphoAspartic acid (pAsp) (Mann *et al.,* 2002). pAsp is also found in some mammalian proteins but it is rare in mammals. These phosphorylated amino acids are shown in Figure 1.18.

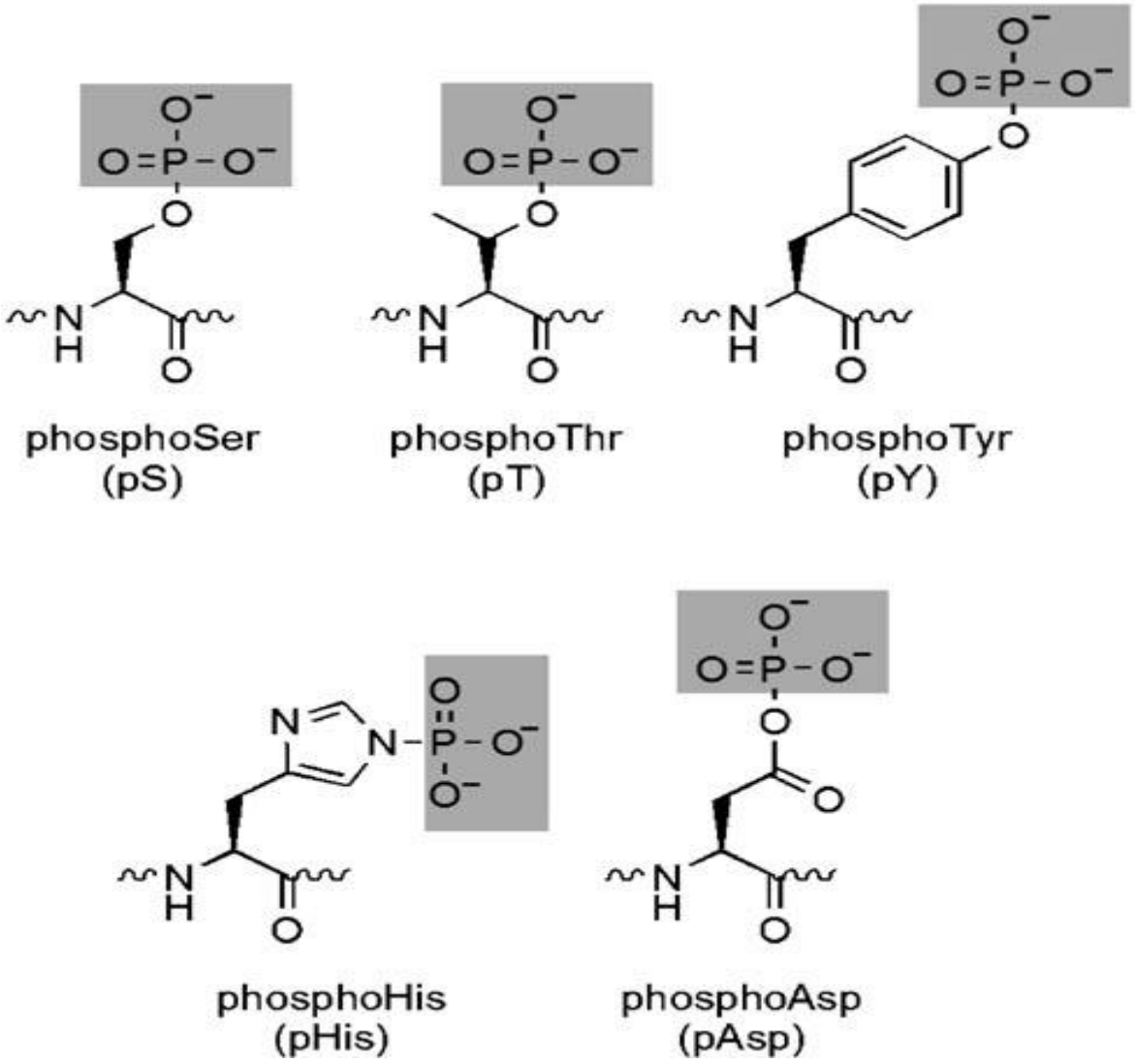

Figure 1.18: Various phosphorylated amino acid structures (Walsh, Garneau-Tsodikova & Gatto, 2005).

One protein can be phosphorylated at multiple residues making the phosphoproteome very complex. Over 500 different kinases (Phoshorylating enzymes) are known in the human proteome capable of influencing the phosphoprotein outcome, making this network even more complex. Introduction of phosphate group(s) to a protein can lead to altered conformation and, thus, also protein function (Johnson & Lewis, 2001) e.g. inducing initiation or termination of a signal in a signalling cascade. There has been only one study in exosomes which has identified phosphoproteins and their site of modification in urinary

exosomes (Gonzales *et al.,* 2009). Gonzalez *et al.,* used mass spectrometric technique of neutral loss scanning with high-stringency target-decoy analysis to identify phosphorylation sites in urinary exosomal proteins. Nineteen phosphorylation sites corresponding to fourteen exosomal proteins were identified in the study. Orphan G-protein coupled receptors (GPRC5B and GPRC5C) were among the proteins on which these sites were identified including one new site in GPRC5B (T389) and three new sites in GPRC5C (T435, S395 and Y426). S811 was identified as a previously unknown phosphorylation site on carboxy terminal of thiazide-sentitive co-transporter (NCC) which is proposed to regulate transport. S256 was identified in aquaporin-2 (AQP-2) which was further confirmed by immunoblotting the exosomal sample with specific anti-AQP-2 antibodies. A list of all the indentified phosphorylation sites with corresponding proteins is shown in Table 1.4.

Table 1.4: Phosphoproteins identified in urinary exosomes by (Gonzales *et al.*, 2009)

Ref Seq	Protein Name, Sequence	Site	Gene	Novel site
NP_061123	G protein-coupled receptor family C, group 5, member C isoform b		GPRC5C	
	R.AEDMYSAQSHQAA (T*)PPKDGK.N	T435		Yes
	K.VP (S*)EGAYDIILPR.A	S395		Yes
	R.AEDM (Y*)SAQSHQAATPPKDGK.N	Y426		Yes
NP_001035149	secreted phosphoprotein 1 isoform C		SPP1	
	K.AIPVAQDLNAPSDWD (S*)R.G	S192		No
	R.GKD (S*)YETSQLDDQSAETHSHK.Q	S197		
	R.GKDSYETSQLDDQ (S*)AETHSHK.Q	S207		
NP_057319	G protein-coupled receptor family C, group 5, member b precursor		GPRC5B	
	R.SNVYQPTEMAVVLNGG (T*)IPTAPPSHTGR.H	T389		Yes
NP_000477	Aquaporin 2		AQP2	
	R.RQ (S*)VELHSPQSLPR.G	S256		No
NP_004860	Vacuolar protein sorting factor 4B		VPS4B	
	K.EGQPSPADEKGND (S*)DGEGESDDPEKKK.L	S102		Yes
NP_054762	Chromatin modifying protein 2B		CHMP2B	
	K.ATI (S*)DEEIER.Q	S199		No
NP_687033	Proteasome α 3 subunit isoform 2		PSAM3	
	K.ESLKEEDE (S*)DDDNM	S243		No
NP_036382	Related RAS viral (r-ras) oncogene homolog 2		RRAS2	
	R.KFQEQECPP (S*)PEPTRK.E	S186		Yes
NP_031381	Heat-shock 90-kD protein 1, β		HSP90AB1	
	K.IEDVG (S*)DEEDDSGKDKK.K	S255		No
NP_612433	Kinesin family member 12		KIF12	
	R.VTTRPQAPK (S*)PVAK.Q	S236		Yes
NP_079119	Cytochrome b reductase 1		CYBRD1	
	R.NLALDEAGQRS (T*)M.	T285		Yes
NP_001037857	Mucin 1 isoform 7 precursor		MUC1	
	R.DTYHPMSEYPTYH (T*)HGR.Y	T118		Yes
NP_000330	Solute carrier family 12 (sodium/chloride transporters), member 3		SLC12A3	
	R.GARP (S*)VSGALDPK.A	S811		Yes
NP_000329	Sodium potassium chloride co-transporter 2		SLC12A1	
	K.IEYYRN (T*)GSISGPK.V	T118		No
	K.IEYYRNTG (S*)ISGPK.V	S120		No

1.4 Reference

Al-Nedawi, K., Meehan, B., Micallef, J., Lhotak, V., May, L., Guha, A. & Rak, J. (2008), "Intercellular transfer of the oncogenic receptor EGFrvIII by microvesicles derived from tumour cells", *NATURE CELL BIOLOGY*, vol. 10, no. 5, pp. 619-624.

Amzallag, N., Passer, B., Allanic, D., Segura, E., Thery, C., Goud, B., Amson, R. & Telerman, A. (2004), "TSAP6 facilitates the secretion of translationally controlled tumor protein/histamine-releasing factor via a nonclassical pathway", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 279, no. 44, pp. 46104-46112.

Andrews, R.K. & Berndt, M.C. (2004), "Platelet physiology and thrombosis", *THROMBOSIS RESEARCH*, vol. 114, no. 5-6, pp. 447-453.

Baroni, M., Pizzirani, C., Pinotti, M., Ferrari, D., Adinolfi, E., Calzavarini, S., Caruso, P., Bernardi, F. & Di Virgilio, F. (2007), "Stimulation of P2 (P2X(7)) receptors in human dendritic cells induces the release of tissue factor-bearing microparticles", *FASEB JOURNAL*, vol. 21, no. 8, pp. 1926-1933.

Barres, C., Blanc, L., Bette-Bobillo, P., Andre, S., Mamoun, R., Gabius, H. & Vidal, M. (2010), "Galectin-5 is bound onto the surface of rat reticulocyte exosomes and modulates vesicle uptake by macrophages", *BLOOD*, vol. 115, no. 3, pp. 696-705.

Barrett, J. & Feehally, J. (2011), "Primary IgA Nephropathy: New Insights Into Pathogenesis", *SEMINARS IN NEPHROLOG*, vol. 31, no. 4, pp. 349-360.

Barry, O., Pratico, D., Lawson, J. & FitzGerald, G. (1997), "Transcellular activation of platelets and endothelial cells by bioactive lipids in platelet microparticles", *JOURNAL OF CLINICAL INVESTIGATION*, vol. 99, no. 9, pp. 2118-2127.

Barry, O., Pratico, D., Savani, R. & FitzGerald, G. (1998), "Modulation of monocyte-endothelial cell interactions by platelet microparticles", *JOURNAL OF CLINICAL INVESTIGATION*, vol. 102, no. 1, pp. 136-144.

Barry, O.P., Kazanietz, M.G., PraticÃ², D. & FitzGerald, G.A. (1999), "Arachidonic Acid in Platelet Microparticles Up-regulates Cyclooxygenase-2-dependent Prostaglandin Formation via a Protein Kinase C/Mitogen-activated Protein Kinase-dependent Pathway", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 274, no. 11, pp. 7545-7556.

Batista, B.S., Eng, W.S., Pilobello, K.T., Hendricks-Munoz, K.D. & Mahal, L.K. (2011), "Identification of a Conserved Glycan Signature for Microvesicles", *JOURNAL OF PROTEOME RESEARCH*, vol. 10, no. 10, pp. 4624-4633.

Berckmans, R.J., Sturk, A., van Tienen, L.M., Schaap, M.C.L. & Nieuwland, R. (2011), "Cell-derived vesicles exposing coagulant tissue factor in saliva", *BLOOD*, vol. 117, no. 11, pp. 3172-3180.

Bianco, F., Perrotta, C., Novellino, L., Francolini, M., Riganti, L., Menna, E., Saglietti, L., Schuchman, E.H., Furlan, R., Clementi, E., Matteoli, M. & Verderio, C. (2009), "Acid sphingomyelinase activity triggers microparticle release from glial cells", *EMBO JOURNAL*, vol. 28, no. 8, pp. 1043-1054.

Bianco, F., Pravettoni, E., Colombo, A., Schenk, U., Moller, T., Matteoli, M. & Verderio, C. (2005), "Astrocyte-Derived ATP Induces Vesicle Shedding and IL-1β Release from Microglia", *The JOURNAL OF IMMUNOLOGY*, vol. 174, no. 11, pp. 7268-7277.

Booth, A., Fang, Y., Fallon, J., Yang, J., Hildreth, J. & Gould, S. (2006), "Exosomes and HIV Gag bud from endosome-like domains of the T cell plasma membrane", *JOURNAL OF CELL BIOLOGY*, vol. 172, no. 6, pp. 923-935.

Brenner, B.M. 1996, *The kidney*, W.B. Saunders.

Bundy, J. & Fenselau, C. (2001), "Lectin and carbohydrate affinity capture surfaces for mass spectrometric analysis of microorganisms", *ANALYTICAL CHEMISTRY*, vol. 73, no. 4, pp. 751-757.

Buschow, S., Liefhebber, J., Wubbolts, R. & Stoorvogel, W. (2005), "Exosomes contain ubiquitinated proteins", *BLOOD CELLS MOLECULES AND DISEASES*, vol. 35, no. 3, pp. 398-403.

Carney, D., Hammer, C. & Shin, M. (1986), "Elimination of terminal complement complexes in the plasma membrane of nucleated cells: influence of extracellular Ca2+ and association with cellular Ca2+", *The Journal of Immunology*, vol. 137, no. 1, pp. 263-270.

Cheruvanky, A., Zhou, H., Pisitkun, T., Kopp, J.B., Knepper, M.A., Yuen, P.S.T. & Star, R.A. (2007), "Rapid isolation of urinary exosomal biomarkers using a nanomembrane ultrafiltration concentrator", *AMERICAN JOURNAL OF PHYSIOLOGY-RENAL PHYSIOLOGY*, vol. 292, no. 5, pp. F1657-F1661.

Cocucci, E., Racchetti, G. & Meldolesi, J. (2009), "Shedding microvesicles: artefacts no more", *TRENDS IN CELL BIOLOGY*, vol. 19, no. 2, pp. 43-51.

Cocucci, E., Racchetti, G., Podini, P. & Meldolesi, J. (2007), "Enlargeosome Traffic: Exocytosis Triggered by Various Signals Is Followed by Endocytosis, Membrane Shedding or Both", *TRAFFIC*, vol. 8, no. 6, pp. 742-757.

Conde-Vancells, J., Rodriguez-Suarez, E., Embade, N., Gil, D., Matthiesen, R., Valle, M., Elortza, F., Lu, S.C., Mato, J.M. & Falcon-Perez, J.M. (2008), "Characterization and Comprehensive Proteome Profiling of Exosomes Secreted by Hepatocytes", *JOURNAL OF PROTEOME RESEARCH*, vol. 7, no. 12, pp. 5157-5166.

Connor, D.E., Exner, T., Ma, D.D.F. & Joseph, J.E. (2010), "The majority of circulating platelet-derived microparticles fail to bind annexin V, lack phospholipid-dependent procoagulant activity and demonstrate greater expression of glycoprotein Ib", *THROMBOSIS AND HAEMOSTASIS*, vol. 103, no. 5, SI, pp. 1044-1052.

Dean, W.L., Lee, M.J., Cummins, T.D., Schultz, D.J. & Powell, D.W. (2009), "Proteomic and functional characterisation of platelet microparticle size classes", *THROMBOSIS AND HAEMOSTASIS*, vol. 102, no. 4, pp. 711-718.

Degrooth, B., TERSTAPPEN, L., PUPPELS, G. & GREVE, J. (1987), "Light-scattering polarization measurements as a new parameter in flow-cytometry", *CYTOMETRY*, vol. 8, no. 6, pp. 539-544.

Del Conde, I., Shrimpton, C., Thiagarajan, P. & Lopez, J. (2005), "Tissue-factor-bearing microvesicles arise from lipid rafts and fuse with activated platelets to initiate coagulation", *BLOOD*, vol. 106, no. 5, pp. 1604-1611.

Dolo, V., Li, R., Dillinger, M., Flati, S., Manela, J., Taylor, B., Pavan, A. & Ladisch, S. (2000), "Enrichment and localization of ganglioside G(D3) and caveolin-1 in shed tumor cell membrane vesicles", *BIOCHIMICA ET BIOPHYSICA ACTA-MOLECULAR AND CELL BIOLOGY OF LIPIDS,* vol. 1486, no. 2-3, pp. 265-274.

Dolo, V., Ginestra, A., Cassara , D., Violini, S., Lucania, G., Torrisi, M.R., Nagase, H., Canevari, S., Pavan, A. & Vittorelli, M.L. (1998), "Selective Localization of Matrix Metalloproteinase 9,β1 Integrins, and Human Lymphocyte Antigen Class I Molecules on Membrane Vesicles Shed by 8701-BC Breast Carcinoma Cells", *CANCER RESEARCH,* vol. 58, no. 19, pp. 4468-4474.

Du Cheyron, D., Daubin, C., Poggioli, J., Ramakers, M., Houillier, P., Charbonneau, P. & Paillard, M. (2003), "Urinary measurement of Na+/H+ exchanger isoform 3 (NHE3) protein as new marker of tubule injury in critically ill patients with ARF", *AMERICAN JOURNAL OF KIDNEY DISEASES,* vol. 42, no. 3, pp. 497-506.

Dumaswala, U.J. & Greenwalt, T.J. (1984), "Human erythrocytes shed exocytic vesicles in vivow", *TRANSFUSION,* vol. 24, no. 6, pp. 490-492.

Eken, C., Gasser, O., Zenhaeusern, G., Oehri, I., Hess, C. & Schifferli, J.A. (2008), "Polymorphonuclear neutrophil-derived ectosomes interfere with the maturation of monocyte-derived dendritic cells", *JOURNAL OF IMMUNOLOGY,* vol. 180, no. 2, pp. 817-824.

Elliot, S., Goldsmith, P., Knepper, M., Haughey, M. & Olson, B. (1996), "Urinary excretion of aquaporin-2 in humans: A potential marker of collecting duct responsiveness to vasopressin", *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY,* vol. 7, no. 3, pp. 403-409.

Escrevente, C., Keller, S., Altevogt, P. & Costa, J. (2011), "Interaction and uptake of exosomes by ovarian cancer cells", *BMC CANCER,* vol. 11, pp. 108-117.

Fang, Y., Wu, N., Gan, X., Yan, W., Morrell, J.C. & Gould, S.J. (2007), "Higher-order oligomerization targets plasma membrane proteins and HIV gag to exosomes", *PLOS BIOLOGY,* vol. 5, no. 6, pp. 1267-1283.

Fernandez-Llama, P., Khositseth, S., Gonzales, P.A., Star, R.A., Pisitkun, T. & Knepper, M.A. (2010), "Tamm-Horsfall protein and urinary exosome isolation", *JOURNAL OF HYPERTENSION,* vol. 28, no. A, pp. E164.

Fevrier, B., Vilette, D., Archer, F., Loew, D., Faigle, W., Vidal, M., Laude, H. & Raposo, G. (2004), "Cells release prions in association with exosomes", *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA,* vol. 101, no. 26, pp. 9683-9688.

Filipe, V., Hawe, A. & Jiskoot, W. (2010), "Critical Evaluation of Nanoparticle Tracking Analysis (NTA) by NanoSight for the Measurement of Nanoparticles and Protein Aggregates", *PHARMACEUTICAL RESEARCH,* vol. 27, no. 5, pp. 796-810.

Frasch, S.C., Henson, P.M., Nagaosa, K., Fessler, M.B., Borregaard, N. & Bratton, D.L. (2004), "Phospholipid Flip-Flop and Phospholipid Scramblase 1 (PLSCR1) Co-localize to Uropod Rafts in Formylated Met-Leu-Phe-stimulated Neutrophils", *JOURNAL OF BIOLOGICAL CHEMISTRY,* vol. 279, no. 17, pp. 17625-17633.

Freyssinet, J. (2003), "Cellular microparticles: what are they bad or good for?", *JOURNAL OF THROMBOSIS AND HAEMOSTASIS,* vol. 1, no. 7, pp. 1655-1662.

Futter, C.E. & White, I.J. (2007), "Annexins and endocytosis", *TRAFFIC*, vol. 8, no. 8, pp. 951-958.

Gasser, O. & Schifferli, J. (2004), "Activated polymorphonuclear neutrophils disseminate anti-inflammatory microparticles by ectocytosis", *BLOOD*, vol. 104, no. 8, pp. 2543-2548.

Geminard, C., de Gassart, A., Blanc, L. & Vidal, M. (2004), "Degradation of AP2 during reticulocyte maturation enhances binding of hsc70 and Alix to a common site on TfR for sorting into exosomes", *TRAFFIC*, vol. 5, no. 3, pp. 181-193.

Goni, F.M. & Alonso, A. (2006), "Biophysics of sphingolipids I. Membrane properties of sphingosine, ceramides and other simple sphingolipids", *BIOCHIMICA ET BIOPHYSICA ACTA-BIOMEMBRANES*, vol. 1758, no. 12, pp. 1902-1921.

Gonzales, P.A., Pisitkun, T., Hoffert, J.D., Tchapyjnikov, D., Star, R.A., Kleta, R., Wang, N.S. & Knepper, M.A. (2009), "Large-Scale Proteomics and Phosphoproteomics of Urinary Exosomes", *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY*, vol. 20, no. 2, pp. 363-379.

Gralnick, H.R., Givelber, H. & Abrams, E. (1978), "Dysfibrinogenemia Associated with Hepatoma", *New England Journal of Medicine*, vol. 299, no. 5, pp. 221-226.

Griffiths, G. & Hoppeler, H. (1986), "Quantitation in immunocytochemistry - correlation of immunogold labeling to absolute number of membrane-antigens", *JOURNAL OF HISTOCHEMISTRY \& CYTOCHEMISTRY*, vol. 34, no. 11, pp. 1389-1398.

Gronborg, M., Kristiansen, T., Stensballe, A., Andersen, J., Ohara, O., Mann, M., Jensen, O. & Pandey, A. (2002), "A mass spectrometry-based proteomic approach for identification of serine/threonine-phosphorylated proteins by enrichment with phospho-specific antibodies - Identification of a novel protein, Frigg, as a protein kinase A substrate", *MOLECULAR \& CELLULAR PROTEOMICS*, vol. 1, no. 7, pp. 517-527.

Hao, S., Bai, O., Li, F., Yuan, J., Laferte, S. & Xiang, J. (2007), "Mature dendritic cells pulsed with exosomes stimulate efficient cytotoxic T-lymphocyte responses and antitumour immunity", *IMMUNOLOGY*, vol. 120, no. 1, pp. 90-102.

Hara, M., Yanagihara, T., Hirayama, Y., Ogasawara, S., Kurosawa, H., Sekine, S. & Kihara, I. (2010), "Podocyte membrane vesicles in urine originate from tip vesiculation of podocyte microvilli", *HUMAN PATHOLOGY*, vol. 41, no. 9, pp. 1265-1275.

Harding, C., Heuser, J. & Stahl, P. (1984), "Endocytosis and intracellular processing of transferrin and colloidal gold-transferrin in rat reticulocytes - demonstration of a pathway for receptor shedding", *EUROPEAN JOURNAL OF CELL BIOLOGY*, vol. 35, no. 2, pp. 256-263.

Hegmans, J., Bard, M., Hemmes, A., Luider, T., Kleijmeer, M., Prins, J., Zitvogel, L., Burgers, S., Hoogsteden, H. & Lambrecht, B. (2004), "Proteomic analysis of exosomes secreted by human mesothelioma cells", *AMERICAN JOURNAL OF PATHOLOGY*, vol. 164, no. 5, pp. 1807-1815.

Heijnen, H.F.G., Schiel, A.E., Fijnheer, R., Geuze, H.J. & Sixma, J.J. (1999), "Activated Platelets Release Two Types of Membrane Vesicles: Microvesicles by Surface Shedding and Exosomes Derived From Exocytosis of Multivesicular Bodies and îœ€-Granules", *BLOOD*, vol. 94, no. 11, pp. 3791-3799.

Hess, C., Sadallah, S., Hefti, A., Landmann, R. & Schifferli, J. (1999), "Ectosomes released by human neutrophils are specialized functional units", *JOURNAL OF IMMUNOLOGY*, vol. 163, no. 8, pp. 4564-4573.

Hirabayashi, J., Kuno, A. & Tateno, H. (2011), "Lectin-based structural glycomics: A practical approach to complex glycans", *ELECTROPHORESIS*, vol. 32, no. 10, Part 2, SI, pp. 1118-1128.

Hogan, M.C., Manganelli, L., Woollard, J.R., Masyuk, A.I., Masyuk, T.V., Tammachote, R., Huang, B.Q., Leontovich, A.A., Beito, T.G., Madden, B.J., Charlesworth, M.C., Torres, V.E., LaRusso, N.F., Harris, P.C. & Ward, C.J. (2009), "Characterization of PKD Protein-Positive Exosome-Like Vesicles", *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY*, vol. 20, no. 2, pp. 278-288.

Hoorn, E.J., Pisitkun, T., Zietse, R., Gross, P., Frokiaer, J., Wang, N.S., Gonzales, P.A., Star, R.A. & Knepper, M.A. (2005), "Prospects for urinary proteomics: Exosomes as a source of urinary biomarkers ", *NEPHROLOGY*, vol. 10, no. 3, pp. 283-290.

Hoover, R.S., Poch, E., Monroy, A., Va¡zquez, N., Nishio, T., Gamba, G. & Hebert, S.C. (2003), "N-Glycosylation at Two Sites Critically Alters Thiazide Binding and Activity of the Rat Thiazide-sensitive Na+:Cl- Cotransporter", *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY*, vol. 14, no. 2, pp. 271-282.

Ishikawa, S. & Schrier, R. (2003), "Pathophysiological roles of arginine vasopressin and aquaporin-2 in impaired water excretion", *CLINICAL ENDOCRINOLOGY*, vol. 58, no. 1, pp. 1-17.

Jaeken, J. & Carchon, H. (1993), "The carbohydrate-deficient glycoprotein syndromes - an overview", *JOURNAL OF INHERITED METABOLIC DISEASE*, vol. 16, no. 5, pp. 813-820.

Jarvis, S.M. & Young, J.D. (1982), "Nucleoside translocation in sheep reticulocytes and fetal erythrocytes: a proposed model for the nucleoside transporter", *THE JOURNAL OF PHYSIOLOGY*, vol. 324, no. 1, pp. 47-66.

Jarvis, S.M., Young, J.D., Ansay, M., Archibald, A.L., Harkness, R.A. & Simmonds, R.J. (1980), "Is inosine the physiological energy source of pig erythrocytes?", *BIOCHIMICA ET BIOPHYSICA ACTA (BBA) - BIOMEMBRANES*, vol. 597, no. 1, pp. 183-188.

Jilling, T. & Kirk, K. (1997), "The biogenesis, traffic, and function of the cystic fibrosis transmembrane conductance regulator" in *INTERNATIONAL REVIEW OF CYTOLOGY - A SURVEY OF CELL BIOLOGY*, vol. 172, pp. 193-241.

Johnson, L.N. & Lewis, R.J. (2001), "Structural Basis for Control by Phosphorylation", *Chemical reviews*, vol. 101, no. 8, pp. 2209-2242.

Johnstone, R. (2006), "Exosomes biological significance: A concise review", *BLOOD CELLS MOLECULES AND DISEASES*, vol. 36, no. 2, pp. 315-321.

Johnstone, R., Adam, M., Hammond, J., Orr, L. & Turbide, C. (1987), "Vesicle formation during reticulocyte maturation - association of plasma-membrane activities with released vesicles (exosomes)", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 262, no. 19, pp. 9412-9420.

Johnstone, R., Bianchini, A. & Teng, K. (1989), "Reticulocyte maturation and exosome release - transferrin receptor containing exosomes shows multiple plasma-membrane functions", *BLOOD*, vol. 74, no. 5, pp. 1844-1851.

Kahner, B.N., Dorsam, R.T. & Kunapuli, S.P. (2008), "Role of P2Y receptor subtypes in platelet-derived microparticle generation", *FRONTIERS IN BIOSCIENCE-LANDMARK*, vol. 13, pp. 433-439.

Kaimori, J., Nagasawa, Y., Menezes, L.F., Garcia-Gonzalez, M.A., Deng, J., Imai, E., Onuchic, L.F., Guay-Woodford, L.M. & Germino, G.G. (2007), "Polyductin undergoes notch-like processing and regulated release from primary cilia", *HUMAN MOLECULAR GENETICS*, vol. 16, no. 8, pp. 942-956.

Kanno, K., Sasaki, S., Hirata, Y., Ishikawa, S., Fushimi, K., Nakanishi, S., Bichet, D. & Marumo, F. 1995, "Urinary-excretion of aquaporin-2 in patients with diabetes-insipidus", *NEW ENGLAND JOURNAL OF MEDICINE*, vol. 332, no. 23, pp. 1540-1545.

Karlsson, M., Lundin, S., Dahlgren, U., Kahu, H., Pettersson, I. & Telemo, E. (2001), "``Tolerosomes" are produced by intestinal epithelial cells", *EUROPEAN JOURNAL OF IMMUNOLOGY*, vol. 31, no. 10, pp. 2892-2900.

Kawamoto, S., Matsumoto, Y., Mizuno, K., Okubo, K. & Matsubara, K. (1996), "Expression profiles of active genes in human and mouse livers", *GENE*, vol. 174, no. 1, pp. 151-158.

Kendall, D. & Macdonald, R. (1982), "A fluorescence assay to monitor vesicle fusion and lysis", *journal of biological chemistry*, vol. 257, no. 23, pp. 3892-3895.

Kim, S.H., Bianco, N.R., Shufesky, W.J., Morelli, A.E. & Robbins, P.D. (2007), "Effective treatment of inflammatory disease models with exosomes derived from dendritic cells genetically modified to express IL-4", *JOURNAL OF IMMUNOLOGY*, vol. 179, no. 4, pp. 2242-2249.

Kim, S., Bianco, N., Menon, R., Lechman, E., Shufesky, W., Morelli, A. & Robbins, P. (2006), "Exosomes derived from genetically modified DC expressing FasL are anti-inflammatory and immunosuppressive", *MOLECULAR THERAPY*, vol. 13, no. 2, pp. 289-300.

Kim, S., Lechman, E., Bianco, N., Menon, R., Keravala, A., Nash, J., Mi, Z., Watkins, S., Gambotto, A. & Robbins, P. (2005), "Exosomes derived from IL-10-treated dendritic cells can suppress inflammation and collagen-induced arthritis", *JOURNAL OF IMMUNOLOGY*, vol. 174, no. 10, pp. 6440-6448.

Knepper, M. & Burg, M. (1983), "Organization of nephron function", *AMERICAN JOURNAL OF PHYSIOLOGY*, vol. 244, no. 6, pp. F579-F589.

Krishnamoorthy, L., Bess, J.,Julian W., Preston, A.B., Nagashima, K. & Mahal, L.K. (2009), "HIV-1 and microvesicles from T cells share a common glycome, arguing for a common origin", *NATURE CHEMICAL BIOLOGY*, vol. 5, no. 4, pp. 244-250.

Lakkaraju, A. & Rodriguez-Boulan, E. (2008), "Itinerant exosomes: emerging roles in cell and tissue polarity", *TRENDS IN CELL BIOLOGY*, vol. 18, no. 5, pp. 199-209.

Larsen, M.R., Larsen, P.M., Fey, S.J. & Roepstorff, P. (2001), "Characterization of differently processed forms of enolase 2 from Saccharomyces cerevisiae by two-dimensional gel electrophoresis and mass spectrometry", *ELECTROPHORESIS*, vol. 22, no. 3, pp. 566-575.

Lee, A., Kolarich, D., Haynes, P.A., Jensen, P.H., Bakert, M.S. & Packer, N.H. (2009), "Rat Liver Membrane Glycoproteome: Enrichment by Phase Partitioning and Glycoprotein Capture", *JOURNAL OF PROTEOME RESEARCH*, vol. 8, no. 2, pp. 770-781.

Lee, T.L., Lin, Y.C., Mochitate, K. & Grinnell, F. (1993), "Stress-relaxation of fibroblasts in collagen matrices triggers ectocytosis of plasma membrane vesicles containing actin, annexins II and VI, and beta 1 integrin receptors", *JOURNAL OF CELL SCIENCE*, vol. 105, no. 1, pp. 167-177.

Lescuyer, P., Pernin, A., Hainard, A., Bigeire, C., Burgess, J.A., Zimmermann-Ivol, C., Sanchez, J., Schifferli, J.A., Hochstrasser, D.F. & Moll, S. (2008), "Proteomic analysis of a podocyte vesicle-enriched fraction from human normal and pathological urine samples", *PROTEOMICS Clinical Applications*, vol. 2, no. 7-8, pp. 1008-1018.

Lin, X., Almqvist, N. & Telemo, E. (2005), "Human small intestinal epithelial cells constitutively express the key elements for antigen processing and the production of exosomes", *BLOOD CELLS MOLECULES AND DISEASES*, vol. 35, no. 2, pp. 122-128.

Liu, Y., Shah, S.V., Xiang, X., Wang, J., Deng, Z., Liu, C., Zhang, L., Wu, J., Edmonds, T., Jambor, C., Kappes, J.C. & Zhang, H. (2009), "COP9-Associated CSN5 Regulates Exosomal Protein Deubiquitination and Sorting", *AMERICAN JOURNAL OF PATHOLOGY*, vol. 174, no. 4, pp. 1415-1425.

Lutz, H., Liu, S. & Palek, J. (1977), "Release of spectrin-free vesicles from human erythrocytes during ATP depletion: 1. characterization of spectrin-free vesicles", *THE JOURNAL OF CELL BIOLOGY*, vol. 73, no. 3, pp. 548-560.

MacKenzie, A., Wilson, H.L., Kiss-Toth, E., Dower, S.K., North, R.A. & Surprenant, A. (2001), "Rapid Secretion of Interleukin-1β by microvesicle Shedding", *IMMUNITY*, vol. 15, no. 5, pp. 825-835.

Maekawa, H., Yamazumi, K., Muramatsu, S., Kaneko, M., Hirata, H., Takahashi, N., Arochapinango, C., Rodriguez, S., Nagy, H., Perezrequejo, J. & Matsuda, M. (1992), "Fibrinogen lima - a homozygous dysfibrinogen with an a-alpha-arginine-141 to serine substitution associated with extra n-glycosylation at a-alpha-asparagine-139 - impaired fibrin gel formation but normal fibrin-facilitated plasminogen activation catalyzed by tissue-type plasminogen-activator", *JOURNAL OF CLINICAL INVESTIGATION*, vol. 90, no. 1, pp. 67-76.

Mallegol, J., van Niel, G. & Heyman, M. (2005), "Phenotypic and functional characterization of intestinal epithelial exosomes", *BLOOD CELLS MOLECULES AND DISEASES*, vol. 35, no. 1, pp. 11-16.

Mallegol, J., Van Niel, G., Lebreton, C., Lepelletier, Y., Candalh, C., Dugave, C., Heath, J.K., Raposo, G., Cerf-Bensussan, N. & Heyman, M. (2007), "T84-intestinal epithelial exosomes bear MHC class II/peptide complexes potentiating antigen presentation by dendritic cells", *GASTROENTEROLOGY*, vol. 132, no. 5, pp. 1866-1876.

Mann, M., Ong, S., Gronborg, M., Steen, H., Jensen, O. & Pandey, A. (2002), "Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome", *TRENDS IN BIOTECHNOLOGY*, vol. 20, no. 6, pp. 261-268.

Marimuthu, A., Oâ€™Meally, R.N., Chaerkady, R., Subbannayya, Y., Nanjappa, V., Kumar, P., Kelkar, D.S., Pinto, S.M., Sharma, R., Renuse, S., Goel, R., Christopher, R., Delanghe, B., Cole, R.N., Harsha, H.C. & Pandey, A. (2011), "A Comprehensive Map of the Human Urinary Proteome", *JOURNAL OF PROTEOME RESEARCH*, vol. 10, no. 6, pp. 2734-2743.

Martin, S., Tesse, A., Hugel, B., Martinez, M., Morel, O., Freyssinet, J. & Andriantsitohaina, R. (2004), "Shed membrane particles from T lymphocytes impair endothelial function and regulate endothelial protein expression", *CIRCULATION*, vol. 109, no. 13, pp. 1653-1659.

Mashburn, L. & Whiteley, M. (2005), "Membrane vesicles traffic signals and facilitate group activities in a prokaryote", *NATURE*, vol. 437, no. 7057, pp. 422-425.

Mears, R., Craven, R., Hanrahan, S., Totty, N., Upton, C., Young, S., Patel, P., Selby, P. & Banks, R. (2004), "Proteomic analysis of melanoma-derived exosomes by two-dimensional polyacrylamide gel electrophoresis and mass spectrometry", *PROTEOMICS*, vol. 4, no. 12, pp. 4019-4031.

Mechref, Y., Zidek, L., Ma, W. & Novotny, M. (2000), "Glycosylated major urinary protein of the house mouse: characterization of its N-linked oligosaccharides", *GLYCOBIOLOGY*, vol. 10, no. 3, pp. 231-235.

Miki, K. & Sudo, A. (1998), "Effect of urine pH, storage time, and temperature on stability of catecholamines, cortisol, and creatinine", *CLINICAL CHEMISTRY*, vol. 44, no. 8, pp. 1759-1762.

MORGAN, B. (1989), "Complement membrane attack on nucleated cells - resistance, recovery and non-lethal effects", *BIOCHEMICAL JOURNAL*, vol. 264, no. 1, pp. 1-14.

Morgan, B., Imagawa, D., Dankert, J. & Ramm, L. 1986, "Complement lysis of U937, a nucleated mammalian-cell line in the absence of C9 - effect of C9 on C5B-8 mediated cell-lysis", *JOURNAL OF IMMUNOLOGY*, vol. 136, no. 9, pp. 3402-3406.

Moskovich, O. & Fishelson, Z. (2007), "Live cell imaging of outward and inward vesiculation induced by the complement C5b-9 complex", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 282, no. 41, pp. 29977-29986.

Muntasell, A., Berger, A.C. & Roche, P.A. (2007), "T cell-induced secretion of MHC class II-peptide complexes on B cell exosomes", *EMBO JOURNAL*, vol. 26, no. 19, pp. 4263-4272.

Nauta, A., Trouw, L., Daha, M., Tijsma, O., Nieuwland, R., Schwaeble, W., Gingras, A., Mantovani, A., Hack, E. & Roos, A. (2002), "Direct binding of C1q to apoptotic cells and cell blebs induces complement activation", *EUROPEAN JOURNAL OF IMMUNOLOGY*, vol. 32, no. 6, pp. 1726-1736.

Nielsen, S., Frokiaer, J., Marples, D., Kwon, T., Agre, P. & Knepper, M. (2002), "Aquaporins in the kidney: From molecules to medicine", *PHYSIOLOGICAL REVIEWS*, vol. 82, no. 1, pp. 205-244.

Obregon, C., Rothen-Rutishauser, B., Gitahi, S.K., Gehr, P. & Nicod, L.P. (2006), "Exovesicles from human activated dendritic cells fuse with resting dendritic cells, allowing them to present alloantigens", *AMERICAN JOURNAL OF PATHOLOGY*, vol. 169, no. 6, pp. 2127-2136.

Palanisamy, V., Sharma, S., Deshpande, A., Zhou, H., Gimzewski, J. & Wong, D.T. (2010), "Nanostructural and Transcriptomic Analyses of Human Saliva Derived Exosome", *PLOS ONE*, vol. 5, no. 1, e8577.

Pan, B.T., Teng, K., Wu, C., Adam, M. & Johnstone, R.M. (1985), "Electron microscopic evidence for externalization of the transferrin receptor in vesicular form in sheep reticulocytes.", *THE JOURNAL OF CELL BIOLOGY*, vol. 101, no. 3, pp. 942-948.

Pandey, A., Podtelejnikov, A., Blagoev, B., Bustelo, X., Mann, M. & Lodish, H. (2000), "Analysis of receptor signaling pathways by mass spectrometry: Identification of Vav-2 as a substrate of the epidermal and platelet-derived growth factor receptors", *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, vol. 97, no. 1, pp. 179-184.

Pascual, M., Steiger, G., Sadallah, S., Paccaud, J.P., Carpentier, J.L., James, R. & Schifferli, J.A. (1994), "Identification of membrane-bound CR1 (CD35) in human urine: evidence for its release by glomerular podocytes.", *THE JOURNAL OF EXPERIMENTAL MEDICINE*, vol. 179, no. 3, pp. 889-899.

Peche, H., Heslan, M., Usal, C., Amigorena, S. & Cuturi, M. (2003), "Presentation of donor major histocompatibility complex antigens by bone marrow dendritic cell-derived exosomes modulates allograft rejection", *TRANSPLANTATION*, vol. 76, no. 10, pp. 1503-1510.

Perez-Pujol, S., Marker, P.H. & Key, N.S. (2007), "Platelet microparticles are heterogeneous and highly dependent on the activation mechanism: Studies using a new digital flow cytometer", *CYTOMETRY PART A*, vol. 71A, no. 1, pp. 38-45.

Pilzer, D. & Fishelson, Z. (2005), "Mortalin/GRP75 promotes release of membrane vesicles from immune attacked cells and protection from complement-mediated lysis", *INTERNATIONAL IMMUNOLOGY*, vol. 17, no. 9, pp. 1239-1248.

Pilzer, D., Gasser, O., Moskovich, O., Schifferli, J. & Fishelson, Z. (2005), "Emission of membrane vesicles: roles in complement resistance, immunity and cancer", *SPRINGER SEMINARS IN IMMUNOPATHOLOGY*, vol. 27, no. 3, pp. 375-387.

Pisitkun, T., Shen, R. & Knepper, M. (2004), "Identification and proteomic profiling of exosomes in human urine", *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, vol. 101, no. 36, pp. 13368-13373.

Pisitkun, T., Johnstone, R. & Knepper, M.A. (2006), "Discovery of urinary biomarkers", *MOLECULAR \& CELLULAR PROTEOMICS*, vol. 5, no. 10, pp. 1760-1771.

Pizzirani, C., Ferrari, D., Chiozzi, P., Adinolfi, E., Sandona , D., Savaglio, E. & Di Virgilio, F. (2007), "Stimulation of P2 receptors causes release of IL-1β–loaded microvesicles from human dendritic cells", *BLOOD*, vol. 109, no. 9, pp. 3856-3864.

Pluskota, E., Woody, N.M., Szpak, D., Ballantyne, C.M., Soloviev, D.A., Simon, D.I. & Plow, E.F. (2008), "Expression, activation, and function of integrin αMβ2 (Mac-1) on neutrophil-derived microparticles", *BLOOD*, vol. 112, no. 6, pp. 2327-2335.

Polgar, J., Matuskova, J. & Wagner, D.D. (2005), "The P-selectin, tissue factor, coagulation triad", *JOURNAL OF THROMBOSIS AND HAEMOSTASIS*, vol. 3, no. 8, pp. 1590-1596.

Ponting, C., Hofmann, K. & Bork, P. (1999), "A latrophilin/CL-1-like GPS domain in polycystin-1", *CURRENT BIOLOGY*, vol. 9, no. 16, pp. R585-R588.

Porto-Carreiro, I., Fevrier, B., Paquet, S., Vilette, D. & Raposo, G. (2005), "Prions and exosomes: From PrPc trafficking to PrPsc propagation", *BLOOD CELLS MOLECULES AND DISEASES*, vol. 35, no. 2, pp. 143-148.

Poutsiaka, D., Schroder, E., Taylor, D., Levy, E. & Black, P. (1985), "Membrane-vesicles shed by murine melanoma-cells selectively inhibit the expression of Ia-antigen by macrophages", *JOURNAL OF IMMUNOLOGY*, vol. 134, no. 1, pp. 138-144.

Putz, U., Howitt, J., Lackovic, J., Foot, N., Kumar, S., Silke, J. & Tan, S. (2008), "Nedd4 Family-interacting Protein 1 (Ndfip1) Is Required for the Exosomal Secretion of Nedd4 Family Proteins", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 283, no. 47, pp. 32621-32627.

Qian, F., Boletta, A., Bhunia, A., Xu, H., Liu, L., Ahrabi, A., Watnick, T., Zhou, F. & Germino, G. (2002), "Cleavage of polycystin-1 requires the receptor for egg jelly domain and is disrupted by human autosomal-dominant polycystic kidney disease 1-associated mutations", *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, vol. 99, no. 26, pp. 16981-16986.

Quah, B. & O'Neill, H. (2005), "The immunogenicity of dendritic cell-derived exosomes", *BLOOD CELLS MOLECULES AND DISEASES*, vol. 35, no. 2, pp. 94-110.

Raposo, G., Nijman, H.W., Stoorvogel, W., Liejendekker, R., Harding, C.V., Melief, C.J. & Geuze, H.J. (1996), "B lymphocytes secrete antigen-presenting vesicles.", *THE JOURNAL OF EXPERIMENTAL MEDICINE*, vol. 183, no. 3, pp. 1161-1172.

Ridgway, H., Brennan, S., Loreth, R. & George, P. (1997), "Fibrinogen Kaiserslautern (gamma 380 Lys to Asn): a new glycosylated fibrinogen variant with delayed polymerization", *BRITISH JOURNAL OF HAEMATOLOGY*, vol. 99, no. 3, pp. 562-569.

Robert, S., Poncelet, P., Lacroix, R., Arnaud, L., Giraudo, L., Hauchard, A., Sampol, J. & Dignat-george, F. (2009), "Standardization of platelet-derived microparticle counting using calibrated beads and a Cytomics FC500 routine flow cytometer: a first step towards multicenter studies?", *JOURNAL OF THROMBOSIS AND HAEMOSTASIS*, vol. 7, no. 1, pp. 190-197.

Rood, I.M., Deegens, J.K.J., Merchant, M.L., Tamboer, W.P.M., Wilkey, D.W., Wetzels, J.F.M. & Klein, J.B. (2010), "Comparison of three methods for isolation of urinary microvesicles to identify biomarkers of nephrotic syndrome", *KIDNEY INTERNATIONAL*, vol. 78, no. 8, pp. 810-816.

Satta, N., Toti, F., Feugeas, O., Bohbot, A., Dachary-Prigent, J., Eschwege, V., Hedman, H. & Freyssinet, J. (1994), "Monocyte vesiculation is a possible mechanism for dissemination of membrane-associated procoagulant activities and adhesion molecules after stimulation by lipopolysaccharide", *THE JOURNAL OF IMMUNOLOGY*, vol. 153, no. 7, pp. 3245-3255.

Scholz, T., Temmler, U., Krause, S., Heptinstall, S. & Losche, W. (2002), "Transfer of tissue factor from platelets to monocytes: Role of platelet-derived microvesicles and CD62P", *THROMBOSIS AND HAEMOSTASIS*, vol. 88, no. 6, pp. 1033-1038.

Schorey, J.S. & Bhatnagar, S. (2008), "Exosome function: From tumor immunology to pathogen biology", *TRAFFIC*, vol. 9, no. 6, pp. 871-881.

Scolding, N., Morgan, B., Houston, W., Linington, C., Campbell, A. & Compston, D. (1989), "Vesicular removal by oligodendrocytes of membrane attack complexes formed by activated complement", *NATURE*, vol. 339, no. 6226, pp. 620-622.

Segura, E., Amigorena, S. & Thery, C. (2005), "Mature dendritic cells secrete exosomes with strong ability to induce antigen-specific effector immune responses", *BLOOD CELLS MOLECULES AND DISEASES*, vol. 35, no. 2, pp. 89-93.

Sharma, S., Rasool, H.I., Palanisamy, V., Mathisen, C., Schmidt, M., Wong, D.T. & Gimzewski, J.K. (2010), "Structural-Mechanical Characterization of Nanoparticle Exosomes in Human Saliva, Using Correlative AFM, FESEM, and Force Spectroscopy", *ACS NANO*, vol. 4, no. 4, pp. 1921-1926.

Sharon, N. & Lis, H. (1989), "Lectins as cell recognition molecules", *SCIENCE*, vol. 246, no. 4927, pp. 227-234.

Sharrow, S., Mathieson, B. & Singer, A. (1981), "Cell surface appearance of unexpected host MHC determinants on thymocytes from radiation bone marrow chimeras", *THE JOURNAL OF IMMUNOLOGY*, vol. 126, no. 4, pp. 1327-1335.

Shen, B., Wu, N., Yang, J. & Gould, S.J. (2011), "Protein Targeting to Exosomes/Microvesicles by Plasma Membrane Anchors", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 286, no. 16, pp. 14383-14395.

Simons, M. & Raposo, G. (2009), "Exosomes - vesicular carriers for intercellular communication", *CURRENT OPINION IN CELL BIOLOGY*, vol. 21, no. 4, pp. 575-581.

Sims, P., Faioni, E., Wiedmer, T. & Shattil, S. (1988), "Complement proteins C5B-9 cause release of membrane-vesicles from the platelet surface that are enriched in the membrane-receptor for coagulation Factor-Va and express prothrombinase activity", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 263, no. 34, pp. 18205-18212.

Slrnalley, D.M., Sheman, N.E., Nelson, K. & Theodorescu, D. (2008), "Isolation and identification of potential urinary microparticle biomarkers of bladder cancer", *JOURNAL OF PROTEOME RESEARCH*, vol. 7, no. 5, pp. 2088-2096.

Sokolova, V., Ludwig, A., Hornung, S., Rotan, O., Horn, P.A., Epple, M. & Glebel, B. (2011), "Characterisation of exosomes derived from human cells by nanoparticle tracking analysis and scanning electron microscopy", *COLLOIDS AND SURFACES B-BIOINTERFACES*, vol. 87, no. 1, pp. 146-150.

Sparbier, K., Wenzel, T. & Kostrzewa, M. (2006), "Exploring the binding profiles of ConA, boronic acid and WGA by MALDI-TOF/TOF MS and magnetic particles", *JOURNAL OF CHROMATOGRAPHY B*, vol. 840, no. 1, pp. 29-36.

Spek, C. (2004), "Tissue factor: from `just one of the coagulation factors' to a major player in physiology", *BLOOD COAGULATION \& FIBRINOLYSIS*, vol. 15, no. 1, pp. S3-S10.

Steen, H. (2004), "Flow cytometer for measurement of the light scattering of viral and other submicroscopic particles", *CYTOMETRY PART A*, vol. 57A, no. 2, pp. 94-99.

Stein, J. & Luzio, J. (1991), "Ectocytosis caused by sublytic autologous complement attack on human neutrophils - the sorting of endogenous plasma-membrane proteins and lipids into shed vesicles", *BIOCHEMICAL JOURNAL*, vol. 274, no. Part 2, pp. 381-386.

Subra, C., Laulagnier, K., Perret, B. & Record, M. (2007), "Exosome lipidomics unravels lipid sorting at the level of multivesicular bodies", *BIOCHIMIE*, vol. 89, no. 2, pp. 205-212.

Sullivan, R., Saez, F., Girouard, J. & Frenette, G. (2005), "Role of exosomes in sperm maturation during the transit along the male reproductive tract", *BLOOD CELLS MOLECULES AND DISEASES*, vol. 35, no. 1, pp. 1-10.

Taylor, D.D., Taylor, C.G., Jiang, C. & Black, P.H. (1988), "Characterization of plasma membrane shedding from murine melanoma cells", *INTERNATIONAL JOURNAL OF CANCER*, vol. 41, no. 4, pp. 629-635.

Theos, A.C., Truschel, S.T., Tenza, D., Hurbain, I., Harper, D.C., Berson, J.F., Thomas, P.C., Raposo, G. & Marks, M.S. (2006), "A Lumenal Domain-Dependent Pathway for Sorting to Intralumenal Vesicles of Multivesicular Endosomes Involved in Organelle Morphogenesis", *DEVELOPMENTAL CELL*, vol. 10, no. 3, pp. 343-354.

Thery, C., Boussac, M., Veron, P., Ricciardi-Castagnoli, P., Raposo, G., Garin, J. & Amigorena, S. (2001), "Proteomic analysis of dendritic cell-derived exosomes: A secreted subcellular compartment distinct from apoptotic vesicles", *JOURNAL OF IMMUNOLOGY*, vol. 166, no. 12, pp. 7309-7318.

Thery, C., Ostrowski, M. & Segura, E. (2009), "Membrane vesicles as conveyors of immune responses", *NATURE REVIEWS IMMUNOLOGY*, vol. 9, no. 8, pp. 581-593.

Thongboonkerd, V. & Malasit, P. (2005), "Renal and urinary proteomics: Current applications and challenges", *PROTEOMICS*, vol. 5, no. 4, pp. 1033-1042.

Tian, Y., Zhou, Y., Elliott, S., Aebersold, R. & Zhang, H. (2007), "Solid-phase extraction of N-linked glycopeptides", *NATURE PROTOCOLS*, vol. 2, no. 2, pp. 334-339.

Trajkovic, K., Hsu, C., Chiantia, S., Rajendran, L., Wenzel, D., Wieland, F., Schwille, P., Bruegger, B. & Simons, M. (2008), "Ceramide triggers budding of exosome vesicles into multivesicular Endosomes", *SCIENCE*, vol. 319, no. 5867, pp. 1244-1247.

Troy, F. (1992), "Polysialylation - from bacteria to brains", *GLYCOBIOLOGY*, vol. 2, no. 1, pp. 5-23.

Valadi, H., Ekstrom, K., Bossios, A., Sjostrand, M., Lee, J.J. & Lotvall, J.O. (2007), "Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells", *NATURE CELL BIOLOGY*, vol. 9, no. 6, pp. 654-659.

Valenti, G., Laera, A., Pace, G., Aceto, G., Lospalluti, M., Penza, R., Selvaggi, F., Chiozza, M. & Svelto, M. (2000), "Urinary aquaporin 2 and calciuria correlate with the severity of enuresis in children", *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY*, vol. 11, no. 10, pp. 1873-1881.

Van Niel, G., Mallegol, J., Bevilacqua, C., Candalh, C., Brugiere, S., Tomaskovic-Crook, E., Heath, J., Cerf-Bensussan, N. & Heyman, M. (2003), "Intestinal epithelial exosomes carry MHC class II/peptides able to inform the immune system in mice", *GUT*, vol. 52, no. 12, pp. 1690-1697.

Vella, L., Sharples, R., Lawson, V., Masters, C., Cappai, R. & Hill, A. (2007), "Packaging of prions into exosomes is associated with a novel pathway of PrP processing", *THE JOURNAL OF PATHOLOGY*, vol. 211, no. 5, pp. 582-590.

Vertegaal, A.C.O., Andersen, J.S., Ogg, S.C., Hay, R.T., Mann, M. & Lamond, A.I. (2006), "Distinct and overlapping sets of SUMO-1 and SUMO-2 target proteins revealed by quantitative proteomics", *MOLECULAR \& CELLULAR PROTEOMICS*, vol. 5, no. 12, pp. 2298-2310.

Vidal, M., Mangeat, P. & Hoekstra, D. (1997), "Aggregation reroutes molecules from a recycling to a vesicle-mediated secretion pathway during reticulocyte maturation", *JOURNAL OF CELL SCIENCE*, vol. 110, no. Part 16, pp. 1867-1877.

Walsh, C., Garneau-Tsodikova, S. & Gatto, G. (2005), "Protein posttranslational modifications: The chemistry of proteome diversifications", *ANGEWANDTE CHEMIE-INTERNATIONAL EDITION*, vol. 44, no. 45, pp. 7342-7372.

Wang, Y., Wu, S. & Hancock, W. (2006), "Approaches to the study of N-linked glycoproteins in human plasma using lectin affinity chromatography and nano-HPLC coupled to electrospray linear ion trap-Fourier transform mass spectrometry", *GLYCOBIOLOGY*, vol. 16, no. 6, pp. 514-523.

Welchman, R., Gordon, C. & Mayer, R. (2005), "Ubiquitin and ubiquitin-like proteins as multifunctional signals", *NATURE REVIEWS MOLECULAR CELL BIOLOGY*, vol. 6, no. 8, pp. 599-609.

Wiener, M. & vanHoek, A. (1996), "A lectin screening method for membrane glycoproteins: Application to the human CHIP28 water channel (AQP-1)", *ANALYTICAL BIOCHEMISTRY*, vol. 241, no. 2, pp. 267-268.

Wilson, H.L., Francis, S.E., Dower, S.K. & Crossman, D.C. (2004), "Secretion of Intracellular IL-1 Receptor Antagonist (Type 1) Is Dependent on P2X7 Receptor Activation", *The Journal of Immunology*, vol. 173, no. 2, pp. 1202-1208.

Wolfers, J., Lozier, A., Raposo, G., Regnault, A., Thery, C., Masurier, C., Flament, C., Pouzieux, S., Faure, F., Tursz, T., Angevin, E., Amigorena, S. & Zitvogel, L. (2001), "Tumor-derived exosomes are a source of shared tumor rejection antigens for CTL cross-priming", *NATURE MEDICINE*, vol. 7, no. 3, pp. 297-303.

Wubbolts, R., Leckie, R.S., Veenhuizen, P.T.M., Schwarzmann, G., Mobius, W., Hoernschemeyer, J., Slot, J., Geuze, H.J. & Stoorvogel, W. (2003), "Proteomic and Biochemical Analyses of Human B Cell-derived Exosomes", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 278, no. 13, pp. 10963-10972.

Yamamoto, K., Tsuji, T. & Osawa, T. (1998), "Analysis of Asparagine-Linked Oligosaccharides by Sequential Lectin-Affinity Chromatography", *Methods In Molecular Biology*, vol. 76, pp. 35-51.

Yang, Z. & Hancock, W. (2004), "Approach to the comprehensive analysis of glycoproteins isolated from human serum using a multi-lectin affinity column", *JOURNAL OF CHROMATOGRAPHY A*, vol. 1053, no. 1-2, pp. 79-88.

Youings, A., Chang, S., Dwek, R. & Scragg, I. (1996), "Site-specific glycosylation of human immunoglobulin G is altered in four rheumatoid arthritis patients", *BIOCHEMICAL JOURNAL*, vol. 314, no. Part 2, pp. 621-630.

Yu, X., Harris, S. & Levine, A. (2006), "The regulation of exosome secretion: a novel function of the p53 protein", *CANCER RESEARCH*, vol. 66, no. 9, pp. 4795-4801.

Zeidler, R.B. & Kim, H.D. (1982), "Pig reticulocytes. IV. In vitro maturation of naturally occurring reticulocytes with permeability loss to glucose", *JOURNAL OF CELLULAR PHYSIOLOGY*, vol. 112, no. 3, pp. 360-366.

Zhang, F., Sun, S., Feng, D., Zhao, W. & Sui, S. (2009), "A Novel Strategy for the Invasive Toxin: Hijacking Exosome-Mediated Intercellular Trafficking", *TRAFFIC*, vol. 10, no. 4, pp. 411-424.

Zhang, H., Li, X., Martin, D. & Aebersold, R. (2003), "Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry", *NATURE BIOTECHNOLOGY*, vol. 21, no. 6, pp. 660-666.

Zhang, H., Liu, C., Su, K., Yu, S., Zhang, L., Zhang, S., Wang, J., Cao, X., Grizzle, W. & Kimberly, R. (2006), "A membrane form of TNF-alpha presented by exosomes delays T cell activation-induced cell death", *JOURNAL OF IMMUNOLOGY*, vol. 176, no. 12, pp. 7385-7393.

Zhang, H., Kim, H., Liu, C., Yu, S., Wang, J., Grizzle, W.E., Kimberly, R.P. & Barnes, S. (2007), "Curcumin reverses breast tumor exosomes mediated immune suppression of NK cell tumor cytotoxicity", *BIOCHIMICA ET BIOPHYSICA ACTA (BBA) - MOLECULAR CELL RESEARCH*, vol. 1773, no. 7, pp. 1116-1123.

Zhou, H., Pisitkun, T., Aponte, A., Yuen, P.S.T., Hoffert, J.D., Yasuda, H., Hu, X., Chawla, L., Shen, R., Knepper, M.A. & Star, R.A. (2006 a), "Exosomal Fetuin-A identified by proteomics: A novel urinary biomarker for detecting acute kidney injury", *KIDNEY INTERNATIONAL*, vol. 70, no. 10, pp. 1847-1857.

Zhou, H., Yuen, P., Pisitkun, T., Gonzales, P., Yasuda, H., Dear, J., Gross, P., Knepper, M. & Star, R. (2006 b), "Collection, storage, preservation, and normalization of human urinary exosomes for biomarker discovery", *KIDNEY INTERNATIONAL*, vol. 69, no. 8, pp. 1471-1476.

Zoeller, M. (2009), "Tetraspanins: push and pull in suppressing and promoting metastasis", *NATURE REVIEWS CANCER*, vol. 9, no. 1, pp. 40-55.

AIMS OF THE STUDY

The aims of the study were following:

1. To develop alternative and/or novel methods for isolation and purification of urinary membrane vesicles including the search for novel affinity-ligand-based methods for urinary membrane vesicles isolation.

2. To characterise urinary membrane vesicles and their content using proteomic ‘gel-based’ and ‘off-gel’ proteomic analysis.

3. To profile glycoproteins of urinary membrane vesicles. To establish the surface glycan signature of urinary membrane vesicles followed by identification of the constituent glycoproteins of urinary membrane vesicles.

4. To profile the other post-translational modifications like palmitoylation and ubiquitination of protein constituents of urinary membrane vesicles and identification of these proteins.

CHAPTER 2

BIOCHEMICAL AND PHYSICAL CHARACTERISATION OF URINARY CHAPS-RESISTANT MEMBRANE VESICLES

2.1 Introduction

The discovery of exosome vesicles in urine (Pisitkun, Shen & Knepper, 2004) has rapidly opened new possibilities for the mechanistic understanding of biological processes and, importantly, has served as a source for novel biomarkers (Simpson *et al.*, 2009). Consequently, exhaustive proteomic profiling of urinary exosomes has identified more than 1100 gene products, including 177 disease-related proteins derived from all nephron segments (Gonzales *et al.*, 2009) and from the urogenital tract (Mitchell *et al.*, 2009; Welton *et al.*, 2010). The identification from urine of distinct exosomal transcription factors (Zhou *et al.*, 2008) and nucleic acids encoding proteins native to all nephron segments (Miranda *et al.*, 2010) is groundbreaking, and highlights the need to precisely understand their biology which plausibly reflects new aspects of disease pathways.

The aim of this study was to optimise the currently-available techniques and remove the abundant confounding urinary proteins which seriously interfere with the exosomal vesicle recovery, thus influencing the final yield and subsequent analytical power. The treatment of the exosomal pellet obtained by the serial centrifugation protocol with dithiothreitol (DTT) has previously been proposed as a solution to reduce such interference (Gonzales *et al.*, 2009). After treatment with DTT, exosomal markers, CD9, Alix and TSG101 pelleted down in the high-speed pellet (P200,000g) but not in the low speed pellet (18,000g). Therefore, the yield of exosomes increased, however, Tamm-Horsfall protein, which is the most abundant protein in urine and a contaminant in urinary exosome isolation, pelleted down at high speed. This might interfere in obtaining pure exosomal pellet (200,000g) and complicate further proteomic analysis.

Furthermore, previous proteomic profiling studies have also identified a number of receptor proteins whose three-dimensional folding is stabilised by disulfide bridges. Accordingly,

sortilin-related receptor, for example, has 33 predicted disulfide bridges and megalin has 159 predicted disulfide bridges (Westergaard *et al.,* 2004; Bajari *et al.,* 2005) fixing their respective molecular structures. Any study attempting to evaluate functions of exosomes should optimally follow an isolation protocol which preserves the correct folding and therefore, the functionality of the respective proteins. DTT is a strong reducing agent, however, and the exosomal proteins may accordingly be reduced and unable to refold properly upon reoxidation, hampering relevant functional studies leading to loss of biomarker promise

Here, we hypothesized that interference of soluble proteins in the exosomal isolation process occurs due to aggregation, non-specific interactions as well as their gelling properties (as e.g. for Tamm-Horsfall glycoprotein {THP}). Any chemical agent which solubilises these aggregates may also reduce contamination of the ultracentrifugation pellet by these proteins. For these benefits, we have used 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonic (CHAPS) as mild detergent which is known to solubilise THP (Kobayashi & Fukuoka, 2001) to largely exclude this interference. Mild non-ionic or zwitterionic detergents can solubilise the polymers of THP which might lead to pelletting down of THP at low speed (18,000g). This would be expected to remove contamination of high speed pellet (200,000g) with THP. We then compared results after CHAPS treatment with those after the DTT treatment.

Furthermore, we also compared the preservation of two key protease activities, including dipeptidyl peptidase IV (DPP IV) and nephrilysin (NEP), both previously shown to be associated with urinary exosomes (Pisitkun, Shen & Knepper, 2004; Gonzales *et al.,* 2009). These two enzymes are stabilised by 5 and 6 disulfide bonds, respectively (Erdos & Skidgel, 1989; Busek, Malik & Sedo, 2004). Furthermore, we reveal for the first time the overlapping proteome subsets following CHAPS and DTT treatments (SN200,000g). Proteomic profiling

of the SN200,000g would shed light on the types of proteins whose interference will be reduced in the P200,000g. The proteomic profiling revealed that both CHAPS and DTT methods result in closely-related protein profiles, fully validating our method while CHAPS was found to be superior in maintaining functional protein integrity. Our method should greatly improve the exosome functional yield and, thus, possibilities to fully exploit the biomarker potential of exosomes.

2.2 Materials and Methods

2.2.1 Urine collection

Urine samples were collected from ten non-smoking healthy laboratory volunteers (5 female and 5 male) whose ages ranged from between 20 to 40 years. There was no history of renal dysfunction in any of the subjects or drug administration during sample collection. The first morning urine was processed within 3 hours of collection. At collection, the first 50 mls were discarded and the remaining void urine was collected for analysis. Each urine specimen was subsequently tested by the Combur 10 Test®D dipstick (Roche Diagnostics; Mannheim, Germany).

2.2.2 Vesicle purification

All the chemicals were purchased from SIGMA, St. Louis unless otherwise specified. A schematic representation of the methodology used to isolate nano vesicles is shown in Figure 2.1. In summary, pooled urine samples were initially centrifuged at a relative centrifugal force (RCF) of 1,000g for 20 minutes. The resultant supernatant (SN) was retained and split into aliquots of 500mL. Upon thawing, these samples presented with cryo-precipitate, which was completely dissolved by vortexing the sample at room temperature (RT) until complete the solution became clear. Next, 500mL aliquots of SN1 were dialysed at 4ºC against deionised water (3 changes of 10L each over 24hours) and the volume was subsequently reduced up to 20mL using vacuum concentration (miVac, GeneVac, Suffolk, Ipswich, UK). The concentrated SN1 was centrifuged (Avanti®J-26 XP centrifuge, Beckman Coulter, Fullerton, CA) at 18,000g for 30 minutes at RT in a fixed angle rotor (Beckman JA-20, Fullerton, CA). The 18,000g SN (SN18) fraction was subjected to an ultracentrifugation step using the OptimaTM L-90 K preparative ultracentrifuge (Beckman Coulter) at 200,000g for 2 hours at RT in a fixed-angle rotor (Beckman 70Ti, Beckman). The pellet (P200) was

resuspended in 3 mL of deionized water. Aliquots (2mg of total protein) of resolubilised crude P18 and P200 were treated with either 1% (w/v) of 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonic acid (CHAPS) overnight (ON) with end-over-end agitation at 4°C (5mL final volume) or with 200mg/mL of dithiothreitol (DTT), independently of one another, for 30 minutes at 37°C in accordance with the recommendations of Fernández-Llama *et al.,* (Fernandez-Llama *et al.,* 2010) with vortexing every five minutes. Here, a final volume of 5mL was also used. These two solutions (i.e. P18 and P200, CHAPS and DTT-treated pellet) were subjected to a second series of centrifugations at 18,000g followed by an ultracentrifugation at 200,000g as shown in Figure 1. After the first centrifugation at low-speed (18,000g), the SNs from same treatment were pooled together (10mL final volume) and centrifuged at 200,000g. For method 2, 0.5mg of crude protein obtained from P18 and P200, respectively, was treated with CHAPS as in method 1 without pooling the SN after the first low-speed pellet (18,000g). Instead, they were subjected individually to an ultracentrifugation step at 200,000g.

2.2.3 Protein quantification, SDS-PAGE and Western Blotting

The protein concentrations in all urine fractions were measured by Coomassie Protein Assay (M., Bradford, 1976). SDS-PAGE was carried out according to the recommendations of Laemmli (Laemmli, 1970). Gels were either stained by homemade colloidal Coomassie blue staining (Candiano *et al.,* 2004) or transferred to a nitrocellulose membrane (Towbin, Staehelin & Gordon, 1979). Membranes were blocked overnight (ON) at room temperature (RT) with the Odyssey blocking buffer solution (LI-COR Biosciences, Lincoln, NE). All incubation steps with antibodies were performed in a 1:1 (v/v) mixture of Odyssey blocking buffer and PBS with 0.1% (v/v) Tween 20 (PBST). Antibodies were diluted in accordance with supplier's guidelines for 2 hours at room temperature. Rabbit anti-CD63 (H193) sc15363 (Santa Cruz, Santa Cruz, CA), dilution 1:500 (v/v); rabbit anti-TSG101 (Sigma)

0.5µg/mL dilution; rabbit anti-MFG-E8/lactadherin (H60); sc-33545v (Santa Cruz) 1:500 (v/v) dilution; rabbit anti-nephrin (in house) 1:500 (v/v). Secondary antibody goat anti-rabbit IRDye-800 (LI-COR Biosciences) was diluted 1:5000 (v/v) in PBS tween 0.1% (v/v) plus 0.01% (w/v) SDS for 2 hours at room temperature. Western blot images were acquired by Odyssey infrared laser scanner (LI-COR Biosciences, Lincoln, NE)

2.2.4 Negative Transmission Electron Microscopy

Fifty µg of vesicle preparations were fixed with 1% (v/v) glutaraldehyde (Sigma Aldrich) in water. Fixed vesicle preparations were spotted onto a Formvar/Carbon 300 mesh grid (Agar Scientific, Stansted, UK) and dried at RT. The grids were washed twice in 0.1M PBS and incubated in 1% (w/v) OsO_4 in 0.1M PBS for 30 min on ice. After five 5 min washes (3 with PBS and 2 with water), exosomes were stained with 5% (w/v) uranyl acetate in water for 10 min (Mathias *et al.*, 2009). After staining, vesicle populations were monitored by JEM-2100 transmission electron microscopy (Jeol Ltd, Tokyo, Japan).

2.2.5 LC-MS/MS analysis

100µg of protein were reduced by 10mM tris (2-carboxyethyl)phosphine (TCEP) in 100mM Tris–HCl, pH 8.8 , 8M Urea, 0.1mM EDTA and 1% (w/v) sodium deoxycholate (DOC) for 1.5 h at room temperature (RT) in the dark. Alkylation was carried out by 20mM iodoacetamide (IAA) in 100mM Tris–HCl, pH 8.8, 0.1mM EDTA for 1.5 h at RT in the dark. Excess of IAA was quenched by 20mM N-acetyl cysteine in 100mM Tris–HCl, pH 8.8, 0.1mM EDTA for 0.5 h at RT in the dark. Reduced and alkylated samples were concentrated/delipidated by chloroform/methanol precipitation according to Wessel and Fugge (Wessel & Flugge, 1984) before trypsin digestion. Briefly, to 150µl of sample solution were added first 400µl of 100% (v/v) MeOH, and then 200µl of chloroform and after vortexing well, 400µl of mQ water was added and then centrifuged at 5,000g for 5 min at RT. The upper layer was removed and 600µl of 100% (v/v) MeOH was added to the tube. Finally,

protein was recovered by centrifugation at 15,000g for 30 min and the pellet dried by speed vac. Digestion was carried out using trypsin (modified sequencing grade; Promega, Madison, WI) at 37°C for 16 h in 50mM Tris-HCL, pH 8.0 and 1% (w/v) DOC which is compatible with trypsin activity (Lin *et al.*, 2008). The digested solutions were acidified with 1% (v/v) formic acid (up to around pH 2) and centrifuged at 15,000g for 20 min. The supernatant was collected and the pellet were washed for with 1% (v/v) formic acid and sonicated for 5 minutes and centrifuged at 15,000g for 20 min. This was repeated twice (Lin *et al.*, 2010). All the SNs were pooled together and were then desalted and concentrated by using reverse phase cartridges (Sep-Pak C18) according to the manufacturer's instructions (Waters, Mississauga, ON). Briefly, the columns were conditioned and equilibrated with 1mL 100% methanol (v/v) and 1mL 80% (v/v) acetonitrile (ACN), respectively, and successively washed with 4 mL of 0.1% (v/v) formic acid. The sample solution was then loaded onto the column, and the columns were washed with 6 mL of 0.1% (v/v) formic acid. The tryptic digestions were eluted with 1 mL 80% (v/v) ACN, 0.1% (v/v) formic acid and lyophilized (Thermo). Lyophilized samples were rehydrated with 0.1% (v/v) formic acid to give 1μg/μL concentration of the equivalent starting protein amount. Nanoflow electrospray ionization tandem mass spectrometric analysis of peptide samples was carried out using LTQ-Orbitrap Velos (Thermo Scientific, Bremen, Germany) interfaced with the Agilent 1200 Series nanoflow LC system. The chromatographic capillary columns were used with a flow rate of 300nL per min. The peptides were eluted using a linear gradient of 7-30% (v/v in water) acetonitrile over 50 min. Mass spectrometry analysis was carried out in a data-dependent manner with full scans acquired using the Orbitrap mass analyzer at a mass resolution of 60,000 at 400 m/z. For each cycle, the twenty most intense precursor ions from a survey scan were selected for MS/MS and detected at a mass resolution of 15,000 at m/z 400. The fragmentation was carried out using higher-energy collision dissociation as the activation

method, with 40% normalized collision energy. The ions selected for fragmentation were excluded for 30 seconds. The automatic gain control for full FT MS was set to 1 million ions and for FT MS/MS was set to 0.1 million ions with a maximum time of accumulation of 750 and 100ms, respectively.

2.2.6 Data Analysis

The mass spectrometry data was processed using Proteome Discoverer (Version 1.2.0.208) software (Thermo Fisher Scientific) and searched using Mascot. The search parameters used were: oxidation of methionine, deamidation at N and Q, protein N-terminal acetylation, N-pyroglutamate for N-terminal Q/N and carbamidomethylation of cysteine residues as variable modifications. A maximum of one missed cleavage was allowed for tryptic peptides with a minimum length of 7 amino acids. The peptide and protein data were extracted using high peptide confidence and top one peptide rank filters. False discovery rate (FDR) was calculated by enabling the peptide sequence analysis using the decoy database. A mass error window of 20 ppm and 0.1Da were allowed for MS and MS/MS, respectively. FDR was kept at 1% to be used as a 'cut-off' value for reporting identified peptides. In addition to the target-decoy approach, the quality of proteins identified was validated by manually checking the spectra of those proteins with at least one unique peptide.

2.2.7 Determination of Dipeptidyl peptidase (DPP IV) and Nephrilysin (NEP) catalytic activity

DPP IV activity was measured in triplicate by using HGly- pro-β-naphthylamide as substrate, following the method of Liu and Hansen (Liu & Hansen, 1995). The NEP assay was carried out by incubating samples with N-dansyl-D-Ala-Gly-p-nitro-Phe-Gly ([D]AG (pN)PG, a dansyl derivative), following the protocol of Florentin *et al.*, (Florentin, Sassi & Roques, 1984).

2.2.8 THP purification

THP was purified from resolubilised P1 (Musante *et al.*, 2012) by filtration through a diatomaceous earth filter (DEF) (Serafinicessi *et al.*, 1989). In summary, PBS containing NaCl (0.2M) and $CaCl_2$ (1mM) was added to P1 before being transferred to a diatomaceous earth layer (5mL) placed on a Büchner funnel. After washing with 500mLs of 0.02mM phosphate buffer (pH 7.5) containing 0.2M NaCl and 1mM $CaCl_2$, the diatomaceous earth was retained and transferred into a 50mL centrifugation tube. This was subsequently washed twice with 50mL of 1mM ethylenediaminetetraacetic acid (EDTA) in water. This facilitated the extraction of THP entrapped within the diatomaceous earth. Non-bound and bound fractions were dialysed three times against deionised water (10L) independently of one another, and the volumes reduced to 5mL by vacuum drying (MiVac, Mason Technology, Suffolk, Ipswich, UK).

2.2.9 Purification of monomeric HSA

High-performance liquid chromatography (HPLC) assays were performed on an Agilent 1200 system (Agilent technologies, Palo Alto, California, USA), and all water selected for use was of HPLC grade (18.2 MΩ) and purified using a TKA-Gen-Pure purification system (Mason Technologies, Suffolk, Ipswich, UK). The species composition of human albumin was determined by SEC. Separations were performed using a Bio-Sil-SEC 400-5 stationary phase (Bio-Rad, Richmond, CA; 300x7.8mm), with 20mM PBS containing 0.15M NaCl (pH 7.5) selected as a mobile phase. For enhanced resolution of monomeric and dimeric fractions of HSA, three stationary phases were inter-connected with steel fittings. All samples were assayed at a flow rate of 1.0mL/minute using UV detection at 280nm, with 100µL sample volumes (1mg of protein) selected for injection. Species composition was determined through the use of a gel filtration standard consisting of thyroglobulin (670kDa), immunoglobulin G (158kDa), ovalbumin (44kDa), myoglobin (17kDa) and vitamin B12 (1.35kDa) (Bio-Rad).

Monomeric fractions of HSA for use in FLISA and SPR binding analysis (*vide infra*) were retained manually and screened by SDS-PAGE.

2.2.10 Fluorophore-Linked Immuno Sorbent Assay (FLISA)

As a preliminary method to ascertain whether HSA was binding to ultracentrifugation pellets prior to more rigorous SPR analysis, the P18+200CP200 vesicle fraction (100µg/mL) was passively immobilised in triplicate on a 96-well plate (Nunc Maxisorp, Nunc, Roskilde, Denmark) in 100µL volumes and incubated overnight at 4°C. Next, plates were washed five times with PBS (pH 7.4) and all wells selected for analysis were subsequently blocked with 400µL of Odyssey blocking solution diluted 1:1 in PBS (pH 7.4) for 1 hour at room temperature. After washing with PBS as before, monomeric HSA (100µl of a 20µg/mL stock solution) was added and wells were incubated for 1 hour at room temperature. Wells were subsequently washed five times with PBS containing 0.1% (v/v) tween (PBST). For detection, monoclonal mouse anti-HSA clone 6501 (100µL of a 10µg/mL stock solution), diluted in Odyssey blocking solution (1:1) was added before incubating for 2 hours at room temperature. Wells selected for analysis were washed again in PBST (5x). Next, 100µL of anti-mouse-IRdye 800 secondary antibody, diluted 1:5000 in Odyssey blocking solution, was added and allowed to incubate for 1 hour at room temperature. Non-bound antibody was removed by additional washes (5x) in PBST. Binding was determined through the use of the Odyssey infrared imaging system (Li-COR Biosciences, Lincoln, NE). Albumin interaction was evaluated at a physiological pH (7.4) and at a pH more representative of that seen in urine (pH 6.0). As a negative control, nano vesicle populations were immobilised as before, and subsequently probed with primary and secondary antibodies in the absence of albumin. The signal obtained here was then subtracted from that seen for analytical samples at pH 6.0 and pH 7.4.

2.2.11 Immunoprecipitation of membrane vesicles using anti-albumin antibody

Anti-albumin antibody was immobilised on Dynabeads (Dynal, M270 amine) according to manufacturer's instructions using 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide) and N-hydroxysulfosuccinimide (EDC/SNHS) chemistry (Invitrogen, Carlsbad, CA). 100µg of antibodies were coupled to 150µL of amine-functionalised Dynabeads (M-270 amine, Dynal, Invitrogen) using EDC-Sulfo-NHS. Briefly, antibodies were diluted to a final concentration of 1mg/mL antibody in MES buffer (0.1M MES + .9% (w/v) NaCl pH 6.0). EDC/SNHS were prepared in MES buffer at concentration of 10/15 mg/mL. EDC/SNHS solution (100µL/1mg protein) was added to the antibody and incubated at RT for 15 min. Following antibody activation by EDC/SNHS, the antibodies were added to 150µL Dynal M270-amine beads ($2x10^9$ beads/mL) and mixture was incubated at 4°C overnight while rotating. Next day, the beads were washed with PBS to remove non-bound antibodies and incubated with 100mM Tris, pH 7.2, for 15 min to quench any remaining active groups. The anti-albumin beads were incubated with 1mg of P200CHAPSP200 in binding buffer (phosphate buffer pH 7.2). The next day non-bound fraction was collected and beads were washed with binding buffer three times. The bound fraction was eluted from the beads by elution with 100mM glycine buffer (pH 2.3) for 1 hour while rotating at RT.

2.2.12 Biacore-based analysis of interactions

All surface-plasmon resonance (SPR) assays were performed on a Biacore 3000TM instrument using a CM5 carboxymethylated dextran sensor chip (GE Healthcare) in accordance with the recommendations of Reeves *et al.,* (Reeves *et al.,* 2010). In summary, HEPES-buffered saline (HBS: 10mM 4-2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 3mM EDTA, 150mM NaCl and 0.005% (v/v) Tween 20, pH 7.4 or 6.3), selected as a running buffer, was filtered through a 0.2µM filter (Supelco Inc., Bellefonte, PA) and de-gassed overnight in a sintered glass funnel apparatus (AGB Scientific Apparatus Ltd.,

Belfast, UK) before use. Pre-concentration analysis in 10mM sodium acetate was used to determine the optimal pH for the capture of 100µg/mL of vesicle preparation on flow cells (FC) 2 (P18+170DP170) and 4 (P18+170CP170), respectively. Subsequent immobilisation was permitted by surface activation using standard EDC/NHS coupling. Here, 70µL of EDC (400mM) was mixed with 70µL of NHS (100mM) and 80µL was injected over the sensor surface at a flow rate of 10µL/min. A 100µg/mL solution of the relevant ligand was then captured on the activated surface for 20 minutes. Any un-reacted succinimide esters were subsequently deactivated by capping for 7 minutes with 1M ethanolamine (pH 8.5). Flow cells 1 and 3 were used for online reference subtraction, and each five-minute injection (50µL of analyte at a flow-rate of 10µL/minute) was immediately followed by regeneration using conditions optimised through regeneration scouting (see below). All analyses were performed in triplicate on three consecutive days (interday), and binding responses were monitored as a function of analyte (e.g. albumin) concentration. A negative control of HBS buffer was included for each assay. The response units (RU) value obtained here was then subtracted from that obtained for each analyte-ligand biomolecular interaction, thereby providing an accurate representation of the true binding event. Biaevaluation software (GE Heathcare, Buckinghamshire, UK) was used for all data analysis.

2.3 Results

2.3.1 Vesicle purification

A schematic representation of the workflow used to isolated urine exosomal vesicles and abbreviations used for each product step is shown in Figure 2.1. For the analysis, pooled urine samples were initially centrifuged at a RCF of 1,000g for 20 min. The resultant SN was dialysed against deionized water water and sample volume reduced to 1:20 by vacuum concentration. The concentrated SN was centrifuged at 18,000g and the resulting supernatant at 200,000g. Both pellets (P18 and P200) were resuspended in a minimal volume of 200mg/mL DTT or 1% (w/v) CHAPS and re-centrifuged again at 18,000g. The two SNs were combined and further centrifuged at 200,000g. Alternatively, CHAPS 18,000g SNs were not pooled but centrifuged individually at 200,000g to check THP sedimentation behaviour following CHAPS treatment.

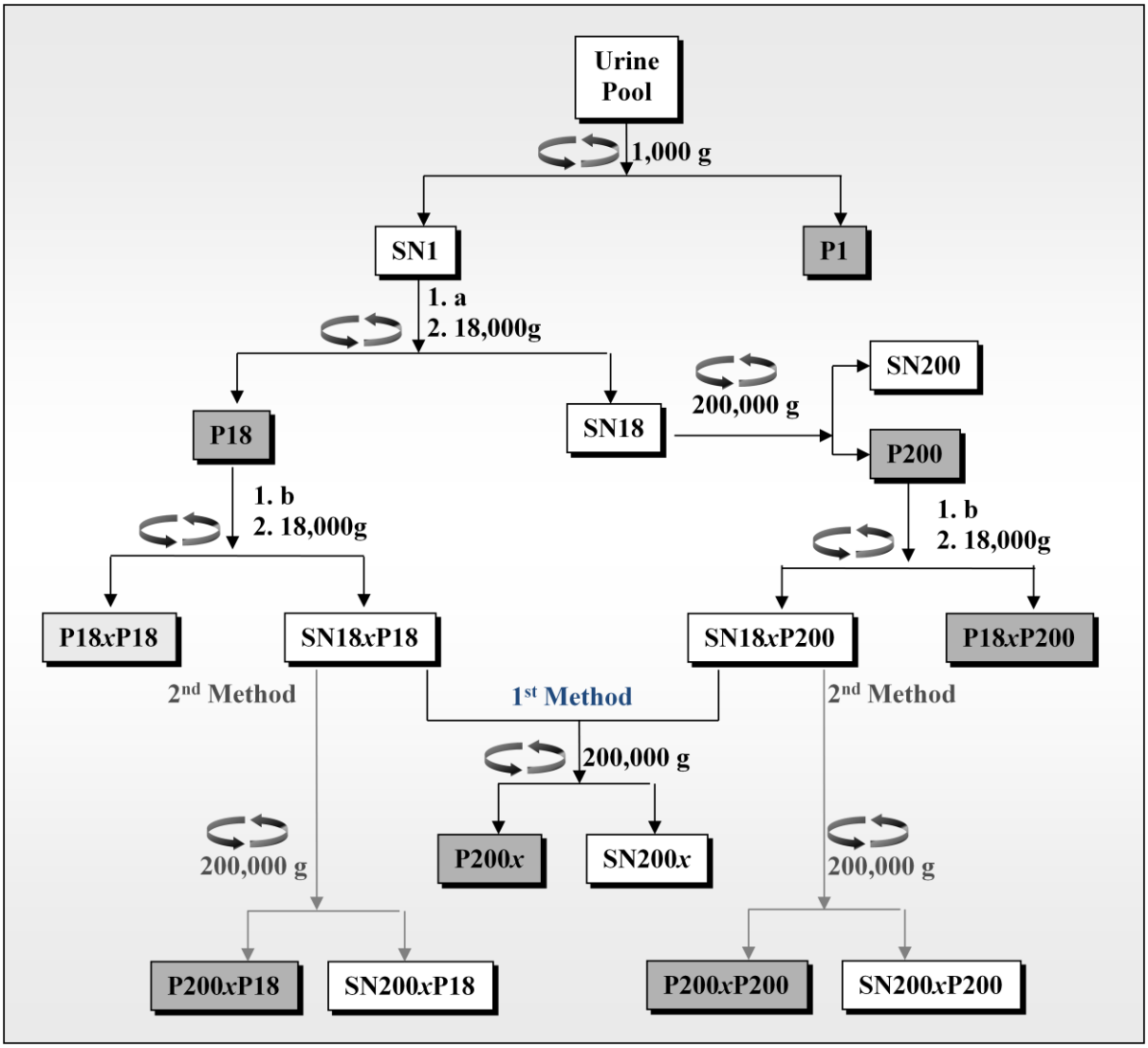

Figure 2.1: Urine vesicle enrichment workflow. **a.** Dialysis followed by vacuum concentration (approx.1:25); **b.** either 1% (w/v) CHAPS, ON at 4ºC or 200mg/mL DTT, 30 minutes at 37ºC. The type of treatment is reflected in the sample name: the x is replaced by C in case of CHAPS treatment and by D in the case of DTT. SAMPLE NOMENCATURE: P: Pellet; SN: Supernatant; x: Detergent used for treatment b C: CHAPS, D: DTT. Generally, P/SN are followed by the shortened centrifugational speed (e.g. P1 means pellet of 1,000g; 18 for 18,000g and 200 for 200,000g), after detergent treatment C/D will be used as prefix of the fraction which underwent treatment. If double centrifugation has been performed, the sample name will reflect the chronological order of events with the last step first.

2.3.2 Transmission electron Microscopy (TEM)

Transmission electron microscopy was used to verify the presence and to evaluate qualitative features of the vesicles present in the crude preparation P18 and P200 (for abbreviations, see Fig 2.1) and in CHAPS/DTT-treated fractions (Figure 2.2), as well as to assess the integrity of these vesicles following the respective treatments.

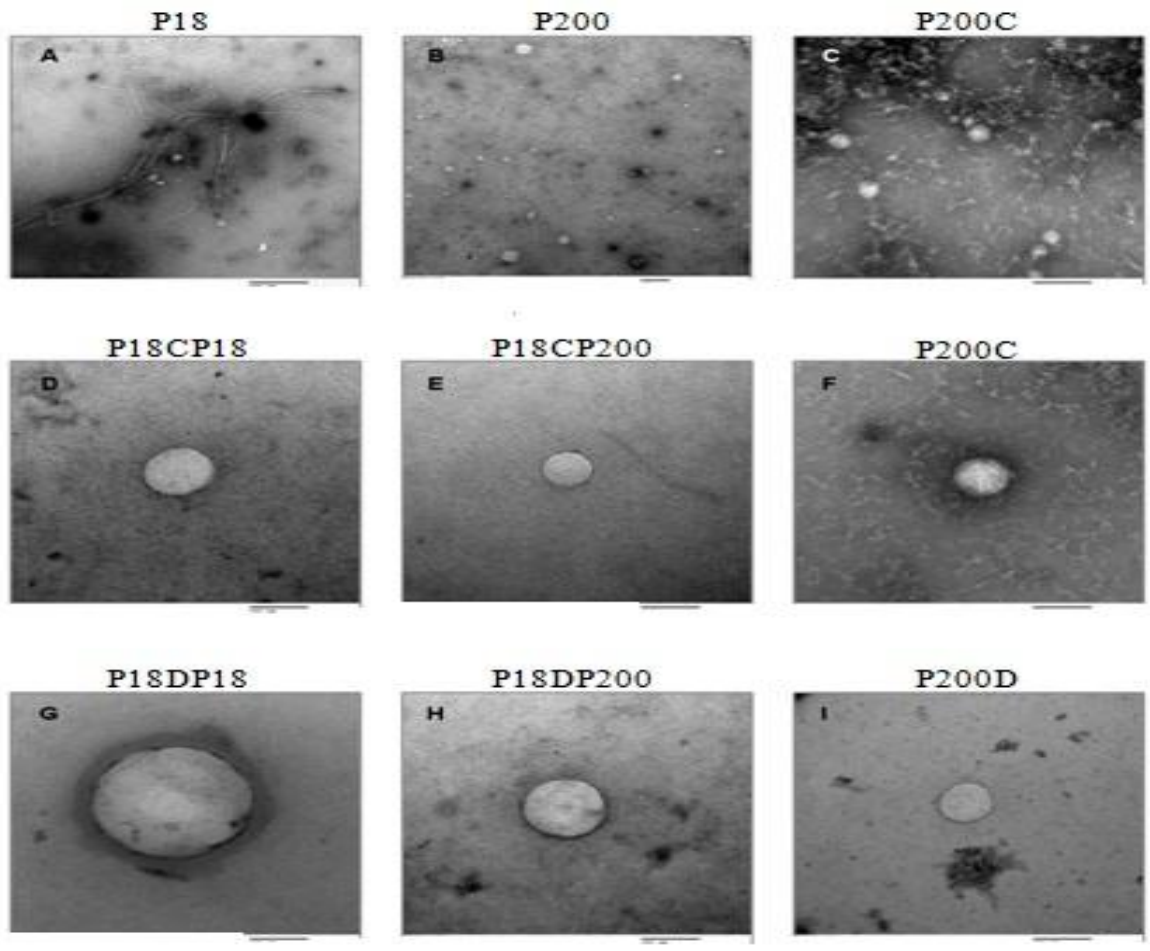

Figure 2.2: TEM analysis Transmission electron micrographs of P18 (Panel A) and P200 (Panel B) at 10,000x and 5,000x magnifications, respectively. High-magnification (50,000x) of CHAPS- (Panels C-F) and DTT-treated (Panels G-I) vesicle preparations are represented. The bar below the pictures is 500nm for panels A and B and 100nm for panels, C-I.

It can be seen in Fig. 2.2A that, the P18 preparation contained a heterogeneous population of vesicles ranging in size from 50-100nm in diameter. Furthermore, long polymeric filaments closely resembling THP polymers were detected along with associated vesicles (Fernandez-Llama *et al.*, 2010), despite the fact that dialysis should remove factors favouring the tendency of THP to aggregate (McQueen & Engel, 1966). Crude P200 (Fig. 2.2, Panel B) showed a heterogeneous population of vesicles whose size distribution (10-300nm) and shape are consistent with those of exosomes and 'exosome-like' vesicles previously described in urine (Pisitkun, Shen & Knepper, 2004; Hogan *et al.*, 2009). TEM images revealed that the preparations obtained after CHAPS or DTT treatments in method 1 yielded intact vesicles ranging from approximately 20 to 200nm, dimensions which are characteristic for urinary exosomes (Pisitkun, Shen & Knepper, 2004), 'exosome-like' (Hogan *et al.*, 2009) and apically-shed vesicles (Hara *et al.*, 2010). High magnification images (50,000x) after CHAPS (Fig. 2.2, panels C, D, E and F) and DTT treatments (Fig. 2.2, panels G, H and I) showed closely related morphologies and sizes. Another interesting observation was the detection of vesicles sized between 5 and 10 nm enriched in the P200C preparation, further highlighting the wide complexity, size distribution and heterogeneity of these fractions (Fig. 2.2, panel C).

2.3.3 SDS-PAGE analysis

Figure 3 shows colloidal Coomassie-stained gels of crude fractions (Fig. 2.3, panel A), DTT fractions and CHAPS fractions obtained in method 1 (Fig. 2.3, Panel B) and CHAPS fractions obtained in method 2 (Fig. 2.3, panel C).

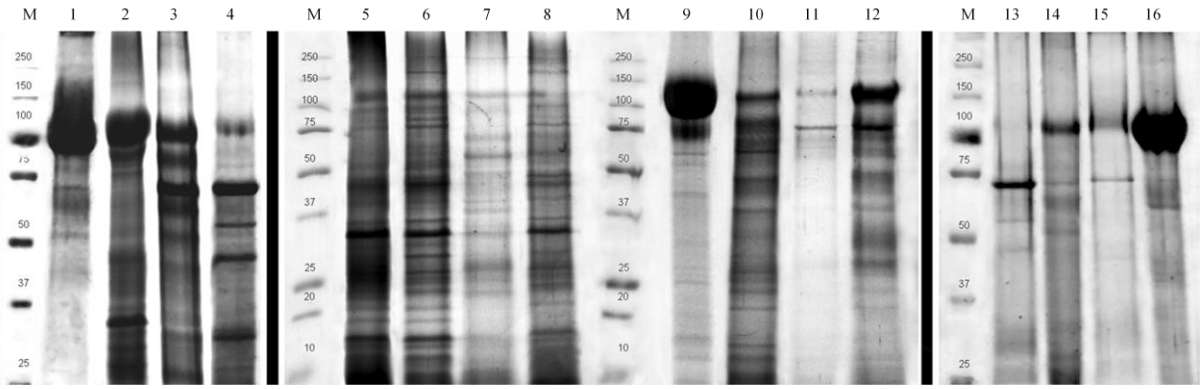

Figure 2.3: SDS-PAGE three panels are separated from each other by thick black lines. From left to right: **Panel A:** Gel Acrylamide T 12% Constant. 15µg of protein per lane of crude preparation **Panel B:** Gel Acrylamide T 8% Constant. 10µg of protein per fraction obtained in Method 1. **Panel C:** Gel Acrylamide T 12% constant. 10µg of protein per fraction obtained in Method 2. M: Molecular weight markers (in kDa), 1: P1, 2: P18, 3: P200, 4: SN200, 5: P18CP18, 6: P18DP18 7: P18CP200, 8: P18DP200, 9: P200C, 10: P200D, 11: SN200C, 12: SN200D, 13: SN200CP200, 14: P200CP200, 15: SN200CP18, 16: P200CP18.

The large band at around 100kDa is the monomeric form of Tamm-Horsfall Protein (THP). It abundantly sediments at a low-speed (1000g) and then progressively, in decreasing amount is recovered in the following centrifugation steps (**P18** and **P200**), with traces left in the final supernatant (**SN200**) (Figure 2.3 Panel A). After either DTT or CHAPS treatment of the crude **P18** and **P200** sediments, respectively, the pellet recovered again at a relatively low RCF (**P18CP18, P18DP18, P18CP200** and **P18DP200; for abbreviations, see Fig. 2.1**) showed only traces of THP. Surprisingly, after CHAPS treatment, in method 1 where the SNs were combined (Fig. 2.1), all the THP was recovered in the ultracentrifugation pellet (Fig. 2.3, Panel B **P200C**) instead of being in the SN (Fig. 2.3, Panel B **SN200C**) as was observed for DTT treatment (Fig. 2.3, Panel B **P200D** and **SN220D**). Interestingly, without pooling the SNs, THP present in the crude pellet (**P18**) after CHAPS treatment was no longer recovered at 18,000g, but at 200,000g (Fig. 2.3, Panel C **P200CP18**). CHAPS treatment of crude P200

pellet without pooling gave a final sediment (Panel C, **P200CP200**) in which Tamm-Horsfall protein was present at the same amount with respect to the DTT pellet (Panel B **P200D**). Moreover, albumin (~66kDa), found in abundance in the crude pellet (Panel A, **P200**), was instead detectable in the supernatant after CHAPS treatment (Fig. 2.3, Panel C **SN200CP200**), with traces remaining in the 200,000g pellet (Panel C, **P200CP200**).

2.3.4 Profiling of urinary markers by Western Blotting

Figure 2.4 shows the Western blotting of exosomal and podocyte markers.

2.3.4.1 CD63

CD63 was abundantly found in **P18** and in the corresponding low-speed detergent-treated and DTT precipitates (Panel A; **Figure 2.4**).

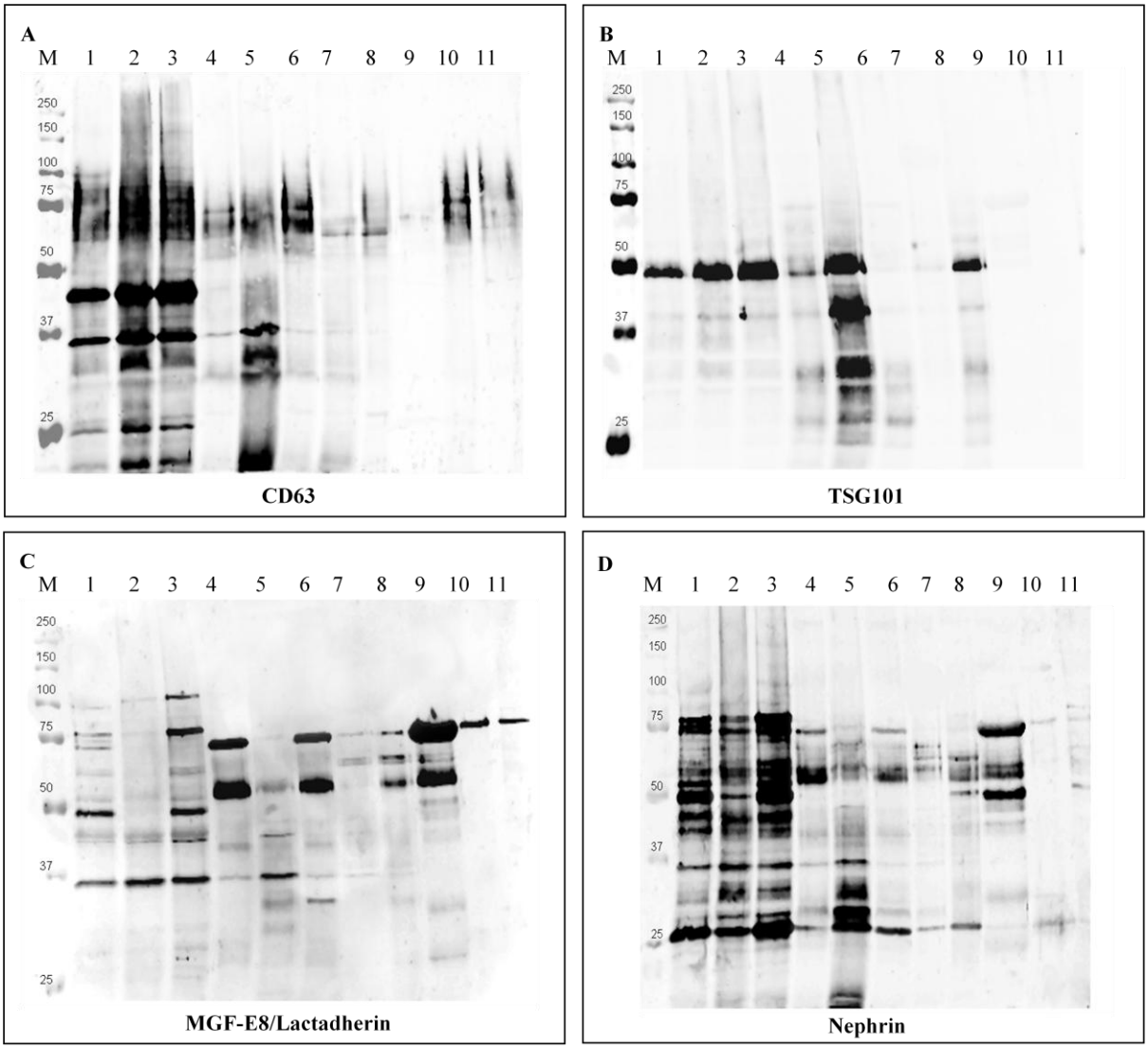

Figure 2.4: Western blotting analysis. A: Rabbit anti-CD63, B: Rabbit anti-TSG101, C: rabbit anti-MGF-E8/Lactaderin and D: rabbit anti-nephrin (Hara *et al.,* 2010). 10µg of protein of fractions obtained in Method 1 were loaded on the gels. M: Molecular weight markers (in kDa), 1: P18, 2: P18CP18, 3:P18DP18, 4: P18DP200, 5: P18CP200, 6: P200D, 7: P200C, 8: P200, 9: SN200 10: SN200C, 11: SN200D.

The CD63 isoform profiles exhibit a distinct distribution in these vesicle preparations, with sizes of approximately 65, 45 and 35kDa, respectively, consistent with a heavy glycosylation pattern (Ageberg & Lindmark, 2003). In treated high-speed pellets (**P200C** and **P200D**) and the crude **P200** fraction, CD63 was observed mainly at higher molecular weights (75 to

100kDa). The broad band seen in Western blotting is most likely due to the high-degree of glycosylation by poly *N*-acetyl lactosamine, as reported by Engering *et al.,* (Engering *et al.,* 2003). Heavy glycosylation can cause streaking and smearing of glycoproteins in the SDS-PAGE system.

2.3.4.2 TSG101

TSG101 was detected in the exosome fraction as a single band of approximately 46 kDa in **P18** and **P200** crude preparations but, significantly, after this fraction was treated (CHAPS or DTT), the antigen was recovered in the low-speed centrifugation preparations, namely **P18CP200** and **P18DP200**. In the latter two samples, in addition to the main band of approximately 46kDa, four other lower molecular weight isoforms were also detected. It appears that the lower molecular weight bands are a specific feature of these two samples, especially of the CHAPS-treated fraction.

2.3.4.3 LACTADHERIN/MFG-E8

In the P18 preparation, two main bands were visible with molecular weights of approximately 48kDa and 35kDa of lactadherin. Interestingly, after CHAPS treatment, only the 35kDa band was detected (**P18CP18** and **P18CP200**). Notably, in the corresponding DTT pellet, (**P18DP18**), both the 46 and 35kDa bands were present, in addition to two extra isoforms of approximately 77 and 100kDa, specifically enriched in this fraction. The analysis of the crude **P200** preparation revealed three main bands whose molecular masses were approximately 53, 62 and 70kDa. Interestingly, the 53 and 70kDa forms were specifically enriched in both DTT preparations originating from the **P200** samples.

2.3.4.4 NEPHRIN

A recent report has described the ability of glomerular podocyte microvilli to secrete non-exosomal vesicles by tip vesiculation (Hara *et al.*, 2010). This finding led us to explore this

podocyte-specific marker of the slit-diaphragm in detail in the vesicle preparations, especially in the CHAPS or DTT-treated vesicles, in order to scout for specific functional features reflecting the critical elements of the glomerular filtration barrier. The whole array of samples was probed with an anti-nephrin antibody raised against the intra-cytoplasmic domain of the protein (Ahola *et al.,* 1999). Surprisingly, the bulk of the nephrin-containing vesicles were recovered at low RCF, (**P18CP18, P18DP18, P18CP200** and **P18DP200**). In the whole sample set, no full-length nephrin was observed, although different nephrin fragments were detected in the diverse preparations. This is in line with previous results from our group (Patari *et al.,* 2003). In the soluble fractions, nephrin was detectable only in the **SN170** fraction as 75, 65 and 48kDa bands, while in the vesicle preparations, a wider range of bands was detected at 75, 65, 48 and 35 and 25kDa.

2.3.5 Dipeptidyl peptidase IV (DDPIV) and nephrilysin (NEP) activity in CHAPS and DTT fractions

Figure 2.5 shows DPP IV and NEP activities measured in pellet and SN after CHAPS and DTT treatments. Before measuring the activity, samples were dialysed with a membrane with a molecular weight 'cut-off' of 300kDa to remove DTT, CHAPS and potential soluble isoforms of these two proteases. Figure 2.5B shows Coomassie staining and the immunodetection of DPP IV after dialysis. In DTT 200,000g pellets and SNs (Anti DPP IV, Panel B lane 1 and 3), DPP IV is visible at 110 kDa, while in CHAPS fractions two specific bands were detected whose molecular weights were approximately 52 and 25kDa (Anti DPP IV, Panel B lane 2 and 4). Activity was recorded as pmol of product/min/mg protein (UP/mg protein). When compared, DPP IV activity in DTT and CHAPS fractions were similar despite the difference in the respective western blotting patterns as shown in Figure 2.5A.

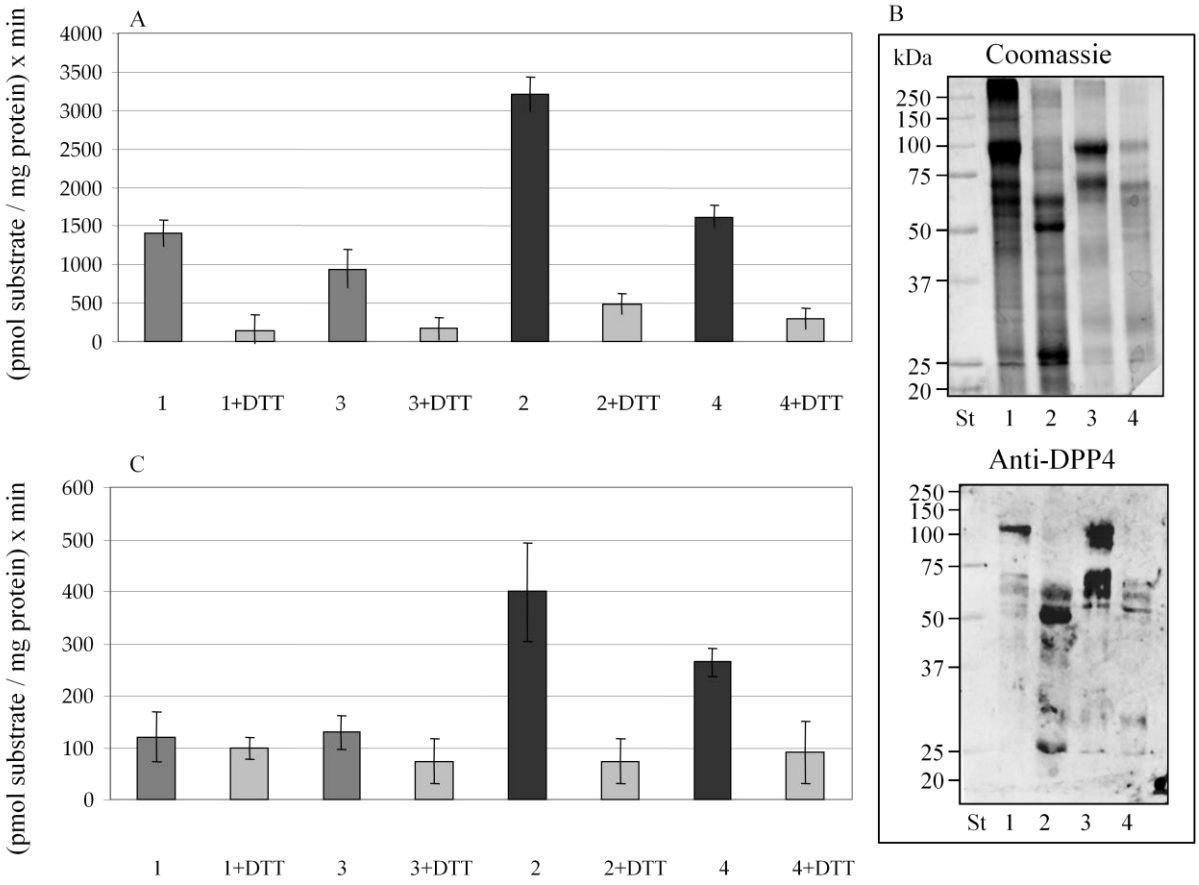

Figure 2.5: Protease activity. Membrane-bound DPP IV (**Panel A**) and NEP (**Panel C**) peptidase activity profiles recorded in absence and presence of 10 mM DTT. Samples were dialysed at a MWCO of 300kDa. DTT pellet 200,000g (sample 1), DTT SN 200,000g (sample 3), CHAPS pellet 200,000g (sample 2) and CHAPS SN 200,000g (sample 4) are represented. Columns compare DTT vs CHAPS after dialysis with a membrane of MWCO 300kDa and in the presence of 10mM DTT. Values represent mean ± SD of units of peptidase (UP) per milligram of protein per minute. **Panel B** represents the Coomassie gel and DPP-IV immunodetection of the same samples. 10µg of protein per fraction obtained in Method 1 were loaded on the gels after 300kDa MWCO dialysis.

Activity drastically decreased in the presence of DTT, suggesting that, although disulfide bond reduction affects the activity, during dialysis DPP IV is able to regain its native

conformation with a good recovery of its activity. For NEP, the activity decreased significantly (around 4-fold) in the DTT fractions when compared with the corresponding CHAPS fraction. This result shows that although DTT was removed by dialysis, there was a detrimental and irreversible effect of the reducing agent on the folding of NEP which led to a loss of the enzymatic activity. Once again, as a control, in the presence of DTT the activity decreased substantially in all the fractions. The activity recorded in the SNs after dialysis with a molecular weight 'cut-off' membrane of 300kDa, along with CD63 signals detected in Western blotting (Figure 2.4) indirectly points to the presence of nano-vesicles in these fractions which were not collected in the pellet after the second ultracentrifugation (200,000g).

2.3.6 MS-based identification of proteins released in the supernatant after DTT and CHAPS treatments

The supplemental tables list the proteins identified by the systematic MS analysis of the final 200,000g supernatant after DTT (Supplementary Table S2.1 and S2.2, 274 unique proteins) and CHAPS (Supplementary Table S2.3, 234 unique proteins) treatments. Figure 2.5 shows the analysis in comparison to the two published data sets of proteins identified in the 200,000g pellet (Gonzales *et al.*, 2009) and SN (Kentsis *et al.*, 2009), respectively.

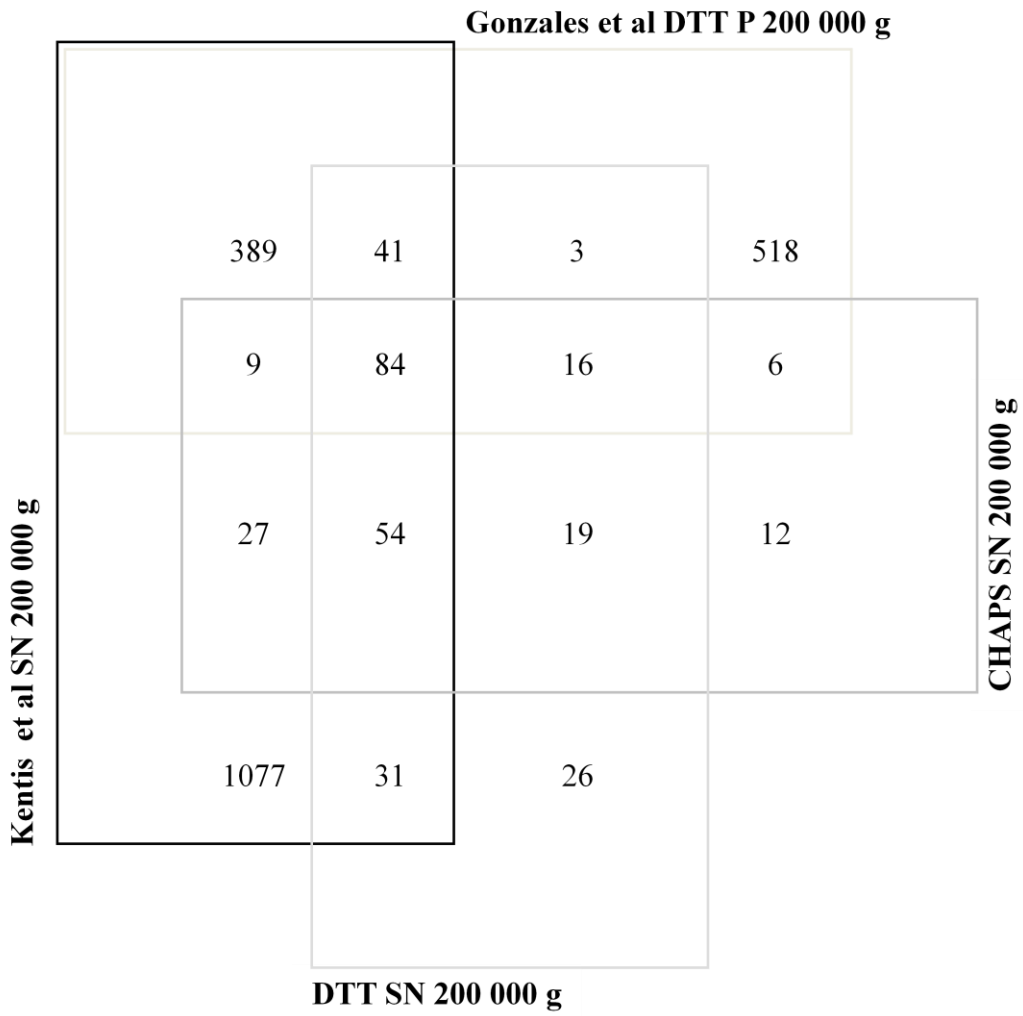

Figure 2.6: Protein identification comparisons. Venn diagram showing the distribution of the number of identified proteins presents in SN 200,000g after CHAPS and DTT treatments. Protein identifications from the current study were compared to two other studies which were carried out using high-resolution mass spectrometers in gels on 200,000g pellets after DTT treatment (Gonzales *et al.,* (Gonzales *et al.,* 2009)) and 200,000g supernatants (Kentsis *et al.,* (Kentsis *et al.,* 2009)).

The Venn diagram shows that 76.2 % of proteins found in the CHAPS fraction are common to the corresponding DTT fraction. This result confirms the close similarity between the two methods in respect to the preservation of major protein components. Interestingly, approximately 9.7% and 6.9% of protein in CHAPS and DTT supernatants, respectively, are

consistent with the observations made by Gonzales (Gonzales *et al.*, 2009), while 33.7% and 31% of proteins in CHAPS and DTT supernatants, respectively, are common to the Kentsis (Kentsis *et al.*, 2009) data-set while 40.9% and 45.6% of the proteins in CHAPS and DTT SN, respectively, overlap with the two reference data sets. Finally, around 10% of proteins detected in our study are unique for both CHAPS and DTT supernatants.

The bulk of identified proteins were found to be already annotated in the whole urine, but others, like Ig gamma-4 chain, Ig gamma-3 chain C region and Ig alpha-2 chain, have specifically been shown to associate with exosomes. Others, including Tumor Necrosis Factor receptor superfamily member 19L or palmitoyl-protein thioesterase 1 have all the potentiality to be part of the exosome proteome (Table 2.1).

Table 2.1: Protein identification in DTT and CHAPS supernatants. Fractions in which proteins were obtained, Uniprot accession, % sequence coverage, peptides, mascot score, protein name description, presence in Exocarta and *Adachi et al.,* (Adachi *et al.*, 2006) *and Marimuthu et al.,* (Marimuthu *et al.*, 2011) is indicated in the table.

Fraction	Accession	Coverage %	Peptides	Mascot score	Descriptions	Exocarta	Adachi	Marimuthu
CHAPS	O95497	10.33	4/4	180.32	Pantetheinase GN=VNN1	N	Y	Y
CHAPS	P50897	9.15	2/2	97.96	Palmitoyl-protein thioesterase 1 GN=PPT	N	N	Y
CHAPS	P01040	30.61	2/2	109.9	Cystatin-A GN=CSTA	N	N	Y
CHAPS	Q86SR0	45.36	2/2	114.84	Secreted Ly-6/uPAR-related protein 2 GN=SLURP	N	N	N

					2			
CHAPS	Q969Z4	2.56	2/2	109.04	Tumor necrosis factor receptor superfamily member 19L GN=RELT	N	Y	Y
CHAPS	P61916	10.60	1/1	66.37	Epididymal secretory protein E1 GN=NPC2	N	Y	Y
CHAPS/ DTT	P01861	18.35	7/4	354.82	Ig gamma-4 chain C region GN=IGHG4	Y	N	N
CHAPS/ DTT	P01877	18.53	7/5	360.53	Ig alpha-2 chain GN=IGHA2	Y	N	N
CHAPS/ DTT	P37235	6.22	1/1	81.02	Hippocalcin-like protein 1 GN=HPCAL 1	Y	N	Y
CHAPS/ DTT	P15289	8.09	2/2	123.11	Arylsulfatase A GN=ARSA	N	N	Y
CHAPS/ DTT	P41222	24.74	6/4	304.59	Prostaglandin-H2 D-isomerase GN=PTGDS	N	Y	Y
CHAPS/ DTT	Q14508	38.71	4/3	216.44	WAP four-disulfide core domain protein 2 GN=WFDC 2	N	Y	Y
CHAPS/ DTT	P31949	15.24	1/1	80.02	Protein S100-A11 GN=S100A1	Y	Y	Y

					1			
DTT	P02788	6.48	5/4	227.66	Lactotransferrin GN=LTF	Y	Y	Y
DTT	P19835	16.61	7/7	339.62	Bile salt-activated lipase GN=CEL	Y	N	Y
DTT	Q9Y646	16.95	6/6	334.27	Plasma glutamate carboxypeptidase GN=PGCP	N	N	Y
DTT	P55259	4.28	2/2	136.92	Pancreatic secretory granule membrane major glycoprotein GP2 GN=GP2	N	N	Y
DTT	P01860	13.79	5/4	256.39	Ig gamma-3 chain C region GN=IGHG3	Y	N	N

These results suggest that in specific subpopulations of exosome vesicles, unique features can still be recovered in the SN after DTT and CHAPS treatments. This finding has practical consequences for their use. Figure 2.7 reports the gene ontology distribution per protein class of identified proteins by the Panther classification system (www.panther.org) (Thomas *et al.*, 2003). Once again the distribution of the protein classes is very similar in the 200,000g CHAPS and DTT supernatants.

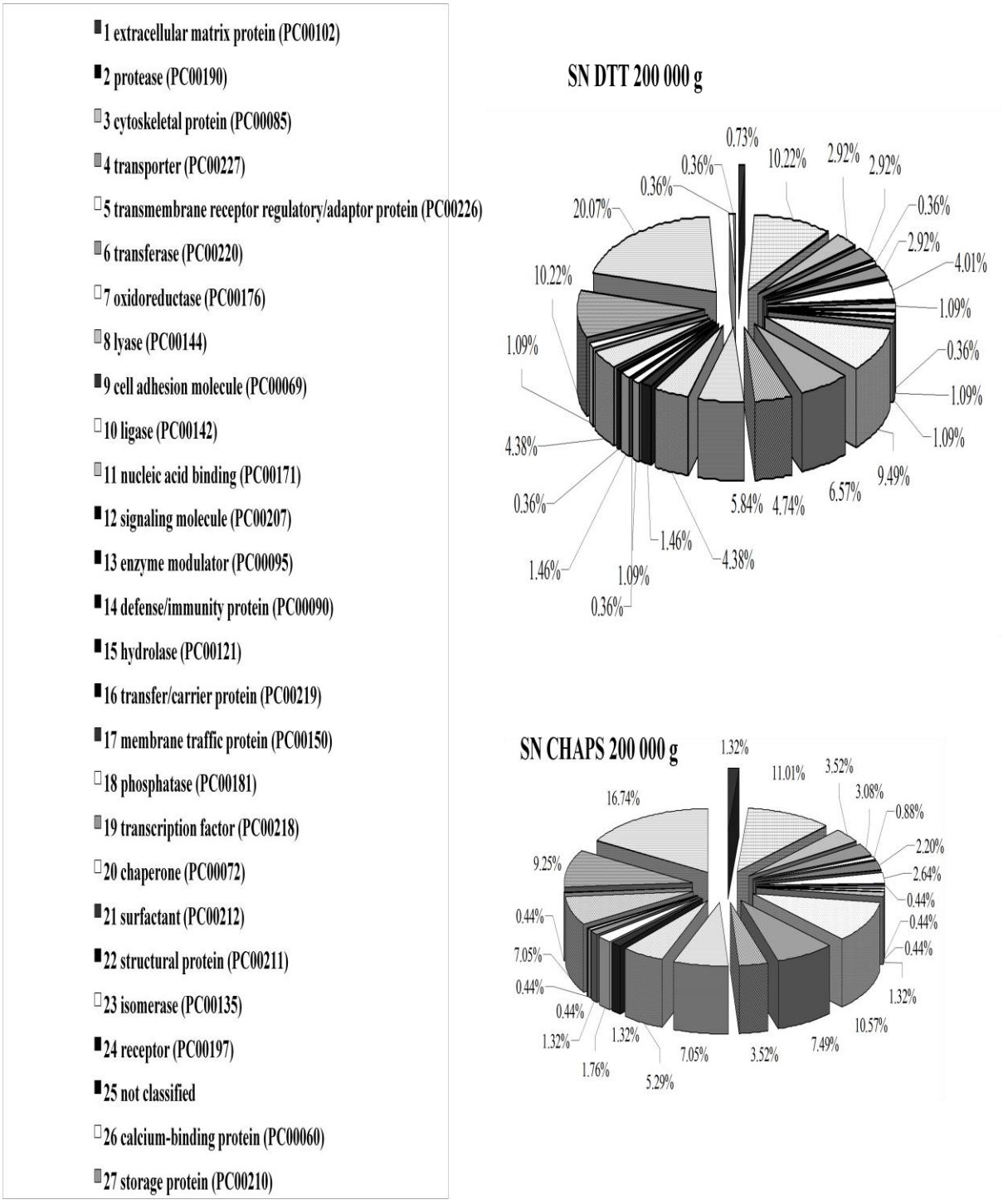

Figure 2.7: Gene ontology based classification (www.panther.org) of proteins identified in DTT and CHAPS SN200,000g**.**

2.3.7 Interaction of abundant proteins like albumin, THP and IgG with nano vesicles treated with either DTT or CHAPS

High-abundance proteins in membrane vesicle fractions were investigated by western blotting in various membrane vesicle fractions. The results are shown in Figure 2.8.

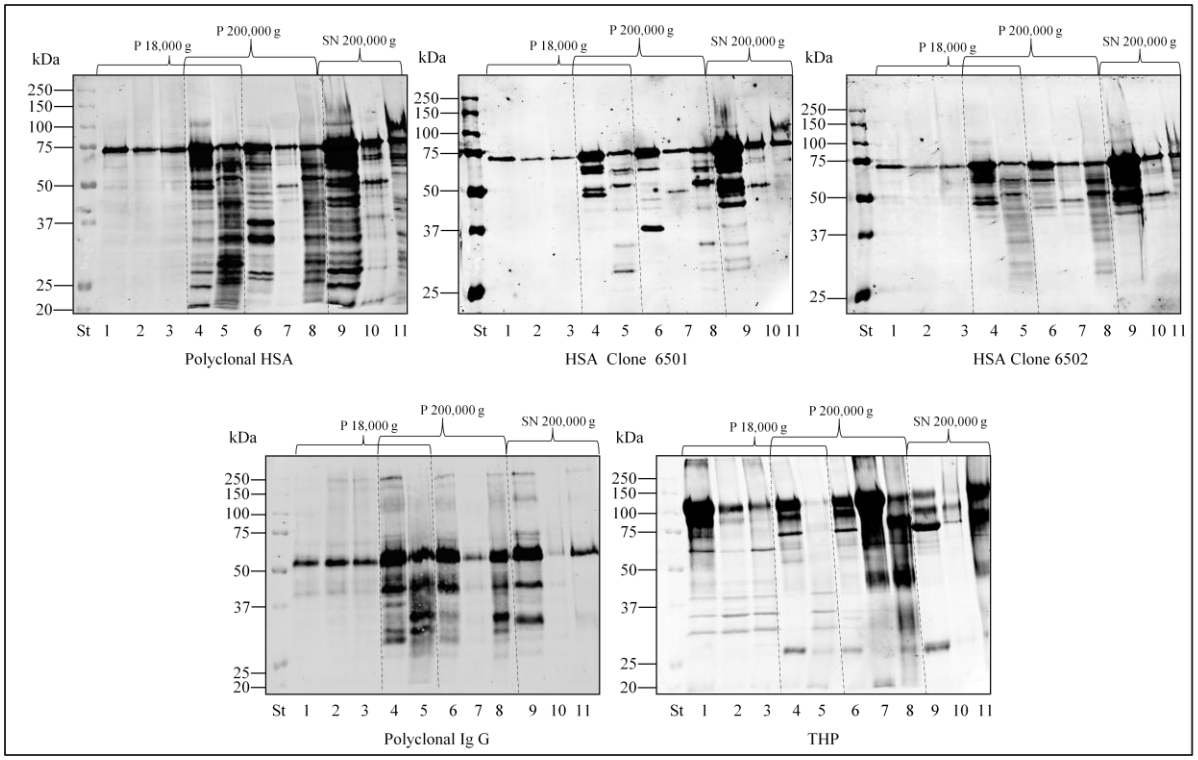

Figure 2.8: Western blot analyses of nano vesicle pellets and SN in crude preparations and after CHAPS and DTT treatment. Lane 1 18,000g pellet; lane 2 18,000g pellet CHAPS 18,000g pellet; lane 3 18,000g pellet DTT 18,000g pellet; lane 4 200,000g pellet DTT 18,000g pellet; lane 5 200,000g pellet CHAPS 18,000g pellet; lane 6 SN 18,000g CHAPS 200,000g pellet; lane 7 SN 18,000g DTT 200,000g pellet; lane 8 200,000g pellet; lane 9 SN 200,000g; lane 10 SN 18,000g CHAPS SN 200,000g; lane 11 SN 18,000g DTT SN 200,000g. Polyclonal anti-HSA and monoclonal anti-HAS clone 6501 and 6502. Polyclonal anti-IgG and polyclonal anti-THP diluted 1:1,000 overnight RT.

Albumin detection with two separate monoclonal antibodies showed that albumin is distributed in all fractions even after CHAPS and DTT treatment. Although the signal is present in the SN 18,000g CHAPS/DTT SN 200,000g (lane 9 and 10), signal was still visible in the respective pellets (lane 6 and 7). Notably, treatment with CHAPS and DTT could not affect the precipitation of albumin into the soluble phase. It is also interesting to note that HSA, which co-precipitated in the 18,000g pellet (lane1, 2 and 3) did not reveal a fragmentation pattern consistent with that of the 18,000g and 200,000g pellets resulting from CHAPS and DTT treatment of the crude 200,000g pellet (lane 4, 5, 6 and 7). This postulates that these different fractions of membrane vesicles have different patterns of albumin fragmentation which co-precipitate specifically with sub populations of nano vesicles. It could be due to a specific protease present in these sub-populations or differential PTMs of albumin. Full-length IgG was found to be present in P18,000g (Lane 1-3) while extensive fragmentation as well as full length IgG was seen in P200,000g pellets both crude and treated (Lane 6 & 8). THP similarly was present to significant levels in the P200,000g (Lane 6-8). This prompted us to investigate whether these abundant proteins are present with membrane vesicle fractions in a non-specific manner or if there is a specific interaction of these proteins with membrane vesicles.

Our preliminary investigations were focused on verifying whether a biomolecular interaction was occurring between both sets of detergent-treated nano vesicle populations (e.g. CHAPS- and DTT-treated) and albumin, as suggested by Western blotting analysis (Figure 2.8 anti-albumin antibodies, polyclonal and monoclonal). For the separation of monomeric and dimeric fractions of HSA by SEC, a significant improvement in resolution was seen when three inter-connected stationary phases were selected instead of one and, hence, this system was implemented during this study. Dimeric and monomeric species with retention times of 26.8 and 28.4 minutes, respectively, were readily identified using the gel filtration standard

for size determination (Figure 2.9 A), while the ratio of monomer:dimer was estimated to be 12.4:1 based on the calculation of peak areas (Figure 2.9 B). Monomeric fractions, estimated to be approximately 40-45kDa in size, were manually collected and analysed by SDS-PAGE to monitor purity and verify the absence of superfluous albumin fractions (Figure 2.9 C).

Interaction analysis studies were firstly performed by FLISA before proceeding to more in-depth SPR analysis. Here, the CHAPS-treated 200,000g pellet (P18+200CP200) was coated in triplicate on individual wells of a 96-well plate and after blocking and washing, was subsequently probed with the retained monomeric albumin fraction and detected as described in the methods section. Comparative analysis of binding at pH 6.0 and pH 7.4 revealed that more favourable binding was seen at the higher pH (Figure 2.9 D), so subsequent SPR interaction studies were performed using a running buffer (HBS) at a physiological pH (7.4).

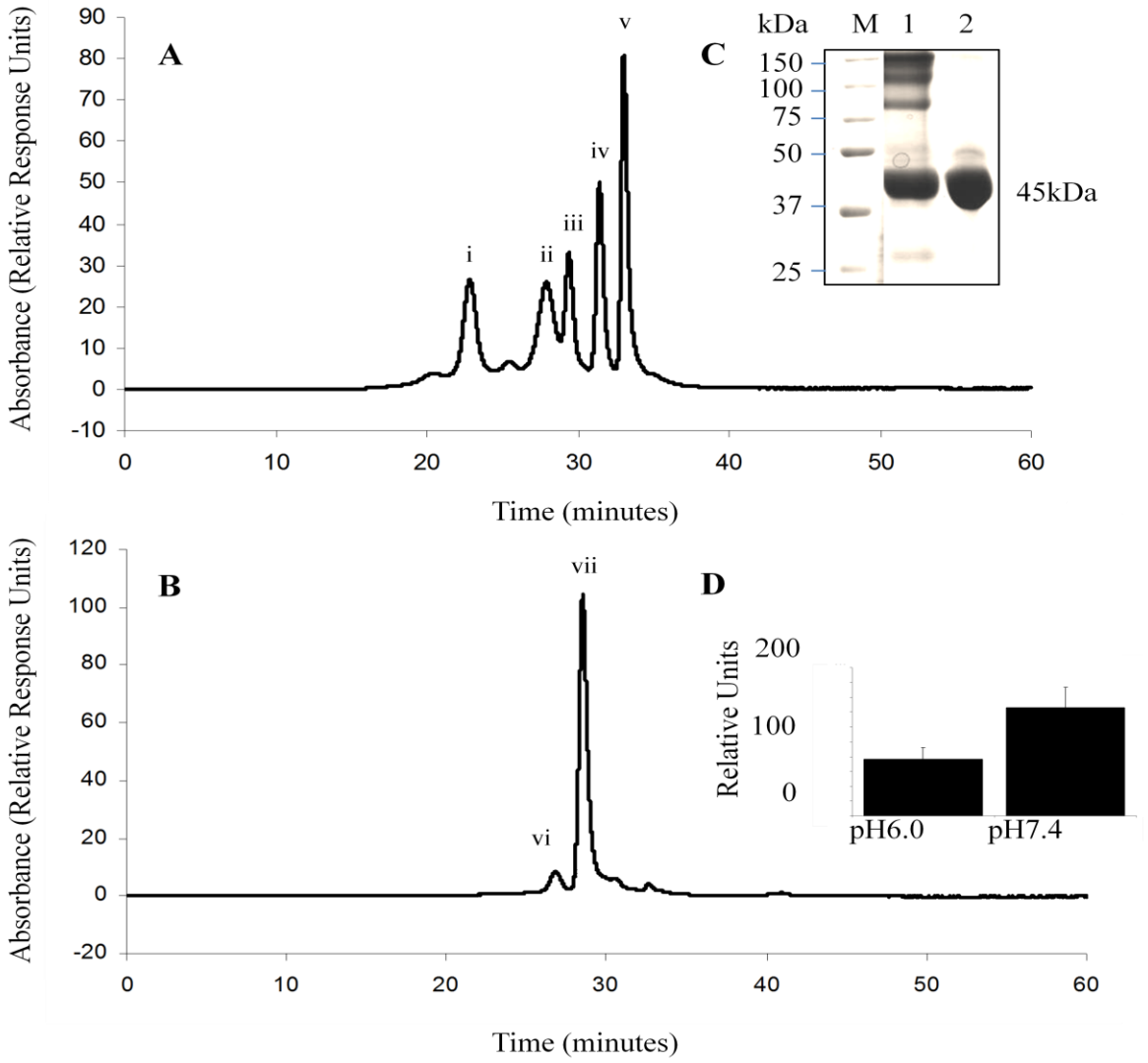

Figure 2.9: Species composition of human albumin and preliminary binding studies. **(**A) Retention times for individual components in the gel filtration standard were as follows: bovine thyroglobulin (i; 22.8mins), IgG (ii; 27.9mins), ovalbumin (iii; 29.5mins) myoglobin (iv; 31.4mins) and vitamin B12 (v; 33.1mins). (B) Dimeric (vi) and monomeric (vii) species of commercial human albumin was readily detected by SEC, and subsequently analysed by SDS-PAGE to ascertain purity of retained fractions (40-45kDa). (D) A pilot study using CHAPS-treated nano vesicle populations demonstrated more favourable binding at physiological pH by FLISA. The background fluorescence signal for the negative control was subtracted from those obtained at pH values of 6.0 and 7.4, respectively.

Aliquots of DTT- and CHAPS-treated nano vesicle preparations were reconstituted in HPLC-grade de-ionised water (filter-sterilised) to 980µg/mL and 509µg/mL, respectively, before being diluted to a final concentration of 100µg/mL. Next, pre-concentration studies were used to determine the optimal pH for the immobilisation of DTT- (pH3.6) and CHAPS-treated populations (pH3.8) on flow cells 2 and 4, respectively. EDC/NHS-based coupling and ethanolamine capping resulted in the capture of 15,952.7RU of DTT-treated exosomes, while 8,530RU of CHAPS-treated exosomes were captured. Nano vesicle populations were previously shown by TEM analysis to be comprised of a panel of entities whose sizes ranged from 20-200nm. It suggests that heterogenous surfaces (having vesicles of multiple sizes) were therefore presented for interaction analysis with high-abundance proteins. Regeneration scouting was used to verify that 2.5mM NaOH could be selected to ensure a stable baseline on both flow cells, with 20 regenerations (5µL injections) resulting in a loss of 15RU and 18RU from FC2 and FC4, respectively. Hence, these conditions were deemed to be appropriate and were selected for albumin binding studies.

We initially investigated whether the nano vesicle populations isolated by CHAPS- and DTT-treatments were already coupled to albumin at the time of their immobilisation, which would also suggest an interaction prior to cross-linking. To test this hypothesis, a dose-response curve with polyclonal anti-albumin antibody at five concentrations, namely 10µg/mL, 20µg/mL, 50µg/mL, 100µg/mL and 200µg/mL, was tested on both flow cells by interday analysis. The antibody bound to both CHAPS- and DTT-treated populations, with apparent saturation seen at the lowest concentration selected (10µg/mL). In contrast, binding to DTT-treated populations was in a concentration-dependent manner, with the introduction of elevated concentrations of analyte correlating with an increase in RU (Figure 2.10 A).

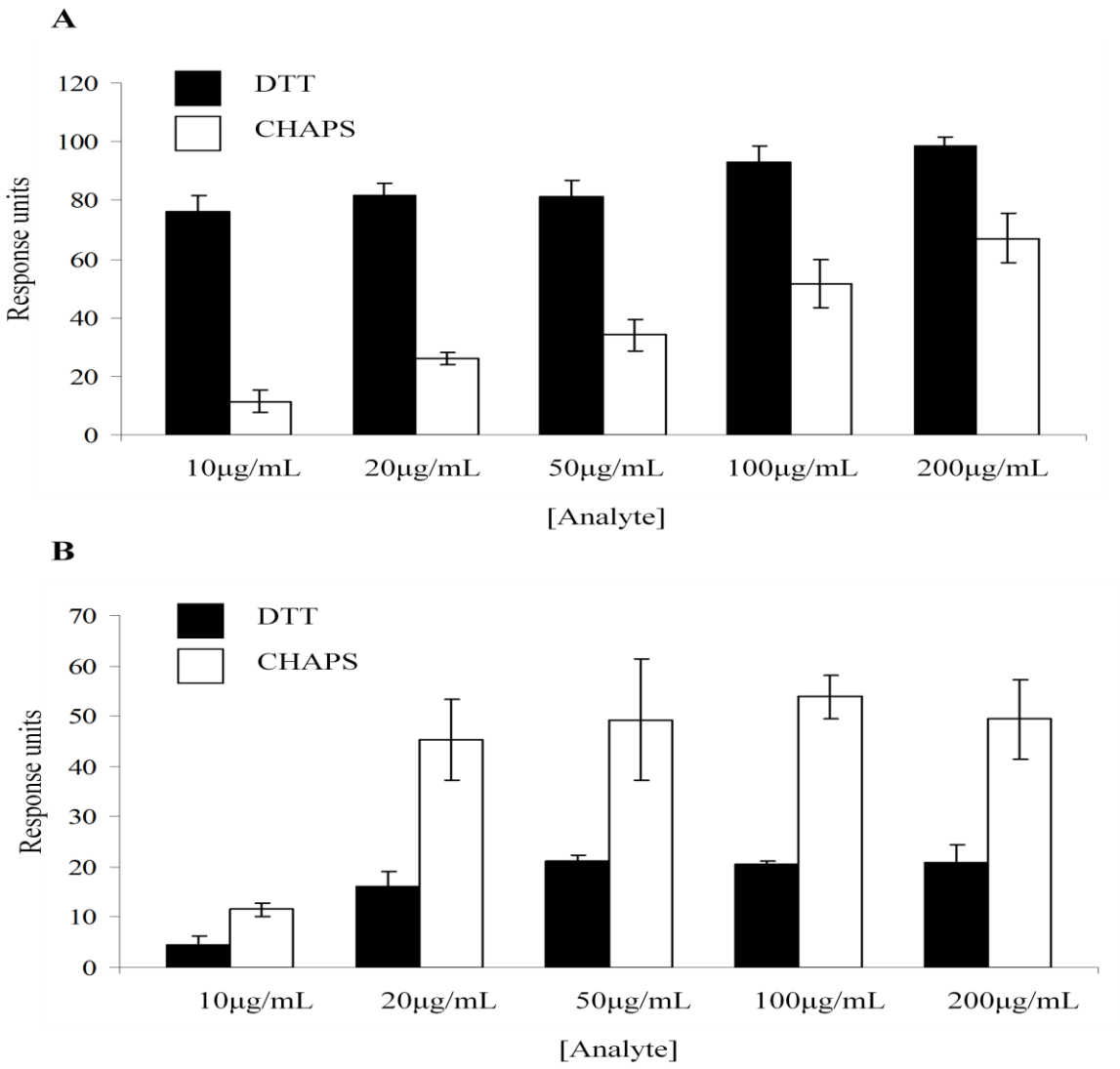

Figure 2.10: Comparative binding analysis of albumin to CHAPS-treated nano vesicle populations at two pH values by SPR. Comparative analysis of the binding of anti-albumin (A) and albumin (B) to DTT-and CHAPS-treated nano vesicle populations immobilised on flow cells 2 and 4, respectively. Signals shown are for interday analysis, with the signal obtained from the HBS negative control (analyte-free) and the reference flow cells (e.g. flow cell 1 for DTT-nano vesicles and flow cell 3 for CHAPS) subtracted so as to show true binding to immobilised populations. Erros bars represent standard deviation from the mean.

These observations suggested that the immobilised nano vesicles were coupled with albumin, and that detergent treatment did not effect the removal of this soluble protein from either population (DTT and CHAPS trated vesicles). Next, as a further verification of the occurrence of a binding event between albumin and both vesicle populations independently of one another, the monomeric fraction of human albumin previously selected for FLISA-based analysis was tested on both flow cells using the same dose-dependence response curve as previously used for anti-albumin. More favourable binding was evident for the CHAPS-treated population, and apparent saturation was seen at an analyte concentration of 20µg/mL for each ligand (Figure 2.10 B).

Based on these observations, we investigated whether an anti-albumin immunoprecipitation (IP) could be performed with the objective of selectively enriching exosomes from the nano vesicle population, with polyclonal anti-HSA antibody selected as a biorecognition element and a monoclonal anti-HSA antibody implemented for detection purposes. Nano vesicles isolated by treatment with CHAPS (P18+170CP170) were selected as an analyte, primarily due to the observation in SPR analysis of more favourable binding to albumin. Immunoprecipitation of the full-lenght HSA molecule, and not proteolysis-derived fragments, was seen when captured fractions were eluted, analysed by SDS-PAGE (Figure 2.11 A) and subsequently probed with monoclonal anti-HSA antibody (Figure 2.11 B). However, proteolysed fragments of albumin were present in the unbound fraction.

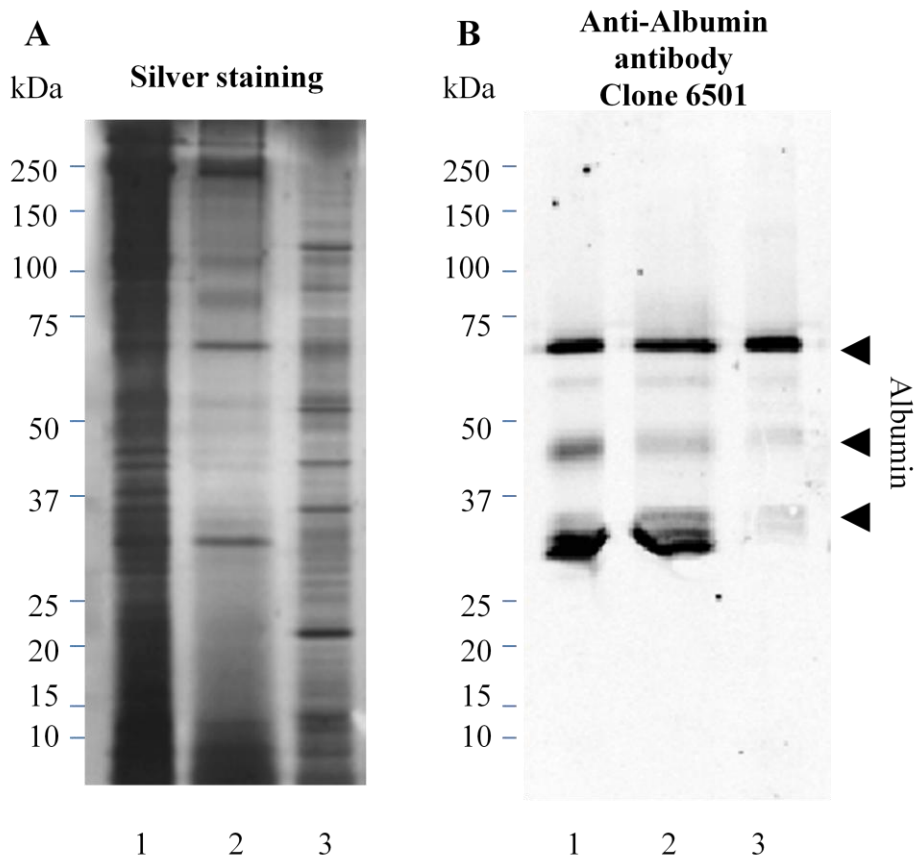

Figure 2.11: Immunoprecipitation of urinary nano vesicles using monoclonal anti-albumin antibody for biorecognition. Silver staining (A) and Western blotting (B) analysis of non-bound (Lane 1), wash (Lane 2) and elution fractions (Lane 3), verifying the presence of three isoforms of albumin.

In our previous report, we outlined that in the aforementioned CHAPS-treated nano vesicle population (P18+200CP200) CD63, a well-established exosomal marker (Logozzi *et al.*, 2009; Nilsson *et al.*, 2009), was detected in relatively low intensities by Western blotting (Musante *et al.*, 2012). Interday SPR analysis using a single concentration of anti-CD63 antibody (10µg/mL) was performed to verify this result with respect to binding to both nano vesicle populations, and determine whether this antibody could also be selected as an alternative to anti-albumin for immunoprecipitation. Here, it was observed that weak binding was occurring between anti-CD63 and DTT-treated exosomes (μ=16.9RU; δ= 1.34), with

comparable results (μ=15.1RU; δ= 4.31) obtained for vesicles treated with CHAPS. Subsequent analysis of immunoprecipitation fractions using anti-CD63 demonstrated a clear enrichment of CD63 in the eluate fraction (approximately 50-55kDa), which was not detected in the non-bound and washing fractions (Figure 2.12).

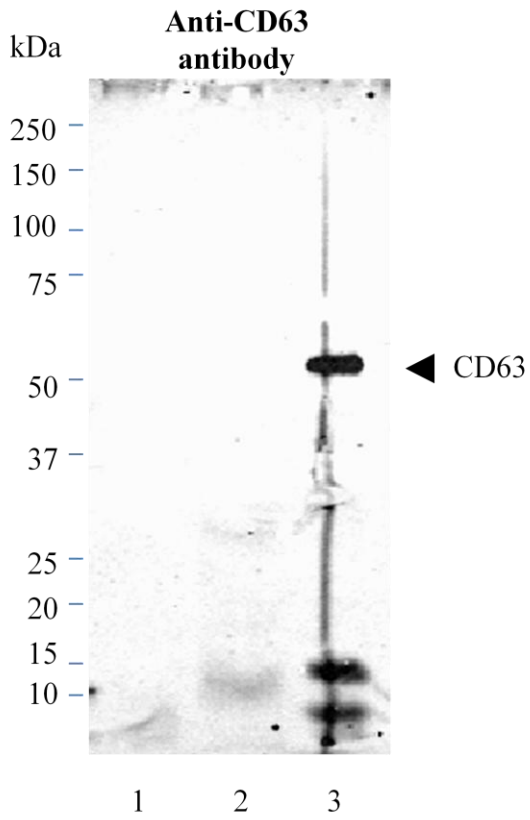

Figure 2.12: Immunoprecipitation of urinary nano vesicles (using anti-albumin antibody) and detection in a western blot using anti-CD63 antibody. Western blotting analysis of non-bound (Lane 1), wash (Lane 2) and elution fractions (Lane 3); CD63 (~55kDa) is present in the elution fraction but not in other fractions.

It is well-established that the co-precipitation of THP, with urinary nano vesicles during exosome/urinary nano vesicle isolation, is problematic when attempting to isolate populations for proteomic, glycomic or transcriptomic analysis. Moreover, THP polymers are known to

interact with urinary nano vesicles (Fernandez-Llama *et al.*, 2010). Furthermore, in our previous analysis we demonstrated that sedimentation of this protein in the 1,000g pellet significantly reduced the quantity of THP in subsequent fractions, which permitted us to obtained improved yields of pure nano vesicles (Figure 2.3). As a demonstration of the ability of this soluble protein to bind to both CHAPS- and DTT-treated populations, DEF-extracted THP was selected as an analyte and assayed on both nano vesicle populations by SPR, using the same dose-response curve as described previously. For DTT-treated vesicles, an increase in analyte concentration correlated with higher RU values. Notably, for vesicles isolated using a CHAPS-treatment, an inverse trend was seen. Here, binding to THP was significantly lower for each concentration selected for analysis. Furthermore, introducing increased concentrations of analyte resulted in reduced binding to surface-immobilised nano vesicles (Figure 2.13).

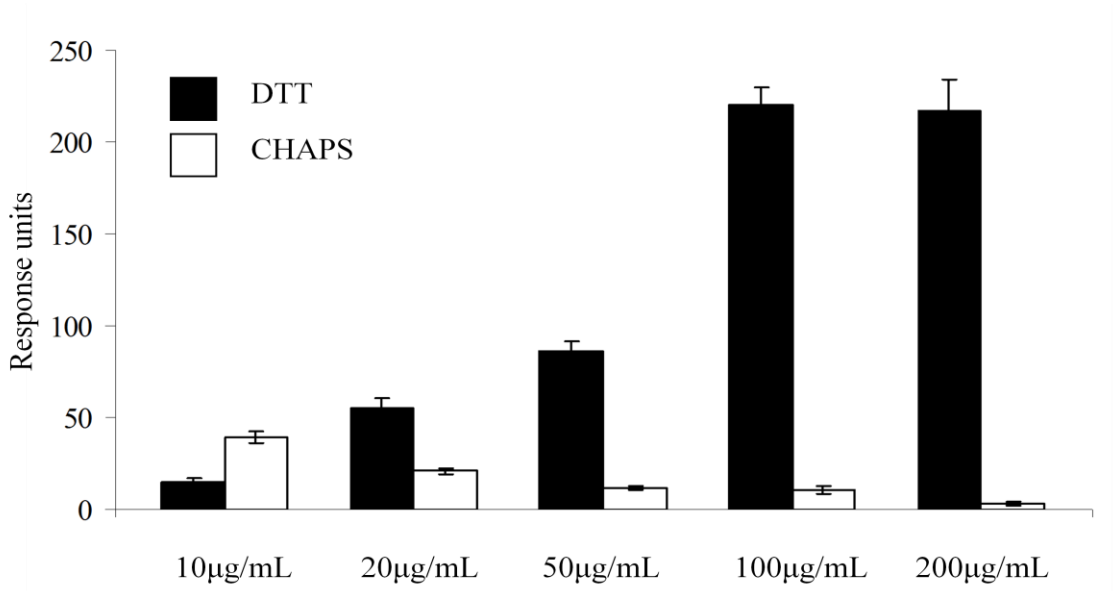

Figure 2.13: Biomolecular interactions with THP. Comparative analysis of the binding of DEF-purified THP in solution to DTT- and CHAPS-treated populations by SPR, using a dose-dependent response curve and interday analysis. X-axis is the THP concentration and Y-axis is the response unit obtained upon interaction with vesicles.

This inverse trend of reduced binding seen upon introduction of increasing concentrations of THP to the CHAPS-treated fraction could be due to the apparent Hooks effect. Western blotting using anti-IgG demonstrated that this immunoglobulin was resident to both CHAPS- and DTT-treated nano vesicle populations (Figure 2.8, polyclonal IgG). To verify whether IgG could interact with the nano vesicle fractions selected herein for binding studies, human IgG was selected as an analyte and assayed using identical conditions to the aforementioned interaction analyses. In contrast to what was observed for THP binding studies, an increase in RU as a function of IgG concentration was observed for ligands on flow cells 2 and 4, respectively, while no apparent saturation was observed during this analysis (Figure 2.14). The reference flow cells 1 and 3 were used as a control by passing equal amounts of IgG over them and subtracting their value from the interaction shown here.

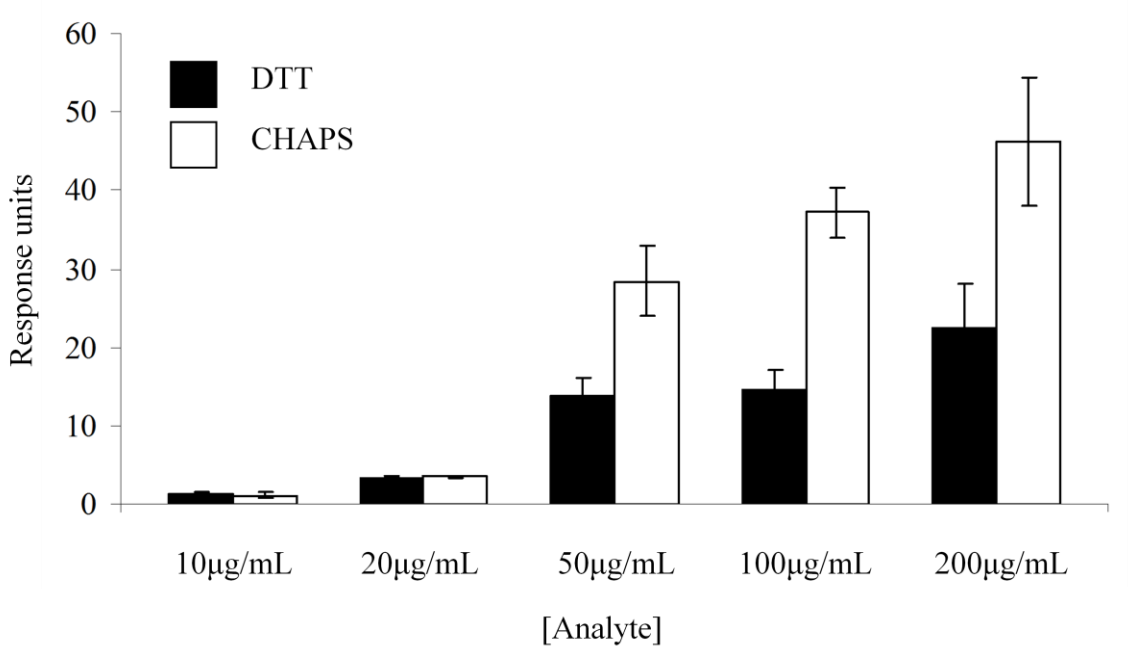

Figure 2.14: Biomolecular interactions with IgG. Comparative analysis of the binding of IgG in solution to both nano vesicle populations (DTT-treated and CHAPS-treated) by interday SPR analysis.

2.4 Discussion

Exosomes found in all biological fluids, including the urine, are emerging as a novel class of cell products plausibly revolutionising our views of cell behaviour and processes like intercellular communication (Van Niel *et al.,* 2006). While an increasing variety of exosome constituents including specific proteins, DNA and RNA species are being identified, little is still understood of their precise biological roles. Likewise, the optimal method for exosomal isolation and to determine of thier interactions, to appreciate their distinct roles in biology remains to be characterised in detail. The aim of this study was to optimize methods which increase the recovery of urinary exosomal vesicles and optimise their preservation for functionality without notable interference with inherent soluble protein constituents of urine.

Tamm-Horsfall protein (THP) is the most abundant glycoprotein normally found in urine (Tamm & Horsfall, 1952). It is expressed abundantly in the thick ascending limb of the loop of Henle and the early distal convoluted tubule (Kumar *et al.,* 1985). THP contains the most varied array of linked glycans of any human glycoprotein, which suggests a capacity for adhesion to a variety of ligands (Hard *et al.,* 1992). In the urine, THP may precipitate due to many factors of the immediate physico-chemical microenvironment and, accordingly, THP is the main constituent of hyaline urinary casts (Fairley, Owen & Birch, 1983). More recently, Fernández-Llama and colleagues (Fernandez-Llama *et al.,* 2010) demonstrated that abundant exosome vesicles are entrapped within urinary THP polymers. In order to overcome such an interference in the analysis of the total exosome proteome, a strong denaturation of its 24-disulfide bridges by DTT have been proposed (Gonzales *et al.*, 2009; Fernandez-Llama *et al.,* 2010). This leads to defolding of the zona pellucida (ZP) domain responsible for protein polymerization (Gonzales *et al.,* 2010). Alternatively, use of ultracentrifugation, incorporating a sucrose gradient or a cushion in deuterated buffer, to remove THP has been proposed (Mitchell *et al.,* 2009; Hogan *et al.,* 2009).

We introduce here distinct modifications, as summarised in Figure 1, to the well-established method of exosome isolation which is based on serial differential centrifugations (Gonzales *et al.,* 2010). Firstly, a very low-speed centrifugation at 1,000g was introduced to remove abundant cells, cell debris, nuclei, bacteria and the bulk of THP which interferes by entrapping the exosomes. As evident by SDS-PAGE analysis of this fraction, the approach was efficient (Figure 2.3). Although we cannot fully appreciate the fact that a fraction of vesicles retained at this point may have been entrapped within THP polymers, the introduction of this low-speed centrifugation step avoided a potential contamination resulting from intracellular vesicle release as a result of cytolysis following freeze-thaw cycles and hypotonic shock during dialysis. Furthermore, a dialysis step was introduced to decrease the salt concentration in order to favour the de-polymerization of THP and avoid its precipitation during the following centrifugation steps. Nonetheless, THP fibrils were still present in the low-speed centrifugation pellet, as shown in Figure 2.2, possibly due to the increased THP concentration following the reduction of the starting volume by sppedvac. This step was introduced with the double aim of handling a large amount of urine in a small volume and, secondly, to mimic a proteinuric condition and, thus, to evaluate the extent of albumin interference which is well-represented in the first crude pellets along with THP (section 2.3.3). To overcome these barriers, our protocol was modified particularly to facilitate the release of entrapped vesicles from THP polymers, and simultaneously removing the interference of soluble proteins like albumin.

After systematic trials, we selected two approaches to manage THP interference. The first involved the well-established DTT treatment (Pisitkun, Shen & Knepper, 2004; Gonzales *et al.,* 2009; Fernandez-Llama *et al.,* 2010; Gonzales *et al.,* 2010) which unfolds the zona pellucida (ZP) domain responsible for protein polymerization (Jovine *et al.,* 2002). The second method involved the addition of a mild detergent which has previously been shown to

solubilise THP efficiently (Kobayashi & Fukuoka, 2001). CHAPS, a non-denaturating zwitterionic detergent is ideally suited for the disruption of non-specific protein interactions, while also protecting the conformation of the protein (s) of interest (Hjelmeland, 1980). The bulk of THP and albumin were successfully removed from the exosomal pellet and after the treatment with CHAPS, were found in the supernatant, as shown in Figure 3. CHAPS is known to break protein-protein interactions (Hjelmeland, 1980; Labeta, Fernandez & Festenstein, 1988). However, depending on the microenvironment, the strongest protein-protein interactions, such as those found in the tetraspanin web (Levy & Shoham, 2005), are preserved.

When analysed by electron microscopy, pellets retained after both low and high-speed centrifugations treated with CHAPS and/or DTT showed a characteristic spherical morphology limited by a bilipidic layer (Figure 2.2) and with diameters ranging from between 30 to 150 nm. These characteristics are consistent with the observed size and morphology of vesicles previously described in urine (Pisitkun, Shen & Knepper, 2004; Cheruvanky *et al.,* 2007; Gonzales *et al.,* 2009; Hogan *et al.,* 2009; Hara *et al.,* 2010; Gonzales *et al.,* 2010; Merchant *et al.,* 2010; Rood *et al.,* 2010). Indeed, it is clear that urinary vesicles represent a tremendous heterogeneity in size and in the exosomes, marker patterns are evident as shown in the Western blot analysis. The concomitant presence of other specific markers, like nephrin and lactadherin, complicate the semantic definition of such urinary fractions which can include what are generally called microparticles, ectosomes, exosomes, 'exosome-like' vesicles and shed vesicles. Furthermore, these results clearly suggest that the whole vesicle is a structure resistant to treatment with a detergent such as CHAPS, which in turn demonstrates that this may be used for vesicle isolation without having a deleterious effect on the structural conformation of vesicles.

In our analysis, after an ON incubation in 1% (w/v) CHAPS, which is able to solubilise a detergent-resistant domain almost completely within one hour of incubation (Garner, Smith & Hooper, 2008), we were able to perform comparative analysis of CHAPS- and DTT-treated vesicle populations by TEM (Figure 2.2). This suggests that urinary vesicles are detergent (CHAPS)-resistant. Furthermore, the MS-based proteomic profiling led to the identification of 247 unique proteins in CHAPS SN and 274 in DTT SN, respectively. This finding indirectly confirms the resistance of these vesicles to detergent solubilisation. The Venn diagram in Figure 2.6 shows that around 75% of identifications in CHAPS or DTT are shared. This evidence clearly highlights the detergent-resistant features of urinary nano vesicles. In fact, if CHAPS had lysed and solubilised the vesicle contents, it would be expected to find a protein data set which would be highly biased toward the exosomal proteome. Furthermore, the distribution of the protein classes by the Panther classification system (Fig. 2.7) was similar between CHAPS and DTT in respect to several exosomal proteins. Comparison of our results with the published proteomic data on exosomes and exosome free supernatants (Gonzales *et al.*, 2009; Kentsis *et al.*, 2009) showed that more than 50% -exosomes pellet- and 70% -SN exosome-free of the identified proteins were shared, respectively. Interestingly, around 40% of identified proteins are common to the 4 data sets.

In spite of differences in the methodological approach, the 'off-gel' utilized in this study with respect to the 'in-gel' method utilized in the comparative protein data set (Gonzales *et al.*, 2009; Kentsis *et al.*, 2009), instrumentation and bioinformatics tools, the 3 data sets showed high concordance in the identifications. One of the reasons for it oculd be the stringency of MS identification criteria used (described in detail in the original publications on each dataset), was reasonably high in each case. Therefore, false-positive identifications should represent less than 1% of the entries. Keeping in mind that these vesicles were found to be resistant to harsh treatments, this evidence shows an incomplete recovery of vesicles by the

classical ultracentrifugation protocols which may lead to a preferential enrichment of vesicular subpopulations which plainly share a common set of proteins (exosomal markers) along with some specific ones (inherent to a possible sub-population). This is of significant practical importance as the increased yield of vesicles isolated and the confidence in the MS identification and quantification of low-abundant proteins from vesicle pellets could be of crucial importance to identify biomarkers represented in very low amounts. Currently, modern proteomics aim for the analysis of sub-cellular proteomes for markers discovery. In this regard, two special kinds of extracellular vesicles such as exosomes and membrane plasma shedding vesicles are emerging as excellent biological sources to be applied in the discovery of non-invasive organ-specific disease biomarkers.

Protease activity of dipeptidyl peptidase (DPP IV) and nephrilysin (NEP) showed that the CHAPS method is a useful alternative to DTT when full preservation of biological activity of all the disulfide bridge-containing proteins is needed. In particular, the NEP structure consists of a short N-terminal cytoplasmic domain, followed by a single trans-membrane helix, and a large C-terminal extracellular domain that contains the active site (Lee *et al.*, 1991; Shimada, Takahashi & Tanzawa, 1994). The extracellular domain of NEP contains 12 cysteine residues, all components of six disulfide bridges. Four of these are located within the catalytic domain. A single disulfide bridge is found within the inter-domain linker fragments, and another one is present within domain 2. All of these participate in maintaining the structure consisting of two multiply-connected folding domains which embrace a large central cavity containing the active site (Emoto & Yanagisawa, 1995). Evaluation of NEP proteolytic activity after reduction and re-oxidation (DTT treatment and subsequent removal of DTT) of disulfide bridges highlights that this may result in a severe misfolding, leading to an impairment of NEP activity. NEP activity is better preserved in the CHAPS protocol as shown by our results.

Western blotting of CHAPS and DTT fractions revealed very similar patterns. Firstly, after both treatments, the bulk of the tetraspanin family exosomal marker associated with late multivesicular endosomes, CD63 (Kobayashi *et al.,* 2000) was recovered at a low centrifugal speed (18,000g) with a broad distribution of isoforms. During urine formation, the final product for voiding contains vesicles originating from a wide range of different cellular origins and hence, it was not a surprise to find such a heterogeneity which is a reflection of specific cell-sorting and trafficking pathways, but also of distinct origin upstream in the podocyte by finding fractions of nephrin in exosomes. Finally, although protease degradation cannot be excluded, at the time of its occurrence this activity seemed to be very specific and related to all those fractions which were retained at a low-speed centrifugation (18,000g). In support of this evidence, immunodetection of cytosolic TSG-101 showed a fragmentation pattern in only one fraction **P18CP200,** which was not seen in other fractions where the signal was equally intense for full length protein (Figure 2.4). Furthermore, the ability to detect the intra-vesicular localised marker TSG101 (Babst *et al.,* 2000; Thery *et al.,* 2001; Khatua *et al.,* 2009) underlines the retained integrity of the vesicles during the whole purification methodology.

Lactadherin (MFG-E8 or SED1) is known to participate in a wide variety of cellular interactions (Raymond, Ensslin & Shur, 2009) and was previously shown to be released in association with exosomes (Oshima *et al.,* 2002). An interaction was reported between the discoidin/C domains and the pellucida zone (Ensslin & Shur, 2003) and one could speculate about a possible interaction with THP (ZP domain). Finally, lactadherin isoforms have been previously identified which are reflections of physiological state and distinct cell type origins (Veron *et al.,* 2005; Watanabe *et al.,* 2005). MFG-E8 immunodetection (Figure 2.4) reveals a wide distribution of bands with specific signatures for each pellet. Proteomic profiling (Pisitkun, Shen & Knepper, 2004; Gonzales *et al.,* 2009) has identified potential interacting

partners of MFGE-E8 (such as β-integrin) in exosomes. On the other hand, shed vesicles and apoptotic bodies have phosphatidyl-serine on the surface and MFG-E8 could mediate interactions between these vesicles and exosomes, and as well as THP. However, this remains to be fully established.

Surprisingly, immunodetection with anti-nephrin antibody revealed a comparable distribution of fragments to those seen in nephrinuria patients (Patari *et al.,* 2003). Examination of the western blot molecular weights and the intracellular location of the specific epitope recognised by our antibody led us to propose that this vesicle preparation contains membrane-associated nephrin fragments. Whether these fragments were orientated 'right-side-out' (extracellular part on surface) or 'inside-out' (cytosolic part on surface) has to be confirmed by further analyses. Identification of shed podocytes in the urines of healthy subjects and patients affected by glomerular diseases has also been proposed previously (Vogelmann *et al.,* 2003; Hara *et al.,* 2005). More recently, it was reported that small podocalyxin-positive vesicles which are negative for exosomal markers originate from microvilli of podocytes in both healthy samples and those from patients with glomerular diseases (Hara *et al.,* 2010). Detection of nephrin and exosomal markers in the same fraction may suggest a protein turnover through clathrin- or raft-mediated endocytosis (Veron *et al.,* 2005). After internalisation by either mechanism, vesicles are sorted into the endosomal pathway which could redirect vesicles to be either degraded or secreted as exosomes. The minute amount of nephrin excreted into urine of healthy subjects could reflect a physiological turnover of the slit diaphragm. In support to this, nephrin has been detected by immuno-electron microscopy not solely at the slit diaphragm area but also along lateral podocyte membranes and, eventually, in the urine (Patari *et al.,* 2003). The detection of the same nephrin fragmentation pattern in the whole urine of diabetic nephropathic patients (Patari *et al.,* 2003) suggests that an increase in this dynamic process may accurately reflect worsening

podocyte injury. A better understanding of the mechanism (s) which accelerates the turnover of this key structure of the glomerular filtration barrier may provide a critical new insight into the identification of early markers of glomerular impairment during disease progression.

It was found that despite the efforts to exclude high-abundant proteins like THP and albumin from nano vesicle isolation, these proteins are present in significant quantities in final pellet (Musante *et al.,* 2012). DTT and CHAPS treatment seem to get rid of these proteins but only to a certain extent. Therefore, the resilient behaviour of these proteins led us to speculate whether these proteins could be interacting specifically with the vesicle populations. We sought to establish this interaction using FLISA, IP and SPR. Preliminary FLISA and IP analysis was performed using CHAPS-treated nano vesicle populations only, and the monomeric fraction of commercial human albumin, as purified by SEC and analysed by SDS-PAGE (Figures 2.9 and 2.12), which demonstrated more favourable binding to albumin at a physiological pH. The demonstration of this biomolecular interaction was one of the primary objectives of these analyses. This FLISA and IP study was not repeated for DTT-populations as a significant amount of effort was required for the isolation of nano vesicles and hence, we elected to conserve this material for SPR analysis, using the FLISA and IP as a preliminary study to ascertain whether there was a justification for proceeding.

For our SPR studies, we elected to capture nano vesicles using a carboxymethylated dextran surface and EDC/NHS-based coupling through the selection of a CM5 sensor chip. It is well established that L1 sensor chips, supplied by the manufacturer of Biacore technology (GE Healthcare), are particularly suited for the capture of lipid-containing ligands (Anderluh *et al.,* 2005). However, our preconcentration and immobilisation studies performed initially on the CM5 surface allowed us to capture suitable quantities of analyte for interaction studies, with nano vesicles retaining stability on immobilised surfaces when subjected to regeneration scouting with NaOH. We propose that the CHAPS- and DTT-treated vesicles selected as

ligands were coupled to this surface via their associated proteins, which in turn permitted covalent capture. It should also be noted that at the completion of this analysis, only a 4.3% loss in captured ligand was observed, which suggested stability. For the purpose of our studies, we elected to proceed with comparative binding of a panel of different analytes with both nano vesicle populations, which explains why these were immobilised as opposed to e.g. albumin or THP. While identical ligand concentrations (100µg/mL) were selected for capture using the same immobilisation and capping cycle times, different amounts of ligand were captured on flow cells 2 and 4, respectively. It is a challenge to have identical amounts of ligand captured on adjacent surfaces. However, the objective of our analysis was simply to demonstrate qualitative binding primarily using a dose-dependent response curve of analyte. Hence, as the same amount of analyte was passing over both surfaces, it was possible to determine relative comparative binding.

For our studies, the immobilisation of nano vesicles previously shown to be comprised of a population of entities with differing sizes (Musante *et al.*, 2012) presented a heterologous surface, which would prevent accurate 1:1 interaction analyses to be performed. In addition to this, our objectives were not to obtain comparative affinity measurements between each analyte selected for analysis and individual nano vesicle populations, but simply to demonstrate the occurrence in nature of this biomolecular interaction and deduce the biological importance of this interaction. Furthermore, if an analyte such as albumin was selected for capture, it would not have been possible to perform kinetic analysis due to the heterogenous nature of these populations, inferring that is would not be possible to determine the size of the analyte. This could only be performed by isolating individual populations of nano vesicles of a defined size, which is problematic given the nature of these analytical matrices. Hence, no kinetic analysis was performed.

Preliminary SPR studies were undertaken to demonstrate an interaction between each nano vesicle population and albumin. More specifically, we initially chose to investigate whether the nano vesicles immobilised were already coupled with albumin by selecting an anti-albumin antibody for biorecognition purposes. For CHAPS-treated exosomes, the observation of a concentration-dependent increase in RU, taking online reference subtraction and the HBS negative control into consideration, suggested that albumin was already present and coupled with these vesicles. In contrast, the introduction of a low analyte concentration (10µg/mL) to DTT nano vesicles resulted in favourable binding (µ=76.2RU), albeit without an increase in binding as elevated amounts of anti-albumin antibody were subsequently introduced (Figure 2.10 A). This suggested that saturation was occurring at this concentration. In contrast, when SEC-purified monomeric albumin was tested (Figure 2.10 B), more favourable binding was observed for CHAPS populations, with signal saturation seen for both ligands at an analyte concentration of 50µg/mL. Based on the observations of these two interday SPR assays, we postulate that there is an interaction between albumin and both nano vesicle populations independent of one another, which is evident at the time of isolation from urine. Furthermore, we can speculate that there is a higher amount of albumin coupled to the DTT-population, which in turn results in increased binding of anti-albumin antibody at the lowest analyte concentration selected for analysis, namely 10µg/mL. Notwithstanding the tendency of human serum albumin to aggregate (Lin *et al.,* 2000), this in theory would mean that there is limited space available for albumin to bind to these vesicles, which is reflected by lower binding and, consequently, reduced RU values. The opposite effect appears to be evident for CHAPS-treated nano vesicles, where significantly increased albumin binding is observed, presumably to the presence of lower amounts on the immobilised vesicles which, in turn, accommodates interaction with free albumin in solution. It should also be mentioned that there is approximately twice as much ligand captured on

FC2 (DTT) than on the corresponding CHAPS surface (FC4). This may, in theory, mean a twofold increase in albumin binding sites availability on this surface which could mean that actual albumin interaction with DTT vesicles was even lower than what was observed. Finally, it must also be acknowledged that the detergent treatments may have different effects on the conformational structure of surface proteins/epitopes which may in turn compromise binding. IP, using anti-albumin antibodies, were able to isolate nano vesicles as evidenced by detection of CD63 in the elution fraction of IP (Figure 2.12). These results of SPR and IP seem to complement each other.

We and others have already outlined the fact that THP co-precipitation is problematic for urinary exosome isolation (Fernandez-Llama *et al.,* 2010; Musante *et al.,* 2012), although we present the first direct binding studies using DEF-purified THP. For CHAPS-treated nano vesicles, increased concentrations of this soluble protein correlated with an increase in binding, as seen previously for albumin. However, a contrasting binding profile was seen for DTT-populations, where analyte concentration was inversely proportional to surface binding. DTT-populations have less THP as compared to the CHAPS-populations (Figure 2.8). This could be due to the apparent Hooks effect.

Subsequent profiling using anti-IgG revealed that both populations contained this immunoglobulin as an external epitope, as TEM analysis revealed that structural integrity was not compromised as a result of detergent treatment (Figure 2.2). A typical concentration-dependent increase in binding was observed, albeit with more favourable binding to CHAPS-treated nano vesicle populations (Figure 2.14). In summary, through the implementation of FLISA and IP, as preliminary methodologies, and SPR analysis, through the application of the Biacore 3000 analytical platform, we have been able to provide key information on the biomolecular interactions that take place between high-abundance proteins such as THP, albumin and IgG and nano vesicles.

In conclusion, an improved method was developed to exclude most of the interference of soluble proteins present in urinary vesicle isolation. CHAPS treatment appears superior in preserving the activity of constituent proteins of membrane vesicles while simultaneously removing interference of soluble proteins (THP, albumin). Thus, our method offers a new protocol to prepare urinary vesicles for –omics, analytical or functional studies. Moreover, interactions between high abundant proteins and membrane vesicles were established. This, in part, explains why these proteins are so difficult to remove from the nano vesicle populations when isolated from whole urine.

2.5 Reference

Adachi, J., Kumar, C., Zhang, Y., Olsen, J.V. & Mann, M. (2006), "The human urinary proteome contains more than 1500 proteins, including a large proportion of membrane proteins", *GENOME BIOLOGY*, vol. 7, no. 9.

Ageberg, M. & Lindmark, A. (2003), "Characterisation of the biosynthesis and processing of the neutrophil granule membrane protein CD63 in myeloid cells", *CLINICAL AND LABORATORY HAEMATOLOGY*, vol. 25, no. 5, pp. 297-306.

Ahola, H., Wang, S., Luimula, P., Solin, M., Holzman, L. & Holthofer, H. (1999), "Cloning and expression of the rat nephrin homolog", *AMERICAN JOURNAL OF PATHOLOGY*, vol. 155, no. 3, pp. 907-913.

Anderluh, G., Besenicar, M., Kladnik, A., Lakey, J. & Macek, P. (2005), "Properties of nonfused liposomes immobilized on an L1 Biacore chip and their permeabilization by a eukaryotic pore-forming toxin", *ANALYTICAL BIOCHEMISTRY*, vol. 344, no. 1, pp. 43-52.

Babst, M., Odorizzi, G., Estepa, E. & Emr, S. (2000), "Mammalian tumor susceptibility gene 101 (TSG101) and the yeast homologue, Vps23p, both function in late endosomal trafficking", *TRAFFIC*, vol. 1, no. 3, pp. 248-258.

Bajari, T., Strasser, V., Nimpf, J. & Schneider, W. (2005), "LDL receptor family: Isolation, production, and ligand binding analysis", *METHODS*, vol. 36, no. 2, pp. 109-116.

Busek, P., Malik, R. & Sedo, A. (2004), "Dipeptidyl peptidase IV activity and/or structure homologues (DASH) and their substrates in cancer", *INTERNATIONAL JOURNAL OF BIOCHEMISTRY \& CELL BIOLOGY*, vol. 36, no. 3, pp. 408-421.

Candiano, G., Bruschi, M., Musante, L., Santucci, L., Ghiggeri, G., Carnemolla, B., Orecchia, P., Zardi, L. & Righetti, P. (2004), "Blue silver: A very sensitive colloidal Coomassie G-250 staining for proteome analysis", *ELECTROPHORESIS*, vol. 25, no. 9, SI, pp. 1327-1333.

Cheruvanky, A., Zhou, H., Pisitkun, T., Kopp, J.B., Knepper, M.A., Yuen, P.S.T. & Star, R.A. (2007), "Rapid isolation of urinary exosomal biomarkers using a nanomembrane ultrafiltration concentrator", *AMERICAN JOURNAL OF PHYSIOLOGY-RENAL PHYSIOLOGY*, vol. 292, no. 5, pp. F1657-F1661.

Emoto, N. & Yanagisawa, M. (1995), "Endothelin-converting enzyme-2 is a membrane-bound, phosphoramidon-sensitive metalloprotease with acidic ph optimum", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 270, no. 25, pp. 15262-15268.

Engering, A., Kuhn, L., Fluitsma, D., Hoefsmit, E. & Pieters, J. (2003), "Differential post-translational modification of CD63 molecules during maturation of human dendritic cells", *EUROPEAN JOURNAL OF BIOCHEMISTRY*, vol. 270, no. 11, pp. 2412-2420.

Ensslin, M. & Shur, B. (2003), "Identification of mouse sperm SED1, a bimotif EGF repeat and discoidin-domain protein involved in sperm-egg binding", *CELL*, vol. 114, no. 4, pp. 405-417.

Erdos, E. & Skidgel, R. (1989), "Neutral endopeptidase 24.11 (enkephalinase) and related regulators of peptide-hormones", *FASEB JOURNAL*, vol. 3, no. 2, pp. 145-151.

Fairley, J., Owen, J. & Birch, D. (1983), "Protein-composition of urinary casts from healthy-subjects and patients with glomerulonephritis", *BRITISH MEDICAL JOURNAL*, vol. 287, no. 6408, pp. 1838-1840.

Fernandez-Llama, P., Khositseth, S., Gonzales, P.A., Star, R.A., Pisitkun, T. & Knepper, M.A. (2010), "Tamm-Horsfall protein and urinary exosome isolation", *JOURNAL OF HYPERTENSION*, vol. 28, no. A, pp. E164.

Florentin, D., Sassi, A. & Roques, B.P. (1984), "A highly sensitive fluorometric assay for enkephalinase□, a neutral metalloendopeptidase that releases tyrosine-glycine-glycine from enkephalins", *ANALYTICAL BIOCHEMISTRY*, vol. 141, no. 1, pp. 62.

Garner, A.E., Smith, D.A. & Hooper, N.M. (2008), "Visualization of detergent solubilization of membranes: Implications for the isolation of rafts", *BIOPHYSICAL JOURNAL*, vol. 94, no. 4, pp. 1326-1340.

Gonzales, P.A., Pisitkun, T., Hoffert, J.D., Tchapyjnikov, D., Star, R.A., Kleta, R., Wang, N.S. & Knepper, M.A. (2009), "Large-Scale Proteomics and Phosphoproteomics of Urinary Exosomes", *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY*, vol. 20, no. 2, pp. 363-379.

Gonzales, P.A., Zhou, H., Pisitkun, T., Wang, N.S., Star, R.A., Knepper, M.A. & Yuen, P.S.T. (2010), "Isolation and Purification of Exosomes in Urine", *METHODS IN MOLECULAR BIOLOGY*, vol. 641, pp. 89-99.

Hard, K., Van zadelhoff, G., Moonen, P., Kamerling, J.P. & Vliegenthart, J.F.G. (1992), "The Asn-linked carbohydrate chains of human Tamm-Horsfall glycoprotein of one male", *EUROPEAN JOURNAL OF BIOCHEMISTRY*, vol. 209, no. 3, pp. 895-915.

Hara, M., Yanagihara, T., Kihara, I., Higashi, K., Fujimoto, K. & Kajita, T. (2005), "Apical cell membranes are shed into urine from injured podocytes: A novel phenomenon of podocyte injury", *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY*, vol. 16, no. 2, pp. 408-416.

Hara, M., Yanagihara, T., Hirayama, Y., Ogasawara, S., Kurosawa, H., Sekine, S. & Kihara, I. (2010), "Podocyte membrane vesicles in urine originate from tip vesiculation of podocyte microvilli", *HUMAN PATHOLOGY*, vol. 41, no. 9, pp. 1265-1275.

Hjelmeland, L.M. (1980), "A nondenaturing zwitterionic detergent for membrane biochemistry: design and synthesis", *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, vol. 77, no. 11, pp. 6368-6370.

Hogan, M.C., Manganelli, L., Woollard, J.R., Masyuk, A.I., Masyuk, T.V., Tammachote, R., Huang, B.Q., Leontovich, A.A., Beito, T.G., Madden, B.J., Charlesworth, M.C., Torres, V.E., LaRusso, N.F., Harris, P.C. & Ward, C.J. (2009), "Characterization of PKD Protein-Positive Exosome-Like Vesicles", *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY*, vol. 20, no. 2, pp. 278-288.

Jovine, L., Qi, H., Williams, Z., Litscher, E. & Wassarman, P. (2002), "The ZP domain is a conserved module for polymerization of extracellular proteins", *NATURE CELL BIOLOGY*, vol. 4, no. 6, pp. 457-461.

Kentsis, A., Monigatti, F., Dorff, K., Campagne, F., Bachur, R. & Steen, H. (2009), "Urine proteomics for profiling of human disease using high accuracy mass spectrometry", *PROTEOMICS CLINICAL APPLICATIONS*, vol. 3, no. 9, pp. 1052-1061.

Khatua, A.K., Taylor, H.E., Hildreth, J.E.K. & Popik, W. January 15, (2009), "Exosomes Packaging APOBEC3G Confer Human Immunodeficiency Virus Resistance to Recipient Cells", *JOURNAL OF VIROLOGY*, vol. 83, no. 2, pp. 512-521.

Kobayashi, K. & Fukuoka, S. (2001), "Conditions for solubilization of Tamm-Horsfall protein/uromodulin in human urine and establishment of a sensitive and accurate enzyme-linked immunosorbent assay (ELISA) method", *ARCHIVES OF BIOCHEMISTRY AND BIOPHYSICS*, vol. 388, no. 1, pp. 113-120.

Kobayashi, T., Vischer, U., Rosnoblet, C., Lebrand, C., Lindsay, M., Parton, R., Kruithof, E. & Gruenberg, J. (2000), "The tetraspanin CD63/lamp3 cycles between endocytic and secretory compartments in human endothelial cells", *MOLECULAR BIOLOGY OF THE CELL*, vol. 11, no. 5, pp. 1829-1843.

Kumar, S., Jasani, B., Hunt, J., Moffat, D. & Asscher, A. (1985), "A system for accurate immunolocalization of Tamm-Horsfall protein in renal biopsies", *HISTOCHEMICAL JOURNAL*, vol. 17, no. 11, pp. 1251-1258.

Labeta, M., Fernandez, N. & Festenstein, H. (1988), "solubilization effect of Nonidet P-40, Triton X-100 and CHAPS in the detection of MHC-like glycoproteins", *JOURNAL OF IMMUNOLOGICAL METHODS*, vol. 112, no. 1, pp. 133-138.

Laemmli, U. (1970), "cleavage of structural proteins during assembly of head of bacteriophage-T4", *NATURE*, vol. 227, no. 5259, pp. 680.

Lee, S., Zambas, E., Marsh, W. & Redman, C. (1991), "Molecular-cloning and primary structure of kell blood-group protein", *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, vol. 88, no. 14, pp. 6353-6357.

Levy, S. & Shoham, T. (2005), "Protein-protein interactions in the tetraspanin web", *PHYSIOLOGY*, vol. 20, pp. 218-224.

Lin, J., Meyer, J., Carpenter, J. & Manning, M. (2000), "Stability of human serum albumin during bioprocessing: Denaturation and aggregation during processing of albumin paste", *PHARMACEUTICAL RESEARCH*, vol. 17, no. 4, pp. 391-396.

Lin, Y., Liu, Y., Li, J., Zhao, Y., He, Q., Han, W., Chen, P., Wang, X. & Liang, S. (2010), "Evaluation and optimization of removal of an acid-insoluble surfactant for shotgun analysis of membrane proteome", *ELECTROPHORESIS*, vol. 31, no. 16, pp. 2705-2713.

Lin, Y., Zhou, J., Bi, D., Chen, P., Wang, X. & Liang, S. (2008), "Sodium-deoxycholate-assisted tryptic digestion and identification of proteolytically resistant proteins", *ANALYTICAL BIOCHEMISTRY*, vol. 377, no. 2, pp. 259-266.

Liu, W.J. & Hansen, P.J. (1995), "Progesterone-induced secretion of dipeptidyl peptidase-IV (cluster differentiation antigen-26) by the uterine endometrium of the ewe and cow that costimulates lymphocyte proliferation.", *ENDOCRINOLOGY*, vol. 136, no. 2, pp. 779-787.

Logozzi, M., De Milito, A., Lugini, L., Borghi, M., Calabro, L., Spada, M., Perdicchio, M., Marino, M.L., Federici, C., Iessi, E., Brambilla, D., Venturi, G., Lozupone, F., Santinami, M., Huber, V., Maio, M., Rivoltini, L. & Fais, S. (2009), "High Levels of Exosomes Expressing CD63 and Caveolin-1 in Plasma of Melanoma Patients", *PLOS ONE*, vol. 4, no. 4, pp. e5219.

M., M. & Bradford (1976), "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding", *ANALYTICAL BIOCHEMISTRY*, vol. 72, no. 12, pp. 248.

Marimuthu, A., O Meally, R.N., Chaerkady, R., Subbannayya, Y., Nanjappa, V., Kumar, P., Kelkar, D.S., Pinto, S.M., Sharma, R., Renuse, S., Goel, R., Christopher, R., Delanghe, B., Cole, R.N., Harsha, H.C. & Pandey, A. (2011), "A Comprehensive Map of the Human Urinary Proteome", *JOURNAL OF PROTEOME RESEARCH*, vol. 10, no. 6, pp. 2734-2743.

Mathias, R.A., Lim, J.W., Ji, H. & Simpson, R.J. (2009), "Isolation of Extracellular Membranous Vesicles for Proteomic Analysis", vol. 528, pp. 227-242.

McQueen, E.G. & Engel, G.B. (1966), "Factors determining the aggregation of urinary mucoprotein", *JOURNAL OF CLINICAL PATHOLOGY*, vol. 19, no. 4, pp. 392-396.

Merchant, M.L., Powell, D.W., Wilkey, D.W., Cummins, T.D., Deegens, J.K., Rood, I.M., McAfee, K.J., Fleischer, C., Klein, E. & Klein, J.B. (2010), "Microfiltration isolation of human urinary exosomes for characterization by MS", *PROTEOMICS CLINICAL APPLICATIONS*, vol. 4, no. 1, pp. 84-96.

Miranda, K.C., Bond, D.T., McKee, M., Skog, J., Paunescu, T.G., Da Silva, N., Brown, D. & Russo, L.M. (2010), "Nucleic acids within urinary exosomes/microvesicles are potential biomarkers for renal disease", *KIDNEY INTERNATIONAL*, vol. 78, no. 2, pp. 191-199.

Mitchell, P.J., Welton, J., Staffurth, J., Court, J., Mason, M.D., Tabi, Z. & Clayton, A. (2009), "Can urinary exosomes act as treatment response markers in prostate cancer?", *JOURNAL OF TRANSLATIONAL MEDICINE*, vol. 7, pp. 4.

Musante, L., Saraswat, M., Duriez, E., Byrne, B., Ravida, A., Domon, B. & Holthofer, H. (2012), "Biochemical and Physical Characterisation of Urinary Nanovesicles following CHAPS Treatment", *PLOS ONE*, vol. 7, no. 7, pp. e37279.

Nilsson, J., Skog, J., Nordstrand, A., Baranov, V., Mincheva-Nilsson, L., Breakefield, X.O. & Widmark, A. (2009), "Prostate cancer-derived urine exosomes: a novel approach to biomarkers for prostate cancer", *BRITISH JOURNAL OF CANCER*, vol. 100, no. 10, pp. 1603-1607.

Oshima, K., Aoki, N., Kato, T., Kitajima, K. & Matsuda, T. (2002), "Secretion of a peripheral membrane protein, MFG-E8, as a complex with membrane vesicles - A possible role in membrane secretion", *EUROPEAN JOURNAL OF BIOCHEMISTRY*, vol. 269, no. 4, pp. 1209-1218.

Patari, A., Forsblom, C., Havana, M., Taipale, H., Groop, P., Holthofer, H. & FinnDiane Study Grp (2003), "Nephrinuria in diabetic nephropathy of type 1 diabetes", *DIABETES*, vol. 52, no. 12, pp. 2969-2974.

Pisitkun, T., Shen, R. & Knepper, M. (2004), "Identification and proteomic profiling of exosomes in human urine", *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, vol. 101, no. 36, pp. 13368-13373.

Raymond, A., Ensslin, M.A. & Shur, B.D. (2009), "SED1/MFG-E8: A Bi-Motif Protein That Orchestrates Diverse Cellular Interactions", *JOURNAL OF CELLULAR BIOCHEMISTRY*, vol. 106, no. 6, pp. 957-966.

Reeves, E.P., Williamson, M., Byrne, B., Bergin, D.A., Smiith, S.G.J., Greally, P., O'Kennedy, R., O'Neill, S.J. & McElvaney, N.G. (2010), "IL-8 Dictates Glycosaminoglycan Binding and Stability of IL-18 in Cystic Fibrosis", *JOURNAL OF IMMUNOLOGY*, vol. 184, no. 3, pp. 1642-1652.

Rood, I.M., Deegens, J.K.J., Merchant, M.L., Tamboer, W.P.M., Wilkey, D.W., Wetzels, J.F.M. & Klein, J.B. (2010), "Comparison of three methods for isolation of urinary microvesicles to identify biomarkers of nephrotic syndrome", *KIDNEY INTERNATIONAL*, vol. 78, no. 8, pp. 810-816.

Serafinicessi, F., Bellabarba, G., Malagolini, N. & Dallolio, F. 1989, "Rapid isolation of Tamm-Horsfall glycoprotein (uromodulin) from human-urine", *JOURNAL OF IMMUNOLOGICAL METHODS*, vol. 120, no. 2, pp. 185-189.

Shimada, K., Takahashi, M. & Tanzawa, K. (1994), "Cloning and functional expression of endothelin-converting enzyme from rat endothelial-cells", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 269, no. 28, pp. 18275-18278.

Simpson, R.J., Lim, J.W.E., Moritz, R.L. & Mathivanan, S. (2009), "Exosomes: proteomic insights and diagnostic potential", *EXPERT REVIEW OF PROTEOMICS*, vol. 6, no. 3, pp. 267-283.

Tamm, I. & Horsfall, F. (1952), "A mucoprotein derived from human urine which reacts with influenza, mumps, and newcastle disease viruses", *JOURNAL OF EXPERIMENTAL MEDICINE*, vol. 95, no. 1, pp. 71-97.

Thery, C., Boussac, M., Veron, P., Ricciardi-Castagnoli, P., Raposo, G., Garin, J. & Amigorena, S. (2001), "Proteomic analysis of dendritic cell-derived exosomes: A secreted subcellular compartment distinct from apoptotic vesicles", *JOURNAL OF IMMUNOLOGY*, vol. 166, no. 12, pp. 7309-7318.

Thomas, P., Campbell, M., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., Diemer, K., Muruganujan, A. & Narechania, A. (2003), "PANTHER: A library of protein families and subfamilies indexed by function", *GENOME RESEARCH*, vol. 13, no. 9, pp. 2129-2141.

Towbin, H., Staehelin, T. & Gordon, J. (1979), "Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets - procedure and some applications", *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, vol. 76, no. 9, pp. 4350-4354.

Van Niel, G., Porto-Carreiro, I., Simoes, S. & Raposo, G. (2006), "Exosomes: A common pathway for a specialized function", *JOURNAL OF BIOCHEMISTRY*, vol. 140, no. 1, pp. 13-21.

Veron, P., Segura, E., Sugano, G., Amigorena, S. & Thery, C. (2005), "Accumulation of MFG-E8/lactadherin on exosomes from immature dendritic cells", *BLOOD CELLS MOLECULES AND DISEASES*, vol. 35, no. 2, pp. 81-88.

Vogelmann, S., Nelson, W., Myers, B. & Lemley, K. (2003), "Urinary excretion of viable podocytes in health and renal disease", *AMERICAN JOURNAL OF PHYSIOLOGY-RENAL PHYSIOLOGY*, vol. 285, no. 1, pp. F40-F48.

Watanabe, T., Totsuka, R., Miyatani, S., Kurata, S., Sato, S., Katoh, I., Kobayashi, S. & Ikawa, Y. (2005), "Production of the long and short forms of MFG-E8 by epidermal keratinocytes", *CELL AND TISSUE RESEARCH*, vol. 321, no. 2, pp. 185-193.

Welton, J.L., Khanna, S., Giles, P.J., Brennan, P., Brewis, I.A., Staffurth, J., Mason, M.D. & Clayton, A. (2010), "Proteomic analysis of bladder cancer exosomes", *MOLECULAR & CELLULAR PROTEOMICS*, vol. 9, no. 6, pp. 1324-1338.

Wessel, D. & Flugge, U. 1984, "A method for the quantitative recovery of protein in dilute-solution in the presence of detergents and lipids", *ANALYTICAL BIOCHEMISTRY*, vol. 138, no. 1, pp. 141-143.

Westergaard, U., Sorensen, E., Hermey, G., Nielsen, M., Nykjaer, A., Kirkegaard, K., Jacobsen, C., Gliemann, J., Madsen, P. & Petersen, C. (2004), "Functional organization of the sortilin Vps10p domain", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 279, no. 48, pp. 50221-50229.

Zhou, H., Cheruvanky, A., Hu, X., Matsumoto, T., Hiramatsu, N., Cho, M.E., Berger, A., Leelahavanichkul, A., Doi, K., Chawla, L.S., Illei, G.G., Kopp, J.B., Balow, J.E., Austin, I.,Howard A., Yuen, P.S.T. & Star, R.A. (2008), "Urinary exosomal transcription factors, a new class of biomarkers for renal disease", *KIDNEY INTERNATIONAL*, vol. 74, no. 5, pp. 613-621.

Supplementary table S2.1: Identification of proteins in (P18+P200) DTT SN200 sample 1. Accession of the protein, sequence coverage, PSM (Match between a fragmentation mass spectrum and peptide), number of peptides found, number of amino acid in the protein, molecular weight, calculated PI and mascot score along with the description of the protein are indicated in the table.

Accession	Coverage %	# PSMs	# Peptides	# AAs	MW [kDa]	calc. pI	Score	Description
P62258	12.16	2	2	255	29.2	4.74	118.92	14-3-3 protein epsilon OS=Homo sapiens GN=YWHAE PE=1 SV=1 - [1433E_HUMAN]
P63104	4.90	1	1	245	27.7	4.79	65.24	14-3-3 protein zeta/delta OS=Homo sapiens GN=YWHAZ PE=1 SV=1 - [1433Z_HUMAN]
P08195	4.44	2	2	630	68.0	5.01	76.24	4F2 cell-surface antigen heavy chain OS=Homo sapiens GN=SLC3A2 PE=1 SV=3 - [4F2_HUMAN]
O95336	6.20	1	1	258	27.5	6.05	98.20	6-phosphogluconolactonase OS=Homo sapiens GN=PGLS PE=1 SV=2 - [6PGL_HUMAN]
Q13510	26.08	9	8	395	44.6	7.62	555.66	Acid ceramidase OS=Homo sapiens GN=ASAH1 PE=1 SV=5 - [ASAH1_HUMAN]
P60709	9.60	3	3	375	41.7	5.48	191.73	Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1 - [ACTB_HUMAN]
P13798	1.64	1	1	732	81.2	5.48	68.22	Acylamino-acid-releasing enzyme OS=Homo sapiens GN=APEH PE=1 SV=4 - [ACPH_HUMAN]
P02763	30.35	5	5	201	23.5	5.02	279.23	Alpha-1-acid glycoprotein 1 OS=Homo sapiens GN=ORM1 PE=1 SV=1 - [A1AG1_HUMAN]
P19652	16.92	3	3	201	23.6	5.11	131.95	Alpha-1-acid glycoprotein 2 OS=Homo sapiens GN=ORM2 PE=1 SV=2 - [A1AG2_HUMAN]
P01011	17.02	5	5	423	47.6	5.52	326.04	Alpha-1-antichymotrypsin OS=Homo sapiens GN=SERPINA3 PE=1 SV=2 - [AACT_HUMAN]
P01009	37.80	14	12	418	46.7	5.59	838.83	Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3 - [A1AT_HUMAN]
P04217	5.86	2	2	495	54.2	5.87	109.18	Alpha-1B-glycoprotein OS=Homo sapiens GN=A1BG PE=1 SV=3 - [A1BG_HUMAN]
P02765	5.72	3	2	367	39.3	5.72	179.74	Alpha-2-HS-glycoprotein OS=Homo sapiens GN=AHSG PE=1 SV=1 - [FETUA_HUMAN]
P01023	1.97	2	2	1474	163.2	6.42	115.80	Alpha-2-macroglobulin OS=Homo sapiens GN=A2M PE=1 SV=2 - [A2MG_HUMAN]
O43707	5.93	5	5	911	104.8	5.44	274.72	Alpha-actinin-4 OS=Homo sapiens GN=ACTN4 PE=1 SV=2 - [ACTN4_HUMAN]
P04745	37.77	12	12	511	57.7	6.93	628.87	Alpha-amylase 1 OS=Homo sapiens GN=AMY1A PE=1 SV=2 - [AMY1_HUMAN]
P19961	31.51	11	11	511	57.7	7.09	601.37	Alpha-amylase 2B OS=Homo sapiens GN=AMY2B PE=1 SV=1 - [AMY2B_HUMAN]
P06733	10.60	3	3	434	47.1	7.39	151.89	Alpha-enolase OS=Homo sapiens

								GN=ENO1 PE=1 SV=2 - [ENOA_HUMAN]
P54802	32.17	19	16	743	82.1	6.54	1058.50	Alpha-N-acetylglucosaminidase OS=Homo sapiens GN=NAGLU PE=1 SV=1 - [ANAG_HUMAN]
P15144	29.16	30	21	967	109.5	5.48	1641.03	Aminopeptidase N OS=Homo sapiens GN=ANPEP PE=1 SV=4 - [AMPN_HUMAN]
P12821	1.07	1	1	1306	149.6	6.39	63.12	Angiotensin-converting enzyme OS=Homo sapiens GN=ACE PE=1 SV=1 - [ACE_HUMAN]
P01019	6.39	3	2	485	53.1	6.32	163.68	Angiotensinogen OS=Homo sapiens GN=AGT PE=1 SV=1 - [ANGT_HUMAN]
P04083	8.09	2	2	346	38.7	7.02	70.17	Annexin A1 OS=Homo sapiens GN=ANXA1 PE=1 SV=2 - [ANXA1_HUMAN]
P07355	4.13	1	1	339	38.6	7.75	95.63	Annexin A2 OS=Homo sapiens GN=ANXA2 PE=1 SV=2 - [ANXA2_HUMAN]
P08758	9.69	2	2	320	35.9	5.05	118.17	Annexin A5 OS=Homo sapiens GN=ANXA5 PE=1 SV=2 - [ANXA5_HUMAN]
P01008	2.80	1	1	464	52.6	6.71	72.00	Antithrombin-III OS=Homo sapiens GN=SERPINC1 PE=1 SV=1 - [ANT3_HUMAN]
P02647	23.97	5	5	267	30.8	5.76	312.30	Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1 - [APOA1_HUMAN]
P02652	20.00	1	1	100	11.2	6.62	91.59	Apolipoprotein A-II OS=Homo sapiens GN=APOA2 PE=1 SV=1 - [APOA2_HUMAN]
P06727	6.31	2	2	396	45.4	5.38	134.56	Apolipoprotein A-IV OS=Homo sapiens GN=APOA4 PE=1 SV=3 - [APOA4_HUMAN]
P05090	35.98	16	8	189	21.3	5.15	831.21	Apolipoprotein D OS=Homo sapiens GN=APOD PE=1 SV=1 - [APOD_HUMAN]
P02649	23.34	6	5	317	36.1	5.73	353.72	Apolipoprotein E OS=Homo sapiens GN=APOE PE=1 SV=1 - [APOE_HUMAN]
P15289	14.00	5	4	507	53.6	6.07	293.32	Arylsulfatase A OS=Homo sapiens GN=ARSA PE=1 SV=3 - [ARSA_HUMAN]
O75882	7.42	10	10	1429	158.4	7.31	500.96	Attractin OS=Homo sapiens GN=ATRN PE=1 SV=2 - [ATRN_HUMAN]
P98160	1.34	5	5	4391	468.5	6.51	241.11	Basement membrane-specific heparan sulfate proteoglycan core protein OS=Homo sapiens GN=HSPG2 PE=1 SV=3 - [PGBM_HUMAN]
P02749	4.35	1	1	345	38.3	7.97	77.13	Beta-2-glycoprotein 1 OS=Homo sapiens GN=APOH PE=1 SV=3 - [APOH_HUMAN]
P61769	8.40	1	1	119	13.7	6.52	75.04	Beta-2-microglobulin OS=Homo sapiens GN=B2M PE=1 SV=1 - [B2MG_HUMAN]
P16278	7.98	4	4	677	76.0	6.57	287.86	Beta-galactosidase OS=Homo sapiens GN=GLB1 PE=1 SV=2 - [BGAL_HUMAN]
P08236	11.21	4	4	651	74.7	7.02	279.58	Beta-glucuronidase OS=Homo sapiens GN=GUSB PE=1 SV=2 - [BGLR_HUMAN]
Q93088	14.53	4	4	406	45.0	7.03	297.42	Betaine--homocysteine S-methyltransferase 1 OS=Homo sapiens GN=BHMT PE=1 SV=2 - [BHMT1_HUMAN]
P19835	14.61	7	7	753	79.3	5.34	339.62	Bile salt-activated lipase OS=Homo sapiens GN=CEL PE=1 SV=3 - [CEL_HUMAN]

P43251	4.79	2	2	543	61.1	6.25	140.86	Biotinidase OS=Homo sapiens GN=BTD PE=1 SV=2 - [BTD_HUMAN]
Q5VW32	3.89	2	1	411	46.4	7.65	80.70	BRO1 domain-containing protein BROX OS=Homo sapiens GN=BROX PE=1 SV=1 - [BROX_HUMAN]
Q8WVV5	2.10	1	1	523	59.0	6.01	66.10	Butyrophilin subfamily 2 member A2 OS=Homo sapiens GN=BTN2A2 PE=2 SV=2 - [BT2A2_HUMAN]
P12830	7.94	4	4	882	97.4	4.73	235.86	Cadherin-1 OS=Homo sapiens GN=CDH1 PE=1 SV=3 - [CADH1_HUMAN]
P55290	1.68	1	1	713	78.2	4.98	67.16	Cadherin-13 OS=Homo sapiens GN=CDH13 PE=1 SV=1 - [CAD13_HUMAN]
P19022	4.08	2	2	906	99.7	4.81	196.12	Cadherin-2 OS=Homo sapiens GN=CDH2 PE=1 SV=4 - [CADH2_HUMAN]
Q9BYE9	3.66	4	4	1310	141.5	4.50	183.49	Cadherin-related family member 2 OS=Homo sapiens GN=CDHR2 PE=1 SV=2 - [CDHR2_HUMAN]
Q9HBB8	2.13	1	1	845	88.2	4.93	76.63	Cadherin-related family member 5 OS=Homo sapiens GN=CDHR5 PE=1 SV=3 - [CDHR5_HUMAN]
P05937	4.98	1	1	261	30.0	4.83	78.51	Calbindin OS=Homo sapiens GN=CALB1 PE=1 SV=2 - [CALB1_HUMAN]
P22792	3.30	1	1	545	60.6	5.99	128.50	Carboxypeptidase N subunit 2 OS=Homo sapiens GN=CPN2 PE=1 SV=2 - [CPN2_HUMAN]
P31944	5.79	1	1	242	27.7	5.58	64.58	Caspase-14 OS=Homo sapiens GN=CASP14 PE=1 SV=2 - [CASPE_HUMAN]
P07339	15.53	4	4	412	44.5	6.54	301.14	Cathepsin D OS=Homo sapiens GN=CTSD PE=1 SV=1 - [CATD_HUMAN]
P16070	1.62	1	1	742	81.5	5.33	95.82	CD44 antigen OS=Homo sapiens GN=CD44 PE=1 SV=2 - [CD44_HUMAN]
P13987	15.63	3	2	128	14.2	6.48	176.85	CD59 glycoprotein OS=Homo sapiens GN=CD59 PE=1 SV=1 - [CD59_HUMAN]
Q8NFZ8	6.19	2	2	388	42.8	6.30	128.19	Cell adhesion molecule 4 OS=Homo sapiens GN=CADM4 PE=1 SV=1 - [CADM4_HUMAN]
P00450	13.90	10	10	1065	122.1	5.72	570.61	Ceruloplasmin OS=Homo sapiens GN=CP PE=1 SV=1 - [CERU_HUMAN]
Q53GD3	1.97	1	1	710	79.2	8.59	99.06	Choline transporter-like protein 4 OS=Homo sapiens GN=SLC44A4 PE=2 SV=1 - [CTL4_HUMAN]
P10909	19.60	6	6	449	52.5	6.27	452.74	Clusterin OS=Homo sapiens GN=CLU PE=1 SV=1 - [CLUS_HUMAN]
Q9UGN4	6.02	1	1	299	33.2	5.49	63.42	CMRF35-like molecule 8 OS=Homo sapiens GN=CD300A PE=1 SV=2 - [CLM8_HUMAN]
P12109	5.45	4	4	1028	108.5	5.43	330.05	Collagen alpha-1(VI) chain OS=Homo sapiens GN=COL6A1 PE=1 SV=3 - [CO6A1_HUMAN]
P39059	0.94	1	1	1388	141.6	5.00	133.46	Collagen alpha-1(XV) chain OS=Homo sapiens GN=COL15A1 PE=1 SV=2 - [COFA1_HUMAN]
P12111	1.51	3	3	3177	343.5	6.68	260.70	Collagen alpha-3(VI) chain OS=Homo sapiens GN=COL6A3 PE=1 SV=4 - [CO6A3_HUMAN]
P01024	1.68	2	2	1663	187.0	6.40	120.26	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2 - [CO3_HUMAN]

P0C0L4	1.49	2	2	1744	192.7	7.08	122.10	Complement C4-A OS=Homo sapiens GN=C4A PE=1 SV=1 - [CO4A_HUMAN]
P08185	3.95	1	1	405	45.1	6.04	75.37	Corticosteroid-binding globulin OS=Homo sapiens GN=SERPINA6 PE=1 SV=1 - [CBG_HUMAN]
P12277	4.46	1	1	381	42.6	5.59	101.91	Creatine kinase B-type OS=Homo sapiens GN=CKB PE=1 SV=1 - [KCRB_HUMAN]
Q86T13	3.47	1	1	490	51.6	6.35	58.87	C-type lectin domain family 14 member A OS=Homo sapiens GN=CLEC14A PE=1 SV=1 - [CLC14_HUMAN]
O60494	9.33	23	23	3623	398.4	5.35	1695.14	Cubilin OS=Homo sapiens GN=CUBN PE=1 SV=4 - [CUBN_HUMAN]
P15924	0.49	1	1	2871	331.6	6.81	69.25	Desmoplakin OS=Homo sapiens GN=DSP PE=1 SV=3 - [DESP_HUMAN]
P16444	11.19	3	3	411	45.6	6.15	189.46	Dipeptidase 1 OS=Homo sapiens GN=DPEP1 PE=1 SV=3 - [DPEP1_HUMAN]
P53634	9.72	3	3	463	51.8	6.99	234.33	Dipeptidyl peptidase 1 OS=Homo sapiens GN=CTSC PE=1 SV=1 - [CATC_HUMAN]
Q9UHL4	4.47	2	2	492	54.3	6.32	90.75	Dipeptidyl peptidase 2 OS=Homo sapiens GN=DPP7 PE=1 SV=3 - [DPP2_HUMAN]
P27487	13.45	9	8	766	88.2	6.04	625.03	Dipeptidyl peptidase 4 OS=Homo sapiens GN=DPP4 PE=1 SV=2 - [DPP4_HUMAN]
Q12805	14.00	5	5	493	54.6	5.07	292.78	EGF-containing fibulin-like extracellular matrix protein 1 OS=Homo sapiens GN=EFEMP1 PE=1 SV=2 - [FBLN3_HUMAN]
Q9UNN8	16.81	3	3	238	26.7	7.18	130.75	Endothelial protein C receptor OS=Homo sapiens GN=PROCR PE=1 SV=1 - [EPCR_HUMAN]
P08294	16.67	3	3	240	25.8	6.61	176.52	Extracellular superoxide dismutase [Cu-Zn] OS=Homo sapiens GN=SOD3 PE=1 SV=2 - [SODE_HUMAN]
P15311	7.34	5	4	586	69.4	6.27	228.20	Ezrin OS=Homo sapiens GN=EZR PE=1 SV=4 - [EZRI_HUMAN]
P02671	1.96	1	1	866	94.9	6.01	86.31	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2 - [FIBA_HUMAN]
Q14314	3.64	1	1	439	50.2	7.39	94.43	Fibroleukin OS=Homo sapiens GN=FGL2 PE=1 SV=1 - [FGL2_HUMAN]
P02751	4.86	8	7	2386	262.5	5.71	496.46	Fibronectin OS=Homo sapiens GN=FN1 PE=1 SV=4 - [FINC_HUMAN]
Q5D862	0.46	1	1	2391	247.9	8.31	62.00	Filaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1 - [FILA2_HUMAN]
Q14315	0.48	1	1	2725	290.8	5.97	74.24	Filamin-C OS=Homo sapiens GN=FLNC PE=1 SV=3 - [FLNC_HUMAN]
P09467	5.03	1	1	338	36.8	6.99	78.68	Fructose-1,6-bisphosphatase 1 OS=Homo sapiens GN=FBP1 PE=1 SV=4 - [F16P1_HUMAN]
P04075	1.92	2	1	364	39.4	8.09	69.67	Fructose-bisphosphate aldolase A OS=Homo sapiens GN=ALDOA PE=1 SV=2 - [ALDOA_HUMAN]
Q08380	29.74	26	12	585	65.3	5.27	1515.69	Galectin-3-binding protein OS=Homo sapiens GN=LGALS3BP PE=1 SV=1 - [LG3BP_HUMAN]
Q92820	16.35	4	4	318	35.9	7.11	280.60	Gamma-glutamyl hydrolase OS=Homo sapiens GN=GGH PE=1 SV=2 - [GGH_HUMAN]

P19440	6.85	3	3	569	61.4	7.12	228.32	Gamma-glutamyltranspeptidase 1 OS=Homo sapiens GN=GGT1 PE=1 SV=2 - [GGT1_HUMAN]
P17900	10.36	1	1	193	20.8	5.31	79.14	Ganglioside GM2 activator OS=Homo sapiens GN=GM2A PE=1 SV=4 - [SAP3_HUMAN]
P06396	9.72	5	5	782	85.6	6.28	384.59	Gelsolin OS=Homo sapiens GN=GSN PE=1 SV=1 - [GELS_HUMAN]
Q16769	26.04	6	6	361	40.9	6.61	419.14	Glutaminyl-peptide cyclotransferase OS=Homo sapiens GN=QPCT PE=1 SV=1 - [QPCT_HUMAN]
Q07075	5.85	4	4	957	109.2	5.47	216.96	Glutamyl aminopeptidase OS=Homo sapiens GN=ENPEP PE=1 SV=3 - [AMPE_HUMAN]
P22352	15.49	3	3	226	25.5	8.13	184.78	Glutathione peroxidase 3 OS=Homo sapiens GN=GPX3 PE=1 SV=2 - [GPX3_HUMAN]
P08263	9.01	2	2	222	25.6	8.88	88.98	Glutathione S-transferase A1 OS=Homo sapiens GN=GSTA1 PE=1 SV=3 - [GSTA1_HUMAN]
P09211	7.62	1	1	210	23.3	5.64	150.04	Glutathione S-transferase P OS=Homo sapiens GN=GSTP1 PE=1 SV=2 - [GSTP1_HUMAN]
P04406	16.42	3	3	335	36.0	8.46	176.18	Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens GN=GAPDH PE=1 SV=3 - [G3P_HUMAN]
P28799	2.53	2	1	593	63.5	6.83	151.06	Granulins OS=Homo sapiens GN=GRN PE=1 SV=2 - [GRN_HUMAN]
Q8NHV1	2.00	2	1	300	34.5	6.46	65.83	GTPase IMAP family member 7 OS=Homo sapiens GN=GIMAP7 PE=2 SV=1 - [GIMA7_HUMAN]
P00738	8.62	3	3	406	45.2	6.58	202.09	Haptoglobin OS=Homo sapiens GN=HP PE=1 SV=1 - [HPT_HUMAN]
P08107	1.25	1	1	641	70.0	5.66	87.09	Heat shock 70 kDa protein 1A/1B OS=Homo sapiens GN=HSPA1A PE=1 SV=5 - [HSP71_HUMAN]
P11142	1.70	1	1	646	70.9	5.52	83.24	Heat shock cognate 71 kDa protein OS=Homo sapiens GN=HSPA8 PE=1 SV=1 - [HSP7C_HUMAN]
P68871	6.80	1	1	147	16.0	7.28	71.11	Hemoglobin subunit beta OS=Homo sapiens GN=HBB PE=1 SV=2 - [HBB_HUMAN]
P02790	2.38	1	1	462	51.6	7.02	67.77	Hemopexin OS=Homo sapiens GN=HPX PE=1 SV=2 - [HEMO_HUMAN]
P04196	5.71	2	2	525	59.5	7.50	137.19	Histidine-rich glycoprotein OS=Homo sapiens GN=HRG PE=1 SV=1 - [HRG_HUMAN]
Q86YZ3	5.09	4	3	2850	282.2	10.04	203.97	Hornerin OS=Homo sapiens GN=HRNR PE=1 SV=2 - [HORN_HUMAN]
Q12794	8.74	2	2	435	48.3	6.77	136.32	Hyaluronidase-1 OS=Homo sapiens GN=HYAL1 PE=1 SV=2 - [HYAL1_HUMAN]
O75144	8.61	2	2	302	33.3	5.31	104.86	ICOS ligand OS=Homo sapiens GN=ICOSLG PE=1 SV=2 - [ICOSL_HUMAN]
P01876	24.36	8	6	353	37.6	6.51	527.15	Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=2 - [IGHA1_HUMAN]
P01877	28.82	7	6	340	36.5	6.10	384.08	Ig alpha-2 chain C region OS=Homo sapiens GN=IGHA2 PE=1 SV=3 - [IGHA2_HUMAN]
P01857	32.42	8	7	330	36.1	8.19	432.82	Ig gamma-1 chain C region OS=Homo sapiens GN=IGHG1 PE=1 SV=1 - [IGHG1_HUMAN]

P01859	7.98	2	2	326	35.9	7.59	204.22	Ig gamma-2 chain C region OS=Homo sapiens GN=IGHG2 PE=1 SV=2 - [IGHG2_HUMAN]
P01825	13.68	1	1	117	12.8	7.08	67.11	Ig heavy chain V-II region NEWM OS=Homo sapiens PE=1 SV=1 - [HV207_HUMAN]
P01781	17.24	2	2	116	12.7	8.48	109.19	Ig heavy chain V-III region GAL OS=Homo sapiens PE=1 SV=1 - [HV320_HUMAN]
P01777	15.97	1	1	119	12.8	8.50	103.30	Ig heavy chain V-III region TEI OS=Homo sapiens PE=1 SV=1 - [HV316_HUMAN]
P01765	16.52	1	1	115	12.3	9.13	118.42	Ig heavy chain V-III region TIL OS=Homo sapiens PE=1 SV=1 - [HV304_HUMAN]
P01764	9.40	1	1	117	12.6	8.28	58.96	Ig heavy chain V-III region VH26 OS=Homo sapiens PE=1 SV=1 - [HV303_HUMAN]
P01834	80.19	6	5	106	11.6	5.87	417.81	Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 SV=1 - [IGKC_HUMAN]
P01593	31.48	2	2	108	12.0	5.99	77.87	Ig kappa chain V-I region AG OS=Homo sapiens PE=1 SV=1 - [KV101_HUMAN]
P01598	26.85	2	2	108	11.8	8.44	137.52	Ig kappa chain V-I region EU OS=Homo sapiens PE=1 SV=1 - [KV106_HUMAN]
P01602	13.68	1	1	117	12.8	6.51	103.59	Ig kappa chain V-I region HK102 (Fragment) OS=Homo sapiens GN=IGKV1-5 PE=4 SV=1 - [KV110_HUMAN]
P01609	27.78	2	2	108	11.8	6.00	78.57	Ig kappa chain V-I region Scw OS=Homo sapiens PE=1 SV=1 - [KV117_HUMAN]
P01611	16.67	1	1	108	11.6	7.28	59.70	Ig kappa chain V-I region Wes OS=Homo sapiens PE=1 SV=1 - [KV119_HUMAN]
P01617	32.74	2	2	113	12.3	6.00	129.82	Ig kappa chain V-II region TEW OS=Homo sapiens PE=1 SV=1 - [KV204_HUMAN]
P01620	31.19	2	2	109	11.8	8.48	159.32	Ig kappa chain V-III region SIE OS=Homo sapiens PE=1 SV=1 - [KV302_HUMAN]
P01625	23.68	3	2	114	12.6	7.93	160.18	Ig kappa chain V-IV region Len OS=Homo sapiens PE=1 SV=2 - [KV402_HUMAN]
P0CG05	47.17	4	3	106	11.3	7.24	182.41	Ig lambda-2 chain C regions OS=Homo sapiens GN=IGLC2 PE=1 SV=1 - [LAC2_HUMAN]
P01871	3.54	1	1	452	49.3	6.77	83.13	Ig mu chain C region OS=Homo sapiens GN=IGHM PE=1 SV=3 - [IGHM_HUMAN]
Q9Y6R7	2.52	5	5	5405	571.6	5.34	331.41	IgGFc-binding protein OS=Homo sapiens GN=FCGBP PE=1 SV=3 - [FCGBP_HUMAN]
P01591	13.84	2	2	159	18.1	5.24	144.29	Immunoglobulin J chain OS=Homo sapiens GN=IGJ PE=1 SV=4 - [IGJ_HUMAN]
Q16270	18.44	3	3	282	29.1	7.90	139.57	Insulin-like growth factor-binding protein 7 OS=Homo sapiens GN=IGFBP7 PE=1 SV=1 - [IBP7_HUMAN]
Q14624	13.12	9	7	930	103.3	6.98	442.57	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4 - [ITIH4_HUMAN]
O75874	4.11	1	1	414	46.6	7.01	110.66	Isocitrate dehydrogenase [NADP] cytoplasmic OS=Homo sapiens GN=IDH1 PE=1 SV=2 - [IDHC_HUMAN]

P53990	7.69	2	2	364	39.7	5.35	133.33	IST1 homolog OS=Homo sapiens GN=KIAA0174 PE=1 SV=1 - [IST1_HUMAN]
P06870	14.50	3	3	262	28.9	4.83	177.21	Kallikrein-1 OS=Homo sapiens GN=KLK1 PE=1 SV=2 - [KLK1_HUMAN]
P29622	5.62	2	2	427	48.5	7.75	92.82	Kallistatin OS=Homo sapiens GN=SERPINA4 PE=1 SV=3 - [KAIN_HUMAN]
P13645	51.20	28	21	584	58.8	5.21	1584.77	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6 - [K1C10_HUMAN]
P13646	10.26	4	4	458	49.6	4.96	307.14	Keratin, type I cytoskeletal 13 OS=Homo sapiens GN=KRT13 PE=1 SV=4 - [K1C13_HUMAN]
P02533	15.89	6	6	472	51.5	5.16	443.14	Keratin, type I cytoskeletal 14 OS=Homo sapiens GN=KRT14 PE=1 SV=4 - [K1C14_HUMAN]
P08779	12.68	5	5	473	51.2	5.05	369.46	Keratin, type I cytoskeletal 16 OS=Homo sapiens GN=KRT16 PE=1 SV=4 - [K1C16_HUMAN]
P35527	38.04	15	12	623	62.0	5.24	1073.37	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3 - [K1C9_HUMAN]
P04264	30.28	22	16	644	66.0	8.12	1362.98	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6 - [K2C1_HUMAN]
P35908	34.12	19	18	639	65.4	8.00	1050.62	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2 - [K22E_HUMAN]
P19013	5.99	3	3	534	57.2	6.61	156.75	Keratin, type II cytoskeletal 4 OS=Homo sapiens GN=KRT4 PE=1 SV=4 - [K2C4_HUMAN]
P13647	13.22	8	7	590	62.3	7.74	381.72	Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 PE=1 SV=3 - [K2C5_HUMAN]
P02538	11.52	7	6	564	60.0	8.00	343.84	Keratin, type II cytoskeletal 6A OS=Homo sapiens GN=KRT6A PE=1 SV=3 - [K2C6A_HUMAN]
P04259	11.52	7	6	564	60.0	8.00	348.29	Keratin, type II cytoskeletal 6B OS=Homo sapiens GN=KRT6B PE=1 SV=5 - [K2C6B_HUMAN]
P01042	20.03	13	11	644	71.9	6.81	644.90	Kininogen-1 OS=Homo sapiens GN=KNG1 PE=1 SV=2 - [KNG1_HUMAN]
P02788	6.06	5	4	710	78.1	8.12	228.88	Lactotransferrin OS=Homo sapiens GN=LTF PE=1 SV=6 - [TRFL_HUMAN]
Q6GTX8	9.06	2	2	287	31.4	5.63	168.04	Leukocyte-associated immunoglobulin-like receptor 1 OS=Homo sapiens GN=LAIR1 PE=1 SV=1 - [LAIR1_HUMAN]
P00338	13.25	3	3	332	36.7	8.27	163.05	L-lactate dehydrogenase A chain OS=Homo sapiens GN=LDHA PE=1 SV=2 - [LDHA_HUMAN]
P07195	8.08	2	2	334	36.6	6.05	119.54	L-lactate dehydrogenase B chain OS=Homo sapiens GN=LDHB PE=1 SV=2 - [LDHB_HUMAN]
P98164	8.31	33	28	4655	521.6	5.08	2042.71	Low-density lipoprotein receptor-related protein 2 OS=Homo sapiens GN=LRP2 PE=1 SV=3 - [LRP2_HUMAN]
P11117	4.49	2	2	423	48.3	6.74	134.44	Lysosomal acid phosphatase OS=Homo sapiens GN=ACP2 PE=1 SV=3 - [PPAL_HUMAN]
P10253	15.65	12	10	952	105.3	5.99	823.12	Lysosomal alpha-glucosidase OS=Homo sapiens GN=GAA PE=1 SV=3 - [LYAG_HUMAN]
P10619	2.71	1	1	480	54.4	6.61	101.19	Lysosomal protective protein OS=Homo sapiens GN=CTSA PE=1

								SV=2 - [PPGB_HUMAN]
P42785	10.48	3	3	496	55.8	7.21	245.30	Lysosomal Pro-X carboxypeptidase OS=Homo sapiens GN=PRCP PE=1 SV=1 - [PCP_HUMAN]
P13473	7.07	3	3	410	44.9	5.63	136.85	Lysosome-associated membrane glycoprotein 2 OS=Homo sapiens GN=LAMP2 PE=1 SV=2 - [LAMP2_HUMAN]
O43451	17.66	22	22	1857	209.7	5.50	1360.42	Maltase-glucoamylase, intestinal OS=Homo sapiens GN=MGAM PE=1 SV=5 - [MGA_HUMAN]
O00187	3.21	2	2	686	75.7	5.77	85.39	Mannan-binding lectin serine protease 2 OS=Homo sapiens GN=MASP2 PE=1 SV=3 - [MASP2_HUMAN]
P02795	19.67	1	1	61	6.0	7.83	68.63	Metallothionein-2 OS=Homo sapiens GN=MT2A PE=1 SV=1 - [MT2_HUMAN]
P26038	3.47	2	2	577	67.8	6.40	150.97	Moesin OS=Homo sapiens GN=MSN PE=1 SV=3 - [MOES_HUMAN]
P08571	10.40	3	2	375	40.1	6.23	135.79	Monocyte differentiation antigen CD14 OS=Homo sapiens GN=CD14 PE=1 SV=2 - [CD14_HUMAN]
P15941	1.75	2	2	1255	122.0	7.47	138.62	Mucin-1 OS=Homo sapiens GN=MUC1 PE=1 SV=3 - [MUC1_HUMAN]
Q9HC84	2.03	3	3	5703	590.1	6.67	174.91	Mucin-5B OS=Homo sapiens GN=MUC5B PE=1 SV=2 - [MUC5B_HUMAN]
Q9H8L6	2.53	1	1	949	104.3	5.86	167.81	Multimerin-2 OS=Homo sapiens GN=MMRN2 PE=1 SV=2 - [MMRN2_HUMAN]
O95865	4.91	1	1	285	29.6	6.01	67.45	N(G),N(G)-dimethylarginine dimethylaminohydrolase 2 OS=Homo sapiens GN=DDAH2 PE=1 SV=1 - [DDAH2_HUMAN]
P34059	4.79	2	2	522	58.0	6.74	113.61	N-acetylgalactosamine-6-sulfatase OS=Homo sapiens GN=GALNS PE=1 SV=1 - [GALNS_HUMAN]
P15586	1.09	1	1	552	62.0	8.31	61.55	N-acetylglucosamine-6-sulfatase OS=Homo sapiens GN=GNS PE=1 SV=3 - [GNS_HUMAN]
O96009	9.52	3	3	420	45.4	6.61	154.35	Napsin-A OS=Homo sapiens GN=NAPSA PE=1 SV=1 - [NAPSA_HUMAN]
P08473	7.60	4	4	750	85.5	5.73	244.91	Neprilysin OS=Homo sapiens GN=MME PE=1 SV=2 - [NEP_HUMAN]
P59665	9.57	1	1	94	10.2	6.99	79.01	Neutrophil defensin 1 OS=Homo sapiens GN=DEFA1 PE=1 SV=1 - [DEF1_HUMAN]
Q92542	1.69	1	1	709	78.4	5.99	71.10	Nicastrin OS=Homo sapiens GN=NCSTN PE=1 SV=2 - [NICA_HUMAN]
P10153	10.56	2	1	161	18.3	8.73	133.16	Non-secretory ribonuclease OS=Homo sapiens GN=RNASE2 PE=1 SV=2 - [RNAS2_HUMAN]
P51688	7.37	3	3	502	56.7	6.95	144.60	N-sulphoglucosamine sulphohydrolase OS=Homo sapiens GN=SGSH PE=1 SV=1 - [SPHM_HUMAN]
P61970	11.02	2	1	127	14.5	5.38	195.33	Nuclear transport factor 2 OS=Homo sapiens GN=NUTF2 PE=1 SV=1 - [NTF2_HUMAN]
Q6UX06	16.27	7	6	510	57.2	5.69	366.99	Olfactomedin-4 OS=Homo sapiens GN=OLFM4 PE=1 SV=1 - [OLFM4_HUMAN]
P04746	35.23	11	11	511	57.7	7.05	566.84	Pancreatic alpha-amylase OS=Homo sapiens GN=AMY2A

								PE=1 SV=2 - [AMYP_HUMAN]
P55259	4.28	2	2	537	59.4	5.24	136.92	Pancreatic secretory granule membrane major glycoprotein GP2 OS=Homo sapiens GN=GP2 PE=2 SV=3 - [GP2_HUMAN]
Q9UBV8	7.39	2	2	284	30.4	6.54	130.49	Peflin OS=Homo sapiens GN=PEF1 PE=1 SV=1 - [PEF1_HUMAN]
O75594	39.29	7	4	196	21.7	8.59	362.59	Peptidoglycan recognition protein 1 OS=Homo sapiens GN=PGLYRP1 PE=1 SV=1 - [PGRP1_HUMAN]
P62937	16.36	2	2	165	18.0	7.81	82.35	Peptidyl-prolyl cis-trans isomerase A OS=Homo sapiens GN=PPIA PE=1 SV=2 - [PPIA_HUMAN]
Q06830	14.57	3	3	199	22.1	8.13	117.57	Peroxiredoxin-1 OS=Homo sapiens GN=PRDX1 PE=1 SV=1 - [PRDX1_HUMAN]
P30086	40.11	5	5	187	21.0	7.53	256.82	Phosphatidylethanolamine-binding protein 1 OS=Homo sapiens GN=PEBP1 PE=1 SV=3 - [PEBP1_HUMAN]
Q9Y646	15.04	5	5	472	51.9	6.18	308.83	Plasma glutamate carboxypeptidase OS=Homo sapiens GN=PGCP PE=1 SV=1 - [PGCP_HUMAN]
P05155	13.80	8	6	500	55.1	6.55	492.13	Plasma protease C1 inhibitor OS=Homo sapiens GN=SERPING1 PE=1 SV=2 - [IC1_HUMAN]
P05154	24.63	11	8	406	45.7	9.26	585.42	Plasma serine protease inhibitor OS=Homo sapiens GN=SERPINA5 PE=1 SV=2 - [IPSP_HUMAN]
P01833	30.24	20	16	764	83.2	5.74	1164.50	Polymeric immunoglobulin receptor OS=Homo sapiens GN=PIGR PE=1 SV=4 - [PIGR_HUMAN]
P0CG48	21.02	1	1	685	77.0	7.66	107.44	Polyubiquitin-C OS=Homo sapiens GN=UBC PE=1 SV=1 - [UBC_HUMAN]
P07602	12.60	7	4	524	58.1	5.17	483.56	Proactivator polypeptide OS=Homo sapiens GN=PSAP PE=1 SV=2 - [SAP_HUMAN]
Q9H3G5	3.99	1	1	476	54.1	5.62	88.11	Probable serine carboxypeptidase CPVL OS=Homo sapiens GN=CPVL PE=1 SV=2 - [CPVL_HUMAN]
P01133	18.31	18	16	1207	133.9	5.85	1086.32	Pro-epidermal growth factor OS=Homo sapiens GN=EGF PE=1 SV=2 - [EGF_HUMAN]
Q8WUM4	3.23	2	2	868	96.0	6.52	126.10	Programmed cell death 6-interacting protein OS=Homo sapiens GN=PDCD6IP PE=1 SV=1 - [PDC6I_HUMAN]
O75340	12.04	2	2	191	21.9	5.40	179.17	Programmed cell death protein 6 OS=Homo sapiens GN=PDCD6 PE=1 SV=1 - [PDCD6_HUMAN]
P12273	18.49	2	2	146	16.6	8.05	147.97	Prolactin-inducible protein OS=Homo sapiens GN=PIP PE=1 SV=1 - [PIP_HUMAN]
O43490	7.86	4	4	865	97.1	7.27	199.21	Prominin-1 OS=Homo sapiens GN=PROM1 PE=1 SV=1 - [PROM1_HUMAN]
P41222	21.05	5	3	190	21.0	7.80	309.59	Prostaglandin-H2 D-isomerase OS=Homo sapiens GN=PTGDS PE=1 SV=1 - [PTGDS_HUMAN]
Q16651	5.54	1	1	343	36.4	5.85	117.83	Prostasin OS=Homo sapiens GN=PRSS8 PE=1 SV=1 - [PRSS8_HUMAN]
O43653	8.13	1	1	123	12.9	5.29	79.41	Prostate stem cell antigen OS=Homo sapiens GN=PSCA PE=1 SV=1 - [PSCA_HUMAN]
P15309	16.32	7	5	386	44.5	6.24	443.71	Prostatic acid phosphatase OS=Homo sapiens GN=ACPP PE=1 SV=3 - [PPAP_HUMAN]

P02760	27.84	16	7	352	39.0	6.25	905.21	Protein AMBP OS=Homo sapiens GN=AMBP PE=1 SV=1 - [AMBP_HUMAN]
P31949	15.24	1	1	105	11.7	7.12	95.47	Protein S100-A11 OS=Homo sapiens GN=S100A11 PE=1 SV=2 - [S10AB_HUMAN]
P06703	8.89	1	1	90	10.2	5.48	91.52	Protein S100-A6 OS=Homo sapiens GN=S100A6 PE=1 SV=1 - [S10A6_HUMAN]
P05109	8.60	1	1	93	10.8	7.03	67.99	Protein S100-A8 OS=Homo sapiens GN=S100A8 PE=1 SV=1 - [S10A8_HUMAN]
P06702	26.32	2	2	114	13.2	6.13	144.59	Protein S100-A9 OS=Homo sapiens GN=S100A9 PE=1 SV=1 - [S10A9_HUMAN]
P25815	13.68	1	1	95	10.4	4.88	90.65	Protein S100-P OS=Homo sapiens GN=S100P PE=1 SV=2 - [S100P_HUMAN]
P00734	6.27	3	3	622	70.0	5.90	236.78	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2 - [THRB_HUMAN]
A6NIZ1	6.52	1	1	184	20.9	5.48	74.07	Ras-related protein Rap-1b-like protein OS=Homo sapiens PE=2 SV=1 - [RP1BL_HUMAN]
Q12913	4.26	4	4	1337	145.9	5.58	230.82	Receptor-type tyrosine-protein phosphatase eta OS=Homo sapiens GN=PTPRJ PE=1 SV=3 - [PTPRJ_HUMAN]
Q13332	0.67	1	1	1948	217.0	6.51	78.02	Receptor-type tyrosine-protein phosphatase S OS=Homo sapiens GN=PTPRS PE=1 SV=2 - [PTPRS_HUMAN]
O75787	2.57	1	1	350	39.0	6.10	78.18	Renin receptor OS=Homo sapiens GN=ATP6AP2 PE=1 SV=2 - [RENR_HUMAN]
Q9HD89	40.74	3	3	108	11.4	6.86	202.08	Resistin OS=Homo sapiens GN=RETN PE=2 SV=1 - [RETN_HUMAN]
P00352	2.99	1	1	501	54.8	6.73	96.77	Retinal dehydrogenase 1 OS=Homo sapiens GN=ALDH1A1 PE=1 SV=2 - [AL1A1_HUMAN]
Q9HB40	5.53	2	2	452	50.8	5.81	124.28	Retinoid-inducible serine carboxypeptidase OS=Homo sapiens GN=SCPEP1 PE=1 SV=1 - [RISC_HUMAN]
P07998	13.46	1	1	156	17.6	8.79	112.46	Ribonuclease pancreatic OS=Homo sapiens GN=RNASE1 PE=1 SV=4 - [RNAS1_HUMAN]
Q8WVN6	14.52	2	2	248	27.0	7.43	147.02	Secreted and transmembrane protein 1 OS=Homo sapiens GN=SECTM1 PE=1 SV=2 - [SCTM1_HUMAN]
P04279	2.60	1	1	462	52.1	9.29	89.93	Semenogelin-1 OS=Homo sapiens GN=SEMG1 PE=1 SV=2 - [SEMG1_HUMAN]
P02787	4.30	2	2	698	77.0	7.12	141.44	Serotransferrin OS=Homo sapiens GN=TF PE=1 SV=2 - [TRFE_HUMAN]
P29508	12.56	4	4	390	44.5	6.81	151.58	Serpin B3 OS=Homo sapiens GN=SERPINB3 PE=1 SV=2 - [SPB3_HUMAN]
P02768	61.25	47	31	609	69.3	6.28	2263.95	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2 - [ALBU_HUMAN]
P02743	10.31	2	2	223	25.4	6.54	137.60	Serum amyloid P-component OS=Homo sapiens GN=APCS PE=1 SV=2 - [SAMP_HUMAN]
Q9HAT2	11.09	5	5	523	58.3	7.33	266.74	Sialate O-acetylesterase OS=Homo sapiens GN=SIAE PE=1 SV=1 - [SIAE_HUMAN]

P48061	15.05	1	1	93	10.7	9.88	70.33	Stromal cell-derived factor 1 OS=Homo sapiens GN=CXCL12 PE=1 SV=1 - [SDF1_HUMAN]
O00391	5.09	4	3	747	82.5	8.92	210.97	Sulfhydryl oxidase 1 OS=Homo sapiens GN=QSOX1 PE=1 SV=3 - [QSOX1_HUMAN]
Q9UGT4	1.70	1	1	822	90.1	6.28	91.80	Sushi domain-containing protein 2 OS=Homo sapiens GN=SUSD2 PE=1 SV=1 - [SUSD2_HUMAN]
O00560	14.09	3	2	298	32.4	7.53	232.53	Syntenin-1 OS=Homo sapiens GN=SDCBP PE=1 SV=1 - [SDCB1_HUMAN]
P10599	12.38	1	1	105	11.7	4.92	73.53	Thioredoxin OS=Homo sapiens GN=TXN PE=1 SV=3 - [THIO_HUMAN]
P07996	0.94	1	1	1170	129.3	4.94	69.26	Thrombospondin-1 OS=Homo sapiens GN=THBS1 PE=1 SV=2 - [TSP1_HUMAN]
Q9UKU6	0.78	1	1	1024	116.9	6.99	63.27	Thyrotropin-releasing hormone-degrading ectoenzyme OS=Homo sapiens GN=TRHDE PE=2 SV=1 - [TRHDE_HUMAN]
P05543	4.58	2	2	415	46.3	6.30	84.27	Thyroxine-binding globulin OS=Homo sapiens GN=SERPINA7 PE=1 SV=2 - [THBG_HUMAN]
P02766	48.30	6	5	147	15.9	5.76	308.51	Transthyretin OS=Homo sapiens GN=TTR PE=1 SV=1 - [TTHY_HUMAN]
P60174	10.44	2	2	249	26.7	6.90	100.56	Triosephosphate isomerase OS=Homo sapiens GN=TPI1 PE=1 SV=2 - [TPIS_HUMAN]
O14773	9.77	4	3	563	61.2	6.48	231.93	Tripeptidyl-peptidase 1 OS=Homo sapiens GN=TPP1 PE=1 SV=2 - [TPP1_HUMAN]
Q99816	2.56	1	1	390	43.9	6.46	77.45	Tumor susceptibility gene 101 protein OS=Homo sapiens GN=TSG101 PE=1 SV=2 - [TS101_HUMAN]
P30530	2.01	1	1	894	98.3	5.43	69.75	Tyrosine-protein kinase receptor UFO OS=Homo sapiens GN=AXL PE=1 SV=3 - [UFO_HUMAN]
Q9H1C7	10.31	1	1	97	10.6	4.32	59.38	UPF0467 protein C5orf32 OS=Homo sapiens GN=C5orf32 PE=2 SV=1 - [CE032_HUMAN]
Q6UX73	3.23	1	1	402	45.4	6.19	59.34	UPF0764 protein C16orf89 OS=Homo sapiens GN=C16orf89 PE=2 SV=2 - [CP089_HUMAN]
P00749	11.83	4	4	431	48.5	8.41	248.62	Urokinase-type plasminogen activator OS=Homo sapiens GN=PLAU PE=1 SV=2 - [UROK_HUMAN]
P07911	39.22	86	18	640	69.7	5.24	5301.67	Uromodulin OS=Homo sapiens GN=UMOD PE=1 SV=1 - [UROM_HUMAN]
Q9H9H4	6.32	1	1	285	31.3	7.34	63.69	Vacuolar protein sorting-associated protein 37B OS=Homo sapiens GN=VPS37B PE=1 SV=1 - [VP37B_HUMAN]
Q9NP79	6.51	1	1	307	33.9	6.29	92.07	Vacuolar protein sorting-associated protein VTA1 homolog OS=Homo sapiens GN=VTA1 PE=1 SV=1 - [VTA1_HUMAN]
Q6EMK4	12.04	5	5	673	71.7	7.39	290.65	Vasorin OS=Homo sapiens GN=VASN PE=1 SV=1 - [VASN_HUMAN]
Q12907	28.37	6	6	356	40.2	6.95	362.95	Vesicular integral-membrane protein VIP36 OS=Homo sapiens GN=LMAN2 PE=1 SV=1 - [LMAN2_HUMAN]
P22891	12.50	3	3	400	44.7	5.97	221.30	Vitamin K-dependent protein Z OS=Homo sapiens GN=PROZ PE=1

								SV=2 - [PROZ_HUMAN]
Q7Z5L0	49.50	6	5	202	21.5	5.07	308.24	Vitelline membrane outer layer protein 1 homolog OS=Homo sapiens GN=VMO1 PE=1 SV=1 - [VMO1_HUMAN]
P04004	5.65	2	2	478	54.3	5.80	157.40	Vitronectin OS=Homo sapiens GN=VTN PE=1 SV=1 - [VTNC_HUMAN]
O43895	4.45	2	2	674	75.6	6.04	88.15	Xaa-Pro aminopeptidase 2 OS=Homo sapiens GN=XPNPEP2 PE=1 SV=3 - [XPP2_HUMAN]
P25311	26.51	7	7	298	34.2	6.05	437.07	Zinc-alpha-2-glycoprotein OS=Homo sapiens GN=AZGP1 PE=1 SV=2 - [ZA2G_HUMAN]
Q96DA0	30.77	5	5	208	22.7	7.39	304.14	Zymogen granule protein 16 homolog B OS=Homo sapiens GN=ZG16B PE=1 SV=3 - [ZG16B_HUMAN]

Supplementary table S2.2: Identification of proteins in (P18+P200) DTT SN200 sample 2. Accession of the protein, sequence coverage, PSM (Match between a fragmentation mass spectrum and peptide), number of peptides found, number of amino acid in the protein, molecular weight, calculated PI and mascot score along with the description of the protein are indicated in the table.

Accession	Coverage %	# PSMs	# Peptides	# AAs	MW [kDa]	calc. pI	Score	Description
P62258	7.45	1	1	255	29.2	4.74	122.69	14-3-3 protein epsilon OS=Homo sapiens GN=YWHAE PE=1 SV=1 - [1433E_HUMAN]
P63104	4.90	1	1	245	27.7	4.79	74.94	14-3-3 protein zeta/delta OS=Homo sapiens GN=YWHAZ PE=1 SV=1 - [1433Z_HUMAN]
P08195	4.44	2	2	630	68.0	5.01	114.93	4F2 cell-surface antigen heavy chain OS=Homo sapiens GN=SLC3A2 PE=1 SV=3 - [4F2_HUMAN]
O95336	6.20	1	1	258	27.5	6.05	77.20	6-phosphogluconolactonase OS=Homo sapiens GN=PGLS PE=1 SV=2 - [6PGL_HUMAN]
Q13510	25.82	7	6	395	44.6	7.62	428.61	Acid ceramidase OS=Homo sapiens GN=ASAH1 PE=1 SV=5 - [ASAH1_HUMAN]
P60709	9.87	3	3	375	41.7	5.48	203.53	Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1 - [ACTB_HUMAN]
P02763	26.37	4	4	201	23.5	5.02	176.91	Alpha-1-acid glycoprotein 1 OS=Homo sapiens GN=ORM1 PE=1 SV=1 - [A1AG1_HUMAN]
P19652	4.48	1	1	201	23.6	5.11	86.48	Alpha-1-acid glycoprotein 2 OS=Homo sapiens GN=ORM2 PE=1 SV=2 - [A1AG2_HUMAN]
P01011	17.02	5	5	423	47.6	5.52	284.75	Alpha-1-antichymotrypsin OS=Homo sapiens GN=SERPINA3 PE=1 SV=2 - [AACT_HUMAN]
P01009	28.95	11	8	418	46.7	5.59	750.55	Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3 - [A1AT_HUMAN]

P04217	6.46	2	2	495	54.2	5.87	109.32	Alpha-1B-glycoprotein OS=Homo sapiens GN=A1BG PE=1 SV=3 - [A1BG_HUMAN]
P02765	14.99	4	4	367	39.3	5.72	155.29	Alpha-2-HS-glycoprotein OS=Homo sapiens GN=AHSG PE=1 SV=1 - [FETUA_HUMAN]
P01023	0.81	1	1	1474	163.2	6.42	91.94	Alpha-2-macroglobulin OS=Homo sapiens GN=A2M PE=1 SV=2 - [A2MG_HUMAN]
O43707	4.83	4	4	911	104.8	5.44	221.54	Alpha-actinin-4 OS=Homo sapiens GN=ACTN4 PE=1 SV=2 - [ACTN4_HUMAN]
P04745	21.92	9	8	511	57.7	6.93	578.84	Alpha-amylase 1 OS=Homo sapiens GN=AMY1A PE=1 SV=2 - [AMY1_HUMAN]
P19961	21.92	9	8	511	57.7	7.09	565.89	Alpha-amylase 2B OS=Homo sapiens GN=AMY2B PE=1 SV=1 - [AMY2B_HUMAN]
P06733	9.22	3	3	434	47.1	7.39	149.46	Alpha-enolase OS=Homo sapiens GN=ENO1 PE=1 SV=2 - [ENOA_HUMAN]
P54802	28.13	15	14	743	82.1	6.54	855.70	Alpha-N-acetylglucosaminidase OS=Homo sapiens GN=NAGLU PE=1 SV=1 - [ANAG_HUMAN]
Q96Q42	0.36	1	1	1657	183.5	6.27	67.00	Alsin OS=Homo sapiens GN=ALS2 PE=1 SV=2 - [ALS2_HUMAN]
P15144	28.44	27	20	967	109.5	5.48	1477.51	Aminopeptidase N OS=Homo sapiens GN=ANPEP PE=1 SV=4 - [AMPN_HUMAN]
P12821	1.07	1	1	1306	149.6	6.39	67.68	Angiotensin-converting enzyme OS=Homo sapiens GN=ACE PE=1 SV=1 - [ACE_HUMAN]
P04083	3.76	1	1	346	38.7	7.02	92.17	Annexin A1 OS=Homo sapiens GN=ANXA1 PE=1 SV=2 - [ANXA1_HUMAN]
P07355	8.85	2	2	339	38.6	7.75	95.33	Annexin A2 OS=Homo sapiens GN=ANXA2 PE=1 SV=2 - [ANXA2_HUMAN]
P08758	5.00	1	1	320	35.9	5.05	73.02	Annexin A5 OS=Homo sapiens GN=ANXA5 PE=1 SV=2 - [ANXA5_HUMAN]
P02647	28.09	6	6	267	30.8	5.76	326.83	Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1 - [APOA1_HUMAN]
P02652	20.00	1	1	100	11.2	6.62	103.82	Apolipoprotein A-II OS=Homo sapiens GN=APOA2 PE=1 SV=1 - [APOA2_HUMAN]
P06727	3.54	1	1	396	45.4	5.38	87.66	Apolipoprotein A-IV OS=Homo sapiens GN=APOA4 PE=1 SV=3 - [APOA4_HUMAN]
P05090	34.92	17	7	189	21.3	5.15	781.10	Apolipoprotein D OS=Homo sapiens GN=APOD PE=1 SV=1 - [APOD_HUMAN]
P02649	26.50	7	7	317	36.1	5.73	400.82	Apolipoprotein E OS=Homo sapiens GN=APOE PE=1 SV=1 - [APOE_HUMAN]
P15289	14.00	4	4	507	53.6	6.07	302.58	Arylsulfatase A OS=Homo sapiens GN=ARSA PE=1 SV=3 - [ARSA_HUMAN]
O75882	5.39	6	6	1429	158.4	7.31	405.43	Attractin OS=Homo sapiens GN=ATRN PE=1 SV=2 - [ATRN_HUMAN]
P98160	0.91	3	3	4391	468.5	6.51	177.74	Basement membrane-specific heparan sulfate proteoglycan core protein OS=Homo sapiens GN=HSPG2 PE=1 SV=3 - [PGBM_HUMAN]
P16278	8.42	5	5	677	76.0	6.57	257.83	Beta-galactosidase OS=Homo sapiens GN=GLB1 PE=1 SV=2 - [BGAL_HUMAN]

P08236	10.29	4	4	651	74.7	7.02	273.62	Beta-glucuronidase OS=Homo sapiens GN=GUSB PE=1 SV=2 - [BGLR_HUMAN]
Q93088	14.53	4	4	406	45.0	7.03	227.22	Betaine--homocysteine S-methyltransferase 1 OS=Homo sapiens GN=BHMT PE=1 SV=2 - [BHMT1_HUMAN]
P19835	10.36	5	5	753	79.3	5.34	284.94	Bile salt-activated lipase OS=Homo sapiens GN=CEL PE=1 SV=3 - [CEL_HUMAN]
P43251	4.79	2	2	543	61.1	6.25	112.26	Biotinidase OS=Homo sapiens GN=BTD PE=1 SV=2 - [BTD_HUMAN]
Q5VW32	3.89	2	1	411	46.4	7.65	112.05	BRO1 domain-containing protein BROX OS=Homo sapiens GN=BROX PE=1 SV=1 - [BROX_HUMAN]
Q8WVV5	2.10	1	1	523	59.0	6.01	67.62	Butyrophilin subfamily 2 member A2 OS=Homo sapiens GN=BTN2A2 PE=2 SV=2 - [BT2A2_HUMAN]
P12830	7.94	4	4	882	97.4	4.73	211.52	Cadherin-1 OS=Homo sapiens GN=CDH1 PE=1 SV=3 - [CADH1_HUMAN]
P55290	1.68	1	1	713	78.2	4.98	63.42	Cadherin-13 OS=Homo sapiens GN=CDH13 PE=1 SV=1 - [CAD13_HUMAN]
P19022	4.08	3	2	906	99.7	4.81	196.74	Cadherin-2 OS=Homo sapiens GN=CDH2 PE=1 SV=4 - [CADH2_HUMAN]
Q9BYE9	1.98	2	2	1310	141.5	4.50	165.71	Cadherin-related family member 2 OS=Homo sapiens GN=CDHR2 PE=1 SV=2 - [CDHR2_HUMAN]
Q9HBB8	2.13	1	1	845	88.2	4.93	65.22	Cadherin-related family member 5 OS=Homo sapiens GN=CDHR5 PE=1 SV=3 - [CDHR5_HUMAN]
P05937	4.98	1	1	261	30.0	4.83	81.70	Calbindin OS=Homo sapiens GN=CALB1 PE=1 SV=2 - [CALB1_HUMAN]
P22792	3.30	1	1	545	60.6	5.99	98.38	Carboxypeptidase N subunit 2 OS=Homo sapiens GN=CPN2 PE=1 SV=2 - [CPN2_HUMAN]
P07339	18.20	5	5	412	44.5	6.54	249.28	Cathepsin D OS=Homo sapiens GN=CTSD PE=1 SV=1 - [CATD_HUMAN]
P16070	1.62	1	1	742	81.5	5.33	101.93	CD44 antigen OS=Homo sapiens GN=CD44 PE=1 SV=2 - [CD44_HUMAN]
P13987	25.00	5	3	128	14.2	6.48	225.02	CD59 glycoprotein OS=Homo sapiens GN=CD59 PE=1 SV=1 - [CD59_HUMAN]
P08962	4.20	1	1	238	25.6	7.81	63.44	CD63 antigen OS=Homo sapiens GN=CD63 PE=1 SV=2 - [CD63_HUMAN]
Q8NFZ8	6.19	2	2	388	42.8	6.30	99.56	Cell adhesion molecule 4 OS=Homo sapiens GN=CADM4 PE=1 SV=1 - [CADM4_HUMAN]
P00450	15.87	11	11	1065	122.1	5.72	563.68	Ceruloplasmin OS=Homo sapiens GN=CP PE=1 SV=1 - [CERU_HUMAN]
Q53GD3	1.97	1	1	710	79.2	8.59	105.10	Choline transporter-like protein 4 OS=Homo sapiens GN=SLC44A4 PE=2 SV=1 - [CTL4_HUMAN]
P10909	19.15	7	6	449	52.5	6.27	467.77	Clusterin OS=Homo sapiens GN=CLU PE=1 SV=1 - [CLUS_HUMAN]
Q9UGN4	6.02	1	1	299	33.2	5.49	67.94	CMRF35-like molecule 8 OS=Homo sapiens GN=CD300A PE=1 SV=2 - [CLM8_HUMAN]
P12109	6.23	5	5	1028	108.5	5.43	332.93	Collagen alpha-1(VI) chain OS=Homo sapiens GN=COL6A1

								PE=1 SV=3 - [CO6A1_HUMAN]
P39059	2.45	3	3	1388	141.6	5.00	125.32	Collagen alpha-1(XV) chain OS=Homo sapiens GN=COL15A1 PE=1 SV=2 - [COFA1_HUMAN]
P12111	1.51	4	3	3177	343.5	6.68	243.42	Collagen alpha-3(VI) chain OS=Homo sapiens GN=COL6A3 PE=1 SV=4 - [CO6A3_HUMAN]
P01024	2.29	3	3	1663	187.0	6.40	169.26	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2 - [CO3_HUMAN]
P0C0L4	1.49	2	2	1744	192.7	7.08	115.04	Complement C4-A OS=Homo sapiens GN=C4A PE=1 SV=1 - [CO4A_HUMAN]
Q12860	1.38	1	1	1018	113.2	5.90	59.76	Contactin-1 OS=Homo sapiens GN=CNTN1 PE=1 SV=1 - [CNTN1_HUMAN]
Q9UBG3	3.03	1	1	495	53.5	6.10	92.70	Cornulin OS=Homo sapiens GN=CRNN PE=1 SV=1 - [CRNN_HUMAN]
P12277	4.46	2	1	381	42.6	5.59	132.34	Creatine kinase B-type OS=Homo sapiens GN=CKB PE=1 SV=1 - [KCRB_HUMAN]
O60494	11.15	27	26	3623	398.4	5.35	1549.07	Cubilin OS=Homo sapiens GN=CUBN PE=1 SV=4 - [CUBN_HUMAN]
P01034	7.53	1	1	146	15.8	8.75	61.81	Cystatin-C OS=Homo sapiens GN=CST3 PE=1 SV=1 - [CYTC_HUMAN]
P16444	2.43	1	1	411	45.6	6.15	141.33	Dipeptidase 1 OS=Homo sapiens GN=DPEP1 PE=1 SV=3 - [DPEP1_HUMAN]
P53634	7.99	2	2	463	51.8	6.99	200.96	Dipeptidyl peptidase 1 OS=Homo sapiens GN=CTSC PE=1 SV=1 - [CATC_HUMAN]
P27487	8.49	5	5	766	88.2	6.04	465.76	Dipeptidyl peptidase 4 OS=Homo sapiens GN=DPP4 PE=1 SV=2 - [DPP4_HUMAN]
Q12805	16.63	6	6	493	54.6	5.07	278.96	EGF-containing fibulin-like extracellular matrix protein 1 OS=Homo sapiens GN=EFEMP1 PE=1 SV=2 - [FBLN3_HUMAN]
Q9UNN8	16.81	3	3	238	26.7	7.18	129.82	Endothelial protein C receptor OS=Homo sapiens GN=PROCR PE=1 SV=1 - [EPCR_HUMAN]
P08294	14.58	3	3	240	25.8	6.61	163.93	Extracellular superoxide dismutase [Cu-Zn] OS=Homo sapiens GN=SOD3 PE=1 SV=2 - [SODE_HUMAN]
P15311	6.83	3	3	586	69.4	6.27	195.43	Ezrin OS=Homo sapiens GN=EZR PE=1 SV=4 - [EZRI_HUMAN]
P02792	8.57	1	1	175	20.0	5.78	68.75	Ferritin light chain OS=Homo sapiens GN=FTL PE=1 SV=2 - [FRIL_HUMAN]
Q14314	8.66	2	2	439	50.2	7.39	126.25	Fibroleukin OS=Homo sapiens GN=FGL2 PE=1 SV=1 - [FGL2_HUMAN]
P02751	5.91	11	9	2386	262.5	5.71	557.56	Fibronectin OS=Homo sapiens GN=FN1 PE=1 SV=4 - [FINC_HUMAN]
P09467	5.03	1	1	338	36.8	6.99	73.25	Fructose-1,6-bisphosphatase 1 OS=Homo sapiens GN=FBP1 PE=1 SV=4 - [F16P1_HUMAN]
P05062	5.49	3	2	364	39.4	7.87	126.83	Fructose-bisphosphate aldolase B OS=Homo sapiens GN=ALDOB PE=1 SV=2 - [ALDOB_HUMAN]
Q08380	32.48	25	12	585	65.3	5.27	1412.53	Galectin-3-binding protein OS=Homo sapiens GN=LGALS3BP PE=1 SV=1 - [LG3BP_HUMAN]

O00182	3.38	1	1	355	39.5	9.17	71.56	Galectin-9 OS=Homo sapiens GN=LGALS9 PE=1 SV=2 - [LEG9_HUMAN]
Q92820	20.13	5	5	318	35.9	7.11	274.29	Gamma-glutamyl hydrolase OS=Homo sapiens GN=GGH PE=1 SV=2 - [GGH_HUMAN]
Q6P531	4.26	1	1	493	50.5	6.07	68.32	Gamma-glutamyltransferase 6 OS=Homo sapiens GN=GGT6 PE=2 SV=2 - [GGT6_HUMAN]
P19440	6.85	3	3	569	61.4	7.12	186.73	Gamma-glutamyltranspeptidase 1 OS=Homo sapiens GN=GGT1 PE=1 SV=2 - [GGT1_HUMAN]
P06396	9.72	5	5	782	85.6	6.28	322.59	Gelsolin OS=Homo sapiens GN=GSN PE=1 SV=1 - [GELS_HUMAN]
Q16769	20.50	6	5	361	40.9	6.61	369.01	Glutaminyl-peptide cyclotransferase OS=Homo sapiens GN=QPCT PE=1 SV=1 - [QPCT_HUMAN]
Q07075	2.61	2	2	957	109.2	5.47	129.06	Glutamyl aminopeptidase OS=Homo sapiens GN=ENPEP PE=1 SV=3 - [AMPE_HUMAN]
P22352	15.49	3	3	226	25.5	8.13	171.85	Glutathione peroxidase 3 OS=Homo sapiens GN=GPX3 PE=1 SV=2 - [GPX3_HUMAN]
P09211	16.67	2	2	210	23.3	5.64	146.27	Glutathione S-transferase P OS=Homo sapiens GN=GSTP1 PE=1 SV=2 - [GSTP1_HUMAN]
P04406	20.60	4	4	335	36.0	8.46	245.42	Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens GN=GAPDH PE=1 SV=3 - [G3P_HUMAN]
P28799	2.53	2	1	593	63.5	6.83	133.91	Granulins OS=Homo sapiens GN=GRN PE=1 SV=2 - [GRN_HUMAN]
Q99988	5.19	1	1	308	34.1	9.66	64.66	Growth/differentiation factor 15 OS=Homo sapiens GN=GDF15 PE=1 SV=2 - [GDF15_HUMAN]
Q8NHV1	2.00	2	1	300	34.5	6.46	81.04	GTPase IMAP family member 7 OS=Homo sapiens GN=GIMAP7 PE=2 SV=1 - [GIMA7_HUMAN]
P00738	17.73	6	6	406	45.2	6.58	272.41	Haptoglobin OS=Homo sapiens GN=HP PE=1 SV=1 - [HPT_HUMAN]
P11142	3.87	2	2	646	70.9	5.52	112.81	Heat shock cognate 71 kDa protein OS=Homo sapiens GN=HSPA8 PE=1 SV=1 - [HSP7C_HUMAN]
P02790	2.38	1	1	462	51.6	7.02	64.63	Hemopexin OS=Homo sapiens GN=HPX PE=1 SV=2 - [HEMO_HUMAN]
P37235	4.15	1	1	193	22.3	5.35	58.38	Hippocalcin-like protein 1 OS=Homo sapiens GN=HPCAL1 PE=1 SV=3 - [HPCL1_HUMAN]
P04196	5.71	2	2	525	59.5	7.50	115.42	Histidine-rich glycoprotein OS=Homo sapiens GN=HRG PE=1 SV=1 - [HRG_HUMAN]
Q86YZ3	1.68	1	1	2850	282.2	10.04	94.15	Hornerin OS=Homo sapiens GN=HRNR PE=1 SV=2 - [HORN_HUMAN]
Q12794	8.74	2	2	435	48.3	6.77	148.48	Hyaluronidase-1 OS=Homo sapiens GN=HYAL1 PE=1 SV=2 - [HYAL1_HUMAN]
O75144	8.61	2	2	302	33.3	5.31	108.54	ICOS ligand OS=Homo sapiens GN=ICOSLG PE=1 SV=2 - [ICOSL_HUMAN]
P01876	24.36	6	5	353	37.6	6.51	425.76	Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=2 - [IGHA1_HUMAN]
P01877	18.53	4	4	340	36.5	6.10	307.52	Ig alpha-2 chain C region OS=Homo sapiens GN=IGHA2

								PE=1 SV=3 - [IGHA2_HUMAN]
P01857	28.79	9	6	330	36.1	8.19	470.52	Ig gamma-1 chain C region OS=Homo sapiens GN=IGHG1 PE=1 SV=1 - [IGHG1_HUMAN]
P01859	16.87	6	4	326	35.9	7.59	296.87	Ig gamma-2 chain C region OS=Homo sapiens GN=IGHG2 PE=1 SV=2 - [IGHG2_HUMAN]
P01860	13.79	5	4	377	41.3	7.90	256.39	Ig gamma-3 chain C region OS=Homo sapiens GN=IGHG3 PE=1 SV=2 - [IGHG3_HUMAN]
P01861	18.35	7	4	327	35.9	7.36	354.82	Ig gamma-4 chain C region OS=Homo sapiens GN=IGHG4 PE=1 SV=1 - [IGHG4_HUMAN]
P01742	10.26	1	1	117	12.5	6.57	79.09	Ig heavy chain V-I region EU OS=Homo sapiens PE=1 SV=1 - [HV101_HUMAN]
P01766	25.00	2	2	120	13.2	6.57	117.64	Ig heavy chain V-III region BRO OS=Homo sapiens PE=1 SV=1 - [HV305_HUMAN]
P01781	17.24	2	2	116	12.7	8.48	108.38	Ig heavy chain V-III region GAL OS=Homo sapiens PE=1 SV=1 - [HV320_HUMAN]
P01765	16.52	2	1	115	12.3	9.13	129.56	Ig heavy chain V-III region TIL OS=Homo sapiens PE=1 SV=1 - [HV304_HUMAN]
P01762	15.57	2	1	122	13.5	9.72	115.81	Ig heavy chain V-III region TRO OS=Homo sapiens PE=1 SV=1 - [HV301_HUMAN]
P01834	63.21	11	5	106	11.6	5.87	554.19	Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 SV=1 - [IGKC_HUMAN]
P01593	16.67	2	1	108	12.0	5.99	165.90	Ig kappa chain V-I region AG OS=Homo sapiens PE=1 SV=1 - [KV101_HUMAN]
P01596	16.82	1	1	107	11.7	9.41	96.06	Ig kappa chain V-I region CAR OS=Homo sapiens PE=1 SV=1 - [KV104_HUMAN]
P01611	16.67	1	1	108	11.6	7.28	61.82	Ig kappa chain V-I region Wes OS=Homo sapiens PE=1 SV=1 - [KV119_HUMAN]
P01616	11.61	1	1	112	12.0	9.29	73.34	Ig kappa chain V-II region MIL OS=Homo sapiens PE=1 SV=1 - [KV203_HUMAN]
P01617	21.24	1	1	113	12.3	6.00	122.31	Ig kappa chain V-II region TEW OS=Homo sapiens PE=1 SV=1 - [KV204_HUMAN]
P01620	14.68	1	1	109	11.8	8.48	78.40	Ig kappa chain V-III region SIE OS=Homo sapiens PE=1 SV=1 - [KV302_HUMAN]
P06312	29.75	3	3	121	13.4	5.25	113.02	Ig kappa chain V-IV region (Fragment) OS=Homo sapiens GN=IGKV4-1 PE=4 SV=1 - [KV401_HUMAN]
P01625	36.84	4	3	114	12.6	7.93	166.81	Ig kappa chain V-IV region Len OS=Homo sapiens PE=1 SV=2 - [KV402_HUMAN]
P01702	7.21	1	1	111	11.4	4.89	58.86	Ig lambda chain V-I region NIG-64 OS=Homo sapiens PE=1 SV=1 - [LV104_HUMAN]
P0CG05	50.94	4	3	106	11.3	7.24	249.63	Ig lambda-2 chain C regions OS=Homo sapiens GN=IGLC2 PE=1 SV=1 - [LAC2_HUMAN]
P01871	3.54	1	1	452	49.3	6.77	65.99	Ig mu chain C region OS=Homo sapiens GN=IGHM PE=1 SV=3 - [IGHM_HUMAN]
Q9Y6R7	2.35	5	5	5405	571.6	5.34	256.55	IgGFc-binding protein OS=Homo sapiens GN=FCGBP PE=1 SV=3 - [FCGBP_HUMAN]
P01591	23.27	3	3	159	18.1	5.24	188.95	Immunoglobulin J chain OS=Homo sapiens GN=IGJ PE=1

								SV=4 - [IGJ_HUMAN]
Q16270	17.73	4	3	282	29.1	7.90	205.49	Insulin-like growth factor-binding protein 7 OS=Homo sapiens GN=IGFBP7 PE=1 SV=1 - [IBP7_HUMAN]
Q9Y287	3.76	1	1	266	30.3	5.14	58.87	Integral membrane protein 2B OS=Homo sapiens GN=ITM2B PE=1 SV=1 - [ITM2B_HUMAN]
Q14624	9.89	6	6	930	103.3	6.98	390.41	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4 - [ITIH4_HUMAN]
P53990	5.22	1	1	364	39.7	5.35	102.01	IST1 homolog OS=Homo sapiens GN=KIAA0174 PE=1 SV=1 - [IST1_HUMAN]
P06870	16.03	3	3	262	28.9	4.83	177.69	Kallikrein-1 OS=Homo sapiens GN=KLK1 PE=1 SV=2 - [KLK1_HUMAN]
P29622	3.51	1	1	427	48.5	7.75	103.42	Kallistatin OS=Homo sapiens GN=SERPINA4 PE=1 SV=3 - [KAIN_HUMAN]
P13645	42.47	24	19	584	58.8	5.21	1250.81	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6 - [K1C10_HUMAN]
P13646	13.54	5	5	458	49.6	4.96	345.59	Keratin, type I cytoskeletal 13 OS=Homo sapiens GN=KRT13 PE=1 SV=4 - [K1C13_HUMAN]
P02533	19.07	7	7	472	51.5	5.16	370.50	Keratin, type I cytoskeletal 14 OS=Homo sapiens GN=KRT14 PE=1 SV=4 - [K1C14_HUMAN]
P08779	14.38	6	6	473	51.2	5.05	311.81	Keratin, type I cytoskeletal 16 OS=Homo sapiens GN=KRT16 PE=1 SV=4 - [K1C16_HUMAN]
P35527	30.18	14	12	623	62.0	5.24	941.86	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3 - [K1C9_HUMAN]
P04264	28.73	23	17	644	66.0	8.12	1050.10	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6 - [K2C1_HUMAN]
P35908	28.64	16	14	639	65.4	8.00	827.60	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2 - [K22E_HUMAN]
P13647	11.86	7	6	590	62.3	7.74	349.88	Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 PE=1 SV=3 - [K2C5_HUMAN]
P04259	8.51	5	4	564	60.0	8.00	295.13	Keratin, type II cytoskeletal 6B OS=Homo sapiens GN=KRT6B PE=1 SV=5 - [K2C6B_HUMAN]
P01042	14.44	12	7	644	71.9	6.81	617.07	Kininogen-1 OS=Homo sapiens GN=KNG1 PE=1 SV=2 - [KNG1_HUMAN]
P02788	6.48	5	4	710	78.1	8.12	227.66	Lactotransferrin OS=Homo sapiens GN=LTF PE=1 SV=6 - [TRFL_HUMAN]
P02750	6.63	1	1	347	38.2	6.95	80.58	Leucine-rich alpha-2-glycoprotein OS=Homo sapiens GN=LRG1 PE=1 SV=2 - [A2GL_HUMAN]
Q6GTX8	9.06	2	2	287	31.4	5.63	120.39	Leukocyte-associated immunoglobulin-like receptor 1 OS=Homo sapiens GN=LAIR1 PE=1 SV=1 - [LAIR1_HUMAN]
P00338	9.94	3	3	332	36.7	8.27	144.81	L-lactate dehydrogenase A chain OS=Homo sapiens GN=LDHA PE=1 SV=2 - [LDHA_HUMAN]
P98164	7.82	28	27	4655	521.6	5.08	1723.08	Low-density lipoprotein receptor-related protein 2 OS=Homo sapiens GN=LRP2 PE=1 SV=3 - [LRP2_HUMAN]

P11117	4.49	2	2	423	48.3	6.74	143.48	Lysosomal acid phosphatase OS=Homo sapiens GN=ACP2 PE=1 SV=3 - [PPAL_HUMAN]
P10253	18.59	17	11	952	105.3	5.99	898.30	Lysosomal alpha-glucosidase OS=Homo sapiens GN=GAA PE=1 SV=3 - [LYAG_HUMAN]
P10619	2.71	1	1	480	54.4	6.61	105.17	Lysosomal protective protein OS=Homo sapiens GN=CTSA PE=1 SV=2 - [PPGB_HUMAN]
P42785	10.48	3	3	496	55.8	7.21	200.30	Lysosomal Pro-X carboxypeptidase OS=Homo sapiens GN=PRCP PE=1 SV=1 - [PCP_HUMAN]
P13473	4.88	2	2	410	44.9	5.63	100.58	Lysosome-associated membrane glycoprotein 2 OS=Homo sapiens GN=LAMP2 PE=1 SV=2 - [LAMP2_HUMAN]
O43451	15.62	21	20	1857	209.7	5.50	1218.94	Maltase-glucoamylase, intestinal OS=Homo sapiens GN=MGAM PE=1 SV=5 - [MGA_HUMAN]
O00187	3.21	2	2	686	75.7	5.77	82.18	Mannan-binding lectin serine protease 2 OS=Homo sapiens GN=MASP2 PE=1 SV=3 - [MASP2_HUMAN]
Q9NR99	0.99	2	2	2828	312.1	8.32	103.53	Matrix-remodeling-associated protein 5 OS=Homo sapiens GN=MXRA5 PE=2 SV=2 - [MXRA5_HUMAN]
P26038	1.91	1	1	577	67.8	6.40	108.62	Moesin OS=Homo sapiens GN=MSN PE=1 SV=3 - [MOES_HUMAN]
P08571	14.40	3	3	375	40.1	6.23	136.45	Monocyte differentiation antigen CD14 OS=Homo sapiens GN=CD14 PE=1 SV=2 - [CD14_HUMAN]
P15941	0.88	1	1	1255	122.0	7.47	135.46	Mucin-1 OS=Homo sapiens GN=MUC1 PE=1 SV=3 - [MUC1_HUMAN]
Q9HC84	2.03	3	3	5703	590.1	6.67	145.40	Mucin-5B OS=Homo sapiens GN=MUC5B PE=1 SV=2 - [MUC5B_HUMAN]
Q9H8L6	4.00	2	2	949	104.3	5.86	196.33	Multimerin-2 OS=Homo sapiens GN=MMRN2 PE=1 SV=2 - [MMRN2_HUMAN]
P34059	3.07	1	1	522	58.0	6.74	70.13	N-acetylgalactosamine-6-sulfatase OS=Homo sapiens GN=GALNS PE=1 SV=1 - [GALNS_HUMAN]
O96009	9.52	3	3	420	45.4	6.61	142.51	Napsin-A OS=Homo sapiens GN=NAPSA PE=1 SV=1 - [NAPSA_HUMAN]
P08473	5.20	3	3	750	85.5	5.73	221.46	Neprilysin OS=Homo sapiens GN=MME PE=1 SV=2 - [NEP_HUMAN]
P80188	7.58	1	1	198	22.6	8.91	119.08	Neutrophil gelatinase-associated lipocalin OS=Homo sapiens GN=LCN2 PE=1 SV=2 - [NGAL_HUMAN]
Q92542	1.69	1	1	709	78.4	5.99	73.57	Nicastrin OS=Homo sapiens GN=NCSTN PE=1 SV=2 - [NICA_HUMAN]
P10153	10.56	2	1	161	18.3	8.73	126.91	Non-secretory ribonuclease OS=Homo sapiens GN=RNASE2 PE=1 SV=2 - [RNAS2_HUMAN]
P51688	7.37	3	3	502	56.7	6.95	163.74	N-sulphoglucosamine sulphohydrolase OS=Homo sapiens GN=SGSH PE=1 SV=1 - [SPHM_HUMAN]
P61970	23.62	3	2	127	14.5	5.38	153.41	Nuclear transport factor 2 OS=Homo sapiens GN=NUTF2 PE=1 SV=1 - [NTF2_HUMAN]

Q6UX06	13.53	5	5	510	57.2	5.69	315.00	Olfactomedin-4 OS=Homo sapiens GN=OLFM4 PE=1 SV=1 - [OLFM4_HUMAN]
P04746	19.37	8	7	511	57.7	7.05	560.45	Pancreatic alpha-amylase OS=Homo sapiens GN=AMY2A PE=1 SV=2 - [AMYP_HUMAN]
P55259	4.28	2	2	537	59.4	5.24	88.99	Pancreatic secretory granule membrane major glycoprotein GP2 OS=Homo sapiens GN=GP2 PE=2 SV=3 - [GP2_HUMAN]
Q9UBV8	11.97	3	3	284	30.4	6.54	165.56	Peflin OS=Homo sapiens GN=PEF1 PE=1 SV=1 - [PEF1_HUMAN]
O75594	39.29	5	4	196	21.7	8.59	301.86	Peptidoglycan recognition protein 1 OS=Homo sapiens GN=PGLYRP1 PE=1 SV=1 - [PGRP1_HUMAN]
Q06830	5.03	1	1	199	22.1	8.13	68.82	Peroxiredoxin-1 OS=Homo sapiens GN=PRDX1 PE=1 SV=1 - [PRDX1_HUMAN]
P30086	32.09	4	4	187	21.0	7.53	229.60	Phosphatidylethanolamine-binding protein 1 OS=Homo sapiens GN=PEBP1 PE=1 SV=3 - [PEBP1_HUMAN]
Q9Y646	16.95	6	6	472	51.9	6.18	334.27	Plasma glutamate carboxypeptidase OS=Homo sapiens GN=PGCP PE=1 SV=1 - [PGCP_HUMAN]
P05155	13.40	7	5	500	55.1	6.55	482.11	Plasma protease C1 inhibitor OS=Homo sapiens GN=SERPING1 PE=1 SV=2 - [IC1_HUMAN]
P05154	21.43	10	7	406	45.7	9.26	546.80	Plasma serine protease inhibitor OS=Homo sapiens GN=SERPINA5 PE=1 SV=2 - [IPSP_HUMAN]
P15151	3.12	1	1	417	45.3	6.52	78.42	Poliovirus receptor OS=Homo sapiens GN=PVR PE=1 SV=2 - [PVR_HUMAN]
P01833	28.93	22	15	764	83.2	5.74	1225.49	Polymeric immunoglobulin receptor OS=Homo sapiens GN=PIGR PE=1 SV=4 - [PIGR_HUMAN]
P0CG48	11.82	1	1	685	77.0	7.66	91.36	Polyubiquitin-C OS=Homo sapiens GN=UBC PE=1 SV=1 - [UBC_HUMAN]
P07602	14.69	8	5	524	58.1	5.17	489.67	Proactivator polypeptide OS=Homo sapiens GN=PSAP PE=1 SV=2 - [SAP_HUMAN]
P01133	16.74	16	14	1207	133.9	5.85	955.71	Pro-epidermal growth factor OS=Homo sapiens GN=EGF PE=1 SV=2 - [EGF_HUMAN]
Q8WUM4	3.23	2	2	868	96.0	6.52	145.44	Programmed cell death 6-interacting protein OS=Homo sapiens GN=PDCD6IP PE=1 SV=1 - [PDC6I_HUMAN]
O75340	16.75	3	3	191	21.9	5.40	178.08	Programmed cell death protein 6 OS=Homo sapiens GN=PDCD6 PE=1 SV=1 - [PDCD6_HUMAN]
P12273	8.22	1	1	146	16.6	8.05	88.34	Prolactin-inducible protein OS=Homo sapiens GN=PIP PE=1 SV=1 - [PIP_HUMAN]
O43490	4.05	2	2	865	97.1	7.27	146.25	Prominin-1 OS=Homo sapiens GN=PROM1 PE=1 SV=1 - [PROM1_HUMAN]
P41222	17.37	3	2	190	21.0	7.80	262.29	Prostaglandin-H2 D-isomerase OS=Homo sapiens GN=PTGDS PE=1 SV=1 - [PTGDS_HUMAN]
Q16651	4.08	1	1	343	36.4	5.85	106.87	Prostasin OS=Homo sapiens GN=PRSS8 PE=1 SV=1 - [PRSS8_HUMAN]

O43653	8.13	1	1	123	12.9	5.29	61.64	Prostate stem cell antigen OS=Homo sapiens GN=PSCA PE=1 SV=1 - [PSCA_HUMAN]
P15309	12.18	7	4	386	44.5	6.24	409.77	Prostatic acid phosphatase OS=Homo sapiens GN=ACPP PE=1 SV=3 - [PPAP_HUMAN]
P02760	30.11	16	8	352	39.0	6.25	950.65	Protein AMBP OS=Homo sapiens GN=AMBP PE=1 SV=1 - [AMBP_HUMAN]
P31949	15.24	1	1	105	11.7	7.12	92.18	Protein S100-A11 OS=Homo sapiens GN=S100A11 PE=1 SV=2 - [S10AB_HUMAN]
P06703	27.78	2	2	90	10.2	5.48	118.96	Protein S100-A6 OS=Homo sapiens GN=S100A6 PE=1 SV=1 - [S10A6_HUMAN]
P06702	26.32	2	2	114	13.2	6.13	120.97	Protein S100-A9 OS=Homo sapiens GN=S100A9 PE=1 SV=1 - [S10A9_HUMAN]
P25815	13.68	1	1	95	10.4	4.88	84.46	Protein S100-P OS=Homo sapiens GN=S100P PE=1 SV=2 - [S100P_HUMAN]
P00734	5.95	2	2	622	70.0	5.90	190.38	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2 - [THRB_HUMAN]
A6NIZ1	6.52	1	1	184	20.9	5.48	74.32	Ras-related protein Rap-1b-like protein OS=Homo sapiens PE=2 SV=1 - [RP1BL_HUMAN]
Q12913	2.32	2	2	1337	145.9	5.58	182.06	Receptor-type tyrosine-protein phosphatase eta OS=Homo sapiens GN=PTPRJ PE=1 SV=3 - [PTPRJ_HUMAN]
Q13332	0.67	1	1	1948	217.0	6.51	62.90	Receptor-type tyrosine-protein phosphatase S OS=Homo sapiens GN=PTPRS PE=1 SV=2 - [PTPRS_HUMAN]
O75787	2.57	1	1	350	39.0	6.10	67.12	Renin receptor OS=Homo sapiens GN=ATP6AP2 PE=1 SV=2 - [RENR_HUMAN]
Q9HD89	40.74	3	3	108	11.4	6.86	152.72	Resistin OS=Homo sapiens GN=RETN PE=2 SV=1 - [RETN_HUMAN]
P00352	2.99	1	1	501	54.8	6.73	105.59	Retinal dehydrogenase 1 OS=Homo sapiens GN=ALDH1A1 PE=1 SV=2 - [AL1A1_HUMAN]
Q9HB40	5.53	2	2	452	50.8	5.81	127.22	Retinoid-inducible serine carboxypeptidase OS=Homo sapiens GN=SCPEP1 PE=1 SV=1 - [RISC_HUMAN]
P07998	17.95	2	2	156	17.6	8.79	93.48	Ribonuclease pancreatic OS=Homo sapiens GN=RNASE1 PE=1 SV=4 - [RNAS1_HUMAN]
Q8WVN6	14.52	2	2	248	27.0	7.43	142.17	Secreted and transmembrane protein 1 OS=Homo sapiens GN=SECTM1 PE=1 SV=2 - [SCTM1_HUMAN]
P04279	7.36	2	2	462	52.1	9.29	97.30	Semenogelin-1 OS=Homo sapiens GN=SEMG1 PE=1 SV=2 - [SEMG1_HUMAN]
P02787	8.02	4	4	698	77.0	7.12	159.56	Serotransferrin OS=Homo sapiens GN=TF PE=1 SV=2 - [TRFE_HUMAN]
P29508	5.90	2	2	390	44.5	6.81	126.41	Serpin B3 OS=Homo sapiens GN=SERPINB3 PE=1 SV=2 - [SPB3_HUMAN]
P02768	49.26	40	23	609	69.3	6.28	2325.90	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2 - [ALBU_HUMAN]
P02743	20.63	4	4	223	25.4	6.54	178.81	Serum amyloid P-component OS=Homo sapiens GN=APCS PE=1 SV=2 - [SAMP_HUMAN]

Q9HAT2	7.65	3	3	523	58.3	7.33	188.56	Sialate O-acetylesterase OS=Homo sapiens GN=SIAE PE=1 SV=1 - [SIAE_HUMAN]
P48061	15.05	1	1	93	10.7	9.88	66.94	Stromal cell-derived factor 1 OS=Homo sapiens GN=CXCL12 PE=1 SV=1 - [SDF1_HUMAN]
P53597	4.34	1	1	346	36.2	8.79	62.08	Succinyl-CoA ligase [GDP-forming] subunit alpha, mitochondrial OS=Homo sapiens GN=SUCLG1 PE=1 SV=4 - [SUCA_HUMAN]
O00391	1.87	1	1	747	82.5	8.92	128.99	Sulfhydryl oxidase 1 OS=Homo sapiens GN=QSOX1 PE=1 SV=3 - [QSOX1_HUMAN]
Q9UGT4	1.70	1	1	822	90.1	6.28	83.61	Sushi domain-containing protein 2 OS=Homo sapiens GN=SUSD2 PE=1 SV=1 - [SUSD2_HUMAN]
O00560	14.43	3	2	298	32.4	7.53	300.83	Syntenin-1 OS=Homo sapiens GN=SDCBP PE=1 SV=1 - [SDCB1_HUMAN]
P10599	12.38	1	1	105	11.7	4.92	85.98	Thioredoxin OS=Homo sapiens GN=TXN PE=1 SV=3 - [THIO_HUMAN]
P05543	6.75	3	3	415	46.3	6.30	120.75	Thyroxine-binding globulin OS=Homo sapiens GN=SERPINA7 PE=1 SV=2 - [THBG_HUMAN]
P02766	48.30	5	5	147	15.9	5.76	247.35	Transthyretin OS=Homo sapiens GN=TTR PE=1 SV=1 - [TTHY_HUMAN]
P60174	10.44	2	2	249	26.7	6.90	80.80	Triosephosphate isomerase OS=Homo sapiens GN=TPI1 PE=1 SV=2 - [TPIS_HUMAN]
O14773	9.77	4	3	563	61.2	6.48	242.71	Tripeptidyl-peptidase 1 OS=Homo sapiens GN=TPP1 PE=1 SV=2 - [TPP1_HUMAN]
P30530	2.01	1	1	894	98.3	5.43	72.45	Tyrosine-protein kinase receptor UFO OS=Homo sapiens GN=AXL PE=1 SV=3 - [UFO_HUMAN]
Q9H1C7	10.31	1	1	97	10.6	4.32	59.23	UPF0467 protein C5orf32 OS=Homo sapiens GN=C5orf32 PE=2 SV=1 - [CE032_HUMAN]
Q6UX73	3.23	1	1	402	45.4	6.19	72.46	UPF0764 protein C16orf89 OS=Homo sapiens GN=C16orf89 PE=2 SV=2 - [CP089_HUMAN]
P00749	4.87	2	2	431	48.5	8.41	134.85	Urokinase-type plasminogen activator OS=Homo sapiens GN=PLAU PE=1 SV=2 - [UROK_HUMAN]
P07911	34.84	84	17	640	69.7	5.24	4820.87	Uromodulin OS=Homo sapiens GN=UMOD PE=1 SV=1 - [UROM_HUMAN]
Q9NP79	6.51	1	1	307	33.9	6.29	66.83	Vacuolar protein sorting-associated protein VTA1 homolog OS=Homo sapiens GN=VTA1 PE=1 SV=1 - [VTA1_HUMAN]
Q6EMK4	10.10	4	4	673	71.7	7.39	242.73	Vasorin OS=Homo sapiens GN=VASN PE=1 SV=1 - [VASN_HUMAN]
Q12907	18.54	4	4	356	40.2	6.95	331.12	Vesicular integral-membrane protein VIP36 OS=Homo sapiens GN=LMAN2 PE=1 SV=1 - [LMAN2_HUMAN]
P22891	13.50	4	4	400	44.7	5.97	184.28	Vitamin K-dependent protein Z OS=Homo sapiens GN=PROZ PE=1 SV=2 - [PROZ_HUMAN]
Q7Z5L0	54.95	8	6	202	21.5	5.07	371.00	Vitelline membrane outer layer protein 1 homolog OS=Homo sapiens GN=VMO1 PE=1 SV=1 - [VMO1_HUMAN]

P04004	10.25	3	3	478	54.3	5.80	177.86	Vitronectin OS=Homo sapiens GN=VTN PE=1 SV=1 - [VTNC_HUMAN]
Q14508	11.29	2	1	124	13.0	4.84	103.89	WAP four-disulfide core domain protein 2 OS=Homo sapiens GN=WFDC2 PE=1 SV=2 - [WFDC2_HUMAN]
O43895	4.45	2	2	674	75.6	6.04	97.58	Xaa-Pro aminopeptidase 2 OS=Homo sapiens GN=XPNPEP2 PE=1 SV=3 - [XPP2_HUMAN]
P25311	32.21	8	8	298	34.2	6.05	412.05	Zinc-alpha-2-glycoprotein OS=Homo sapiens GN=AZGP1 PE=1 SV=2 - [ZA2G_HUMAN]
Q96DA0	30.77	5	5	208	22.7	7.39	270.94	Zymogen granule protein 16 homolog B OS=Homo sapiens GN=ZG16B PE=1 SV=3 - [ZG16B_HUMAN]

Supplementary table S2.3: Identification of proteins in (P18+P200) CHAPS SN200 sample 2. Accession of the protein, sequence coverage, PSM (Match between a fragmentation mass spectrum and peptide), number of peptides found, number of amino acids in the protein, molecular weight, calculated PI and mascot score along with the description of the protein are indicated in the table.

Accession	Coverage %	# PSMs	# Peptides	# AAs	MW [kDa]	calc. pI	Score	Description
P63104	4.90	1	1	245	27.7	4.79	59.74	14-3-3 protein zeta/delta OS=Homo sapiens GN=YWHAZ PE=1 SV=1 - [1433Z_HUMAN]
Q13510	30.13	13	9	395	44.6	7.62	679.07	Acid ceramidase OS=Homo sapiens GN=ASAH1 PE=1 SV=5 - [ASAH1_HUMAN]
Q92485	2.64	1	1	455	50.8	5.64	68.27	Acid sphingomyelinase-like phosphodiesterase 3b OS=Homo sapiens GN=SMPDL3B PE=2 SV=2 - [ASM3B_HUMAN]
P60709	9.60	3	3	375	41.7	5.48	189.11	Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1 - [ACTB_HUMAN]
P02763	32.34	6	5	201	23.5	5.02	242.57	Alpha-1-acid glycoprotein 1 OS=Homo sapiens GN=ORM1 PE=1 SV=1 - [A1AG1_HUMAN]
P19652	12.94	2	2	201	23.6	5.11	92.87	Alpha-1-acid glycoprotein 2 OS=Homo sapiens GN=ORM2 PE=1 SV=2 - [A1AG2_HUMAN]
P01011	13.71	4	4	423	47.6	5.52	279.42	Alpha-1-antichymotrypsin OS=Homo sapiens GN=SERPINA3 PE=1 SV=2 - [AACT_HUMAN]
P01009	33.73	17	12	418	46.7	5.59	877.80	Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3 - [A1AT_HUMAN]
P04217	9.90	4	4	495	54.2	5.87	186.68	Alpha-1B-glycoprotein OS=Homo sapiens GN=A1BG PE=1 SV=3 - [A1BG_HUMAN]
P02765	16.62	5	5	367	39.3	5.72	335.80	Alpha-2-HS-glycoprotein OS=Homo sapiens GN=AHSG PE=1 SV=1 - [FETUA_HUMAN]

P12814	1.12	1	1	892	103.0	5.41	80.88	Alpha-actinin-1 OS=Homo sapiens GN=ACTN1 PE=1 SV=2 - [ACTN1_HUMAN]
P04745	10.37	4	4	511	57.7	6.93	336.72	Alpha-amylase 1 OS=Homo sapiens GN=AMY1A PE=1 SV=2 - [AMY1_HUMAN]
P19961	10.37	4	4	511	57.7	7.09	337.82	Alpha-amylase 2B OS=Homo sapiens GN=AMY2B PE=1 SV=1 - [AMY2B_HUMAN]
P06733	8.53	2	2	434	47.1	7.39	181.18	Alpha-enolase OS=Homo sapiens GN=ENO1 PE=1 SV=2 - [ENOA_HUMAN]
P54802	23.15	12	11	743	82.1	6.54	578.03	Alpha-N-acetylglucosaminidase OS=Homo sapiens GN=NAGLU PE=1 SV=1 - [ANAG_HUMAN]
P15144	27.92	31	20	967	109.5	5.48	1761.40	Aminopeptidase N OS=Homo sapiens GN=ANPEP PE=1 SV=4 - [AMPN_HUMAN]
P01019	10.52	4	3	485	53.1	6.32	177.07	Angiotensinogen OS=Homo sapiens GN=AGT PE=1 SV=1 - [ANGT_HUMAN]
P04083	3.76	1	1	346	38.7	7.02	70.05	Annexin A1 OS=Homo sapiens GN=ANXA1 PE=1 SV=2 - [ANXA1_HUMAN]
P08758	34.38	7	7	320	35.9	5.05	321.92	Annexin A5 OS=Homo sapiens GN=ANXA5 PE=1 SV=2 - [ANXA5_HUMAN]
P02647	15.73	3	3	267	30.8	5.76	220.49	Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1 - [APOA1_HUMAN]
P02652	20.00	1	1	100	11.2	6.62	70.43	Apolipoprotein A-II OS=Homo sapiens GN=APOA2 PE=1 SV=1 - [APOA2_HUMAN]
P05090	24.34	16	6	189	21.3	5.15	840.55	Apolipoprotein D OS=Homo sapiens GN=APOD PE=1 SV=1 - [APOD_HUMAN]
P02649	27.13	8	7	317	36.1	5.73	493.24	Apolipoprotein E OS=Homo sapiens GN=APOE PE=1 SV=1 - [APOE_HUMAN]
P15289	8.09	2	2	507	53.6	6.07	146.85	Arylsulfatase A OS=Homo sapiens GN=ARSA PE=1 SV=3 - [ARSA_HUMAN]
O75882	8.68	10	10	1429	158.4	7.31	517.68	Attractin OS=Homo sapiens GN=ATRN PE=1 SV=2 - [ATRN_HUMAN]
P98160	1.48	4	4	4391	468.5	6.51	261.62	Basement membrane-specific heparan sulfate proteoglycan core protein OS=Homo sapiens GN=HSPG2 PE=1 SV=3 - [PGBM_HUMAN]
P02749	9.28	2	2	345	38.3	7.97	108.24	Beta-2-glycoprotein 1 OS=Homo sapiens GN=APOH PE=1 SV=3 - [APOH_HUMAN]
P61769	8.40	1	1	119	13.7	6.52	69.68	Beta-2-microglobulin OS=Homo sapiens GN=B2M PE=1 SV=1 - [B2MG_HUMAN]
P16278	5.02	3	3	677	76.0	6.57	196.01	Beta-galactosidase OS=Homo sapiens GN=GLB1 PE=1 SV=2 - [BGAL_HUMAN]
P06865	2.08	1	1	529	60.7	5.16	72.45	Beta-hexosaminidase subunit alpha OS=Homo sapiens GN=HEXA PE=1 SV=1 - [HEXA_HUMAN]
P43251	7.92	3	3	543	61.1	6.25	161.44	Biotinidase OS=Homo sapiens GN=BTD PE=1 SV=2 - [BTD_HUMAN]
Q5VW32	3.89	1	1	411	46.4	7.65	65.53	BRO1 domain-containing protein BROX OS=Homo sapiens GN=BROX PE=1 SV=1 - [BROX_HUMAN]

P12830	5.44	3	3	882	97.4	4.73	157.63	Cadherin-1 OS=Homo sapiens GN=CDH1 PE=1 SV=3 - [CADH1_HUMAN]
P19022	6.18	3	3	906	99.7	4.81	177.43	Cadherin-2 OS=Homo sapiens GN=CDH2 PE=1 SV=4 - [CADH2_HUMAN]
Q9BYE9	2.06	2	2	1310	141.5	4.50	122.44	Cadherin-related family member 2 OS=Homo sapiens GN=CDHR2 PE=1 SV=2 - [CDHR2_HUMAN]
P14384	4.06	2	2	443	50.5	7.36	91.85	Carboxypeptidase M OS=Homo sapiens GN=CPM PE=1 SV=2 - [CBPM_HUMAN]
P22792	17.98	9	7	545	60.6	5.99	464.82	Carboxypeptidase N subunit 2 OS=Homo sapiens GN=CPN2 PE=1 SV=2 - [CPN2_HUMAN]
P07858	5.31	1	1	339	37.8	6.30	88.55	Cathepsin B OS=Homo sapiens GN=CTSB PE=1 SV=3 - [CATB_HUMAN]
P07339	21.84	6	6	412	44.5	6.54	345.17	Cathepsin D OS=Homo sapiens GN=CTSD PE=1 SV=1 - [CATD_HUMAN]
Q8N6Q3	3.43	1	1	437	46.3	6.29	58.93	CD177 antigen OS=Homo sapiens GN=CD177 PE=1 SV=2 - [CD177_HUMAN]
Q9Y5K6	2.19	1	1	639	71.4	6.40	111.79	CD2-associated protein OS=Homo sapiens GN=CD2AP PE=1 SV=1 - [CD2AP_HUMAN]
P16070	2.96	3	2	742	81.5	5.33	122.18	CD44 antigen OS=Homo sapiens GN=CD44 PE=1 SV=2 - [CD44_HUMAN]
P13987	25.00	5	4	128	14.2	6.48	232.29	CD59 glycoprotein OS=Homo sapiens GN=CD59 PE=1 SV=1 - [CD59_HUMAN]
P08962	4.20	1	1	238	25.6	7.81	82.96	CD63 antigen OS=Homo sapiens GN=CD63 PE=1 SV=2 - [CD63_HUMAN]
P21926	10.96	1	1	228	25.4	7.15	65.01	CD9 antigen OS=Homo sapiens GN=CD9 PE=1 SV=4 - [CD9_HUMAN]
P11597	3.04	1	1	493	54.7	6.09	59.48	Cholesteryl ester transfer protein OS=Homo sapiens GN=CETP PE=1 SV=2 - [CETP_HUMAN]
P10909	23.61	10	8	449	52.5	6.27	519.59	Clusterin OS=Homo sapiens GN=CLU PE=1 SV=1 - [CLUS_HUMAN]
Q6UXG3	4.82	1	1	332	36.0	5.92	97.17	CMRF35-like molecule 9 OS=Homo sapiens GN=CD300LG PE=1 SV=2 - [CLM9_HUMAN]
P12109	7.20	6	5	1028	108.5	5.43	231.90	Collagen alpha-1(VI) chain OS=Homo sapiens GN=COL6A1 PE=1 SV=3 - [CO6A1_HUMAN]
P39059	0.94	1	1	1388	141.6	5.00	104.31	Collagen alpha-1(XV) chain OS=Homo sapiens GN=COL15A1 PE=1 SV=2 - [COFA1_HUMAN]
P12111	2.64	8	6	3177	343.5	6.68	426.02	Collagen alpha-3(VI) chain OS=Homo sapiens GN=COL6A3 PE=1 SV=4 - [CO6A3_HUMAN]
P01024	0.96	1	1	1663	187.0	6.40	94.26	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2 - [CO3_HUMAN]
P0C0L4	0.86	1	1	1744	192.7	7.08	94.14	Complement C4-A OS=Homo sapiens GN=C4A PE=1 SV=1 - [CO4A_HUMAN]
P08174	2.10	1	1	381	41.4	7.59	93.84	Complement decay-accelerating factor OS=Homo sapiens GN=CD55 PE=1 SV=4 - [DAF_HUMAN]
P05156	4.97	2	2	583	65.7	7.50	110.32	Complement factor I OS=Homo sapiens GN=CFI PE=1 SV=1 - [CFAI_HUMAN]

P08185	8.89	3	3	405	45.1	6.04	121.20	Corticosteroid-binding globulin OS=Homo sapiens GN=SERPINA6 PE=1 SV=1 - [CBG_HUMAN]
O60494	11.40	26	26	3623	398.4	5.35	1386.85	Cubilin OS=Homo sapiens GN=CUBN PE=1 SV=4 - [CUBN_HUMAN]
P01040	30.61	2	2	98	11.0	5.50	122.24	Cystatin-A OS=Homo sapiens GN=CSTA PE=1 SV=1 - [CYTA_HUMAN]
P01034	7.53	1	1	146	15.8	8.75	85.99	Cystatin-C OS=Homo sapiens GN=CST3 PE=1 SV=1 - [CYTC_HUMAN]
Q15828	50.34	5	5	149	16.5	8.09	217.97	Cystatin-M OS=Homo sapiens GN=CST6 PE=1 SV=1 - [CYTM_HUMAN]
P24855	7.80	2	2	282	31.4	4.91	131.31	Deoxyribonuclease-1 OS=Homo sapiens GN=DNASE1 PE=1 SV=1 - [DNAS1_HUMAN]
P81605	12.73	1	1	110	11.3	6.54	89.46	Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2 - [DCD_HUMAN]
P15924	1.32	4	3	2871	331.6	6.81	256.24	Desmoplakin OS=Homo sapiens GN=DSP PE=1 SV=3 - [DESP_HUMAN]
P16444	43.07	16	13	411	45.6	6.15	791.26	Dipeptidase 1 OS=Homo sapiens GN=DPEP1 PE=1 SV=3 - [DPEP1_HUMAN]
P53634	11.23	4	4	463	51.8	6.99	233.92	Dipeptidyl peptidase 1 OS=Homo sapiens GN=CTSC PE=1 SV=1 - [CATC_HUMAN]
Q9UHL4	2.44	1	1	492	54.3	6.32	84.87	Dipeptidyl peptidase 2 OS=Homo sapiens GN=DPP7 PE=1 SV=3 - [DPP2_HUMAN]
P27487	2.48	2	2	766	88.2	6.04	210.81	Dipeptidyl peptidase 4 OS=Homo sapiens GN=DPP4 PE=1 SV=2 - [DPP4_HUMAN]
Q12805	28.60	12	10	493	54.6	5.07	548.10	EGF-containing fibulin-like extracellular matrix protein 1 OS=Homo sapiens GN=EFEMP1 PE=1 SV=2 - [FBLN3_HUMAN]
Q9HCU0	1.85	1	1	757	80.8	5.35	127.02	Endosialin OS=Homo sapiens GN=CD248 PE=1 SV=1 - [CD248_HUMAN]
Q9UNN8	12.61	2	2	238	26.7	7.18	124.53	Endothelial protein C receptor OS=Homo sapiens GN=PROCR PE=1 SV=1 - [EPCR_HUMAN]
P61916	10.60	1	1	151	16.6	7.65	69.34	Epididymal secretory protein E1 OS=Homo sapiens GN=NPC2 PE=1 SV=1 - [NPC2_HUMAN]
P02751	5.74	9	9	2386	262.5	5.71	491.06	Fibronectin OS=Homo sapiens GN=FN1 PE=1 SV=4 - [FINC_HUMAN]
Q9UBX5	10.04	5	4	448	50.1	4.73	218.41	Fibulin-5 OS=Homo sapiens GN=FBLN5 PE=1 SV=1 - [FBLN5_HUMAN]
Q5D862	0.46	1	1	2391	247.9	8.31	63.33	Filaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1 - [FILA2_HUMAN]
P15328	4.28	1	1	257	29.8	7.97	64.95	Folate receptor alpha OS=Homo sapiens GN=FOLR1 PE=1 SV=3 - [FOLR1_HUMAN]
P09467	5.03	1	1	338	36.8	6.99	66.39	Fructose-1,6-bisphosphatase 1 OS=Homo sapiens GN=FBP1 PE=1 SV=4 - [F16P1_HUMAN]
Q08380	28.38	16	11	585	65.3	5.27	947.38	Galectin-3-binding protein OS=Homo sapiens GN=LGALS3BP PE=1 SV=1 - [LG3BP_HUMAN]
Q92820	8.18	2	2	318	35.9	7.11	134.49	Gamma-glutamyl hydrolase OS=Homo sapiens GN=GGH

								PE=1 SV=2 - [GGH_HUMAN]
P19440	12.83	6	5	569	61.4	7.12	342.53	Gamma-glutamyltranspeptidase 1 OS=Homo sapiens GN=GGT1 PE=1 SV=2 - [GGT1_HUMAN]
P17900	15.54	2	2	193	20.8	5.31	97.81	Ganglioside GM2 activator OS=Homo sapiens GN=GM2A PE=1 SV=4 - [SAP3_HUMAN]
P06396	7.67	4	4	782	85.6	6.28	269.39	Gelsolin OS=Homo sapiens GN=GSN PE=1 SV=1 - [GELS_HUMAN]
Q16769	14.40	3	3	361	40.9	6.61	161.96	Glutaminyl-peptide cyclotransferase OS=Homo sapiens GN=QPCT PE=1 SV=1 - [QPCT_HUMAN]
P09211	16.67	2	2	210	23.3	5.64	118.20	Glutathione S-transferase P OS=Homo sapiens GN=GSTP1 PE=1 SV=2 - [GSTP1_HUMAN]
P51654	2.59	2	1	580	65.5	6.37	191.49	Glypican-3 OS=Homo sapiens GN=GPC3 PE=1 SV=1 - [GPC3_HUMAN]
Q8NBJ4	2.49	1	1	401	45.3	4.97	60.96	Golgi membrane protein 1 OS=Homo sapiens GN=GOLM1 PE=1 SV=1 - [GOLM1_HUMAN]
P28799	2.53	2	1	593	63.5	6.83	143.36	Granulins OS=Homo sapiens GN=GRN PE=1 SV=2 - [GRN_HUMAN]
Q8NHV1	2.00	1	1	300	34.5	6.46	67.91	GTPase IMAP family member 7 OS=Homo sapiens GN=GIMAP7 PE=2 SV=1 - [GIMA7_HUMAN]
P00738	11.08	5	4	406	45.2	6.58	261.27	Haptoglobin OS=Homo sapiens GN=HP PE=1 SV=1 - [HPT_HUMAN]
P11142	2.17	1	1	646	70.9	5.52	59.46	Heat shock cognate 71 kDa protein OS=Homo sapiens GN=HSPA8 PE=1 SV=1 - [HSP7C_HUMAN]
Q96RW7	0.44	2	2	5635	613.0	6.49	116.06	Hemicentin-1 OS=Homo sapiens GN=HMCN1 PE=1 SV=2 - [HMCN1_HUMAN]
P02790	14.72	5	4	462	51.6	7.02	239.97	Hemopexin OS=Homo sapiens GN=HPX PE=1 SV=2 - [HEMO_HUMAN]
P37235	6.22	1	1	193	22.3	5.35	81.02	Hippocalcin-like protein 1 OS=Homo sapiens GN=HPCAL1 PE=1 SV=3 - [HPCL1_HUMAN]
Q86YZ3	2.35	2	2	2850	282.2	10.04	159.14	Hornerin OS=Homo sapiens GN=HRNR PE=1 SV=2 - [HORN_HUMAN]
Q12794	19.08	4	4	435	48.3	6.77	175.75	Hyaluronidase-1 OS=Homo sapiens GN=HYAL1 PE=1 SV=2 - [HYAL1_HUMAN]
O75144	8.61	2	2	302	33.3	5.31	123.06	ICOS ligand OS=Homo sapiens GN=ICOSLG PE=1 SV=2 - [ICOSL_HUMAN]
P01876	9.07	3	3	353	37.6	6.51	218.37	Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=2 - [IGHA1_HUMAN]
P01877	9.41	3	3	340	36.5	6.10	167.35	Ig alpha-2 chain C region OS=Homo sapiens GN=IGHA2 PE=1 SV=3 - [IGHA2_HUMAN]
P01857	14.85	3	3	330	36.1	8.19	157.85	Ig gamma-1 chain C region OS=Homo sapiens GN=IGHG1 PE=1 SV=1 - [IGHG1_HUMAN]
P01861	14.98	3	3	327	35.9	7.36	153.39	Ig gamma-4 chain C region OS=Homo sapiens GN=IGHG4 PE=1 SV=1 - [IGHG4_HUMAN]
P01777	15.97	1	1	119	12.8	8.50	94.72	Ig heavy chain V-III region TEI OS=Homo sapiens PE=1 SV=1 - [HV316_HUMAN]

P01765	16.52	1	1	115	12.3	9.13	80.98	Ig heavy chain V-III region TIL OS=Homo sapiens PE=1 SV=1 - [HV304_HUMAN]
P01834	80.19	7	6	106	11.6	5.87	386.88	Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 SV=1 - [IGKC_HUMAN]
P01593	31.48	2	2	108	12.0	5.99	102.61	Ig kappa chain V-I region AG OS=Homo sapiens PE=1 SV=1 - [KV101_HUMAN]
P01617	32.74	2	2	113	12.3	6.00	119.05	Ig kappa chain V-II region TEW OS=Homo sapiens PE=1 SV=1 - [KV204_HUMAN]
P01620	39.45	3	3	109	11.8	8.48	151.14	Ig kappa chain V-III region SIE OS=Homo sapiens PE=1 SV=1 - [KV302_HUMAN]
P04434	7.76	1	1	116	12.7	5.94	65.50	Ig kappa chain V-III region VH (Fragment) OS=Homo sapiens PE=4 SV=1 - [KV310_HUMAN]
P01625	23.68	3	2	114	12.6	7.93	124.44	Ig kappa chain V-IV region Len OS=Homo sapiens PE=1 SV=2 - [KV402_HUMAN]
P0CG05	47.17	4	3	106	11.3	7.24	190.30	Ig lambda-2 chain C regions OS=Homo sapiens GN=IGLC2 PE=1 SV=1 - [LAC2_HUMAN]
P01591	7.55	1	1	159	18.1	5.24	120.62	Immunoglobulin J chain OS=Homo sapiens GN=IGJ PE=1 SV=4 - [IGJ_HUMAN]
P18065	3.69	1	1	325	34.8	7.50	63.26	Insulin-like growth factor-binding protein 2 OS=Homo sapiens GN=IGFBP2 PE=1 SV=2 - [IBP2_HUMAN]
Q16270	17.02	3	3	282	29.1	7.90	167.76	Insulin-like growth factor-binding protein 7 OS=Homo sapiens GN=IGFBP7 PE=1 SV=1 - [IBP7_HUMAN]
Q14624	6.88	5	4	930	103.3	6.98	253.53	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4 - [ITIH4_HUMAN]
P53990	2.20	1	1	364	39.7	5.35	80.70	IST1 homolog OS=Homo sapiens GN=KIAA0174 PE=1 SV=1 - [IST1_HUMAN]
P06870	16.03	3	3	262	28.9	4.83	182.31	Kallikrein-1 OS=Homo sapiens GN=KLK1 PE=1 SV=2 - [KLK1_HUMAN]
P13645	50.86	38	23	584	58.8	5.21	1956.71	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6 - [K1C10_HUMAN]
P13646	5.02	2	2	458	49.6	4.96	158.12	Keratin, type I cytoskeletal 13 OS=Homo sapiens GN=KRT13 PE=1 SV=4 - [K1C13_HUMAN]
P02533	33.69	13	12	472	51.5	5.16	680.04	Keratin, type I cytoskeletal 14 OS=Homo sapiens GN=KRT14 PE=1 SV=4 - [K1C14_HUMAN]
P08779	38.27	17	16	473	51.2	5.05	769.78	Keratin, type I cytoskeletal 16 OS=Homo sapiens GN=KRT16 PE=1 SV=4 - [K1C16_HUMAN]
Q04695	16.44	6	6	432	48.1	5.02	364.57	Keratin, type I cytoskeletal 17 OS=Homo sapiens GN=KRT17 PE=1 SV=2 - [K1C17_HUMAN]
P35527	54.90	27	21	623	62.0	5.24	1631.29	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3 - [K1C9_HUMAN]
P04264	46.27	39	26	644	66.0	8.12	1899.51	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6 - [K2C1_HUMAN]
P35908	46.48	26	24	639	65.4	8.00	1556.88	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2 - [K22E_HUMAN]

P12035	6.20	5	4	629	64.5	6.48	282.40	Keratin, type II cytoskeletal 3 OS=Homo sapiens GN=KRT3 PE=1 SV=2 - [K2C3_HUMAN]
P19013	3.75	3	2	534	57.2	6.61	132.19	Keratin, type II cytoskeletal 4 OS=Homo sapiens GN=KRT4 PE=1 SV=4 - [K2C4_HUMAN]
P13647	29.15	18	17	590	62.3	7.74	794.02	Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 PE=1 SV=3 - [K2C5_HUMAN]
P02538	30.67	19	18	564	60.0	8.00	850.05	Keratin, type II cytoskeletal 6A OS=Homo sapiens GN=KRT6A PE=1 SV=3 - [K2C6A_HUMAN]
P04259	27.84	19	17	564	60.0	8.00	844.08	Keratin, type II cytoskeletal 6B OS=Homo sapiens GN=KRT6B PE=1 SV=5 - [K2C6B_HUMAN]
P48668	30.67	19	18	564	60.0	8.00	843.15	Keratin, type II cytoskeletal 6C OS=Homo sapiens GN=KRT6C PE=1 SV=3 - [K2C6C_HUMAN]
Q8N1N4	3.85	2	2	520	56.8	6.02	100.64	Keratin, type II cytoskeletal 78 OS=Homo sapiens GN=KRT78 PE=1 SV=2 - [K2C78_HUMAN]
P01042	18.32	12	10	644	71.9	6.81	670.29	Kininogen-1 OS=Homo sapiens GN=KNG1 PE=1 SV=2 - [KNG1_HUMAN]
Q9H756	3.51	1	1	370	42.3	5.12	66.66	Leucine-rich repeat-containing protein 19 OS=Homo sapiens GN=LRRC19 PE=2 SV=1 - [LRC19_HUMAN]
Q6GTX8	4.53	1	1	287	31.4	5.63	97.51	Leukocyte-associated immunoglobulin-like receptor 1 OS=Homo sapiens GN=LAIR1 PE=1 SV=1 - [LAIR1_HUMAN]
P05451	6.02	1	1	166	18.7	5.94	64.64	Lithostathine-1-alpha OS=Homo sapiens GN=REG1A PE=1 SV=3 - [REG1A_HUMAN]
P00338	4.52	1	1	332	36.7	8.27	106.81	L-lactate dehydrogenase A chain OS=Homo sapiens GN=LDHA PE=1 SV=2 - [LDHA_HUMAN]
P07195	3.29	2	1	334	36.6	6.05	136.07	L-lactate dehydrogenase B chain OS=Homo sapiens GN=LDHB PE=1 SV=2 - [LDHB_HUMAN]
P98164	6.51	25	23	4655	521.6	5.08	1213.09	Low-density lipoprotein receptor-related protein 2 OS=Homo sapiens GN=LRP2 PE=1 SV=3 - [LRP2_HUMAN]
P11117	11.11	4	4	423	48.3	6.74	230.34	Lysosomal acid phosphatase OS=Homo sapiens GN=ACP2 PE=1 SV=3 - [PPAL_HUMAN]
P10253	18.49	17	12	952	105.3	5.99	788.99	Lysosomal alpha-glucosidase OS=Homo sapiens GN=GAA PE=1 SV=3 - [LYAG_HUMAN]
P10619	5.63	2	2	480	54.4	6.61	234.31	Lysosomal protective protein OS=Homo sapiens GN=CTSA PE=1 SV=2 - [PPGB_HUMAN]
P42785	7.06	3	3	496	55.8	7.21	117.18	Lysosomal Pro-X carboxypeptidase OS=Homo sapiens GN=PRCP PE=1 SV=1 - [PCP_HUMAN]
P11279	3.84	1	1	417	44.9	8.75	100.31	Lysosome-associated membrane glycoprotein 1 OS=Homo sapiens GN=LAMP1 PE=1 SV=3 - [LAMP1_HUMAN]
P13473	9.27	3	3	410	44.9	5.63	166.00	Lysosome-associated membrane glycoprotein 2 OS=Homo sapiens GN=LAMP2 PE=1 SV=2 - [LAMP2_HUMAN]
P09603	5.78	2	2	554	60.1	5.29	112.68	Macrophage colony-stimulating factor 1 OS=Homo sapiens GN=CSF1 PE=1 SV=1 - [CSF1_HUMAN]

P40925	7.78	2	2	334	36.4	7.36	126.16	Malate dehydrogenase, cytoplasmic OS=Homo sapiens GN=MDH1 PE=1 SV=4 - [MDHC_HUMAN]
O43451	6.41	7	7	1857	209.7	5.50	596.95	Maltase-glucoamylase, intestinal OS=Homo sapiens GN=MGAM PE=1 SV=5 - [MGA_HUMAN]
O00187	5.39	3	3	686	75.7	5.77	131.45	Mannan-binding lectin serine protease 2 OS=Homo sapiens GN=MASP2 PE=1 SV=3 - [MASP2_HUMAN]
P33908	1.84	1	1	653	72.9	6.47	58.32	Mannosyl-oligosaccharide 1,2-alpha-mannosidase IA OS=Homo sapiens GN=MAN1A1 PE=1 SV=3 - [MA1A1_HUMAN]
P26038	1.56	1	1	577	67.8	6.40	61.39	Moesin OS=Homo sapiens GN=MSN PE=1 SV=3 - [MOES_HUMAN]
P15941	3.11	3	3	1255	122.0	7.47	129.17	Mucin-1 OS=Homo sapiens GN=MUC1 PE=1 SV=3 - [MUC1_HUMAN]
Q9H8L6	5.27	5	3	949	104.3	5.86	293.31	Multimerin-2 OS=Homo sapiens GN=MMRN2 PE=1 SV=2 - [MMRN2_HUMAN]
Q9UNW1	8.83	3	3	487	55.0	7.81	157.16	Multiple inositol polyphosphate phosphatase 1 OS=Homo sapiens GN=MINPP1 PE=1 SV=1 - [MINP1_HUMAN]
P34059	7.66	2	2	522	58.0	6.74	109.55	N-acetylgalactosamine-6-sulfatase OS=Homo sapiens GN=GALNS PE=1 SV=1 - [GALNS_HUMAN]
P15586	3.26	2	2	552	62.0	8.31	112.45	N-acetylglucosamine-6-sulfatase OS=Homo sapiens GN=GNS PE=1 SV=3 - [GNS_HUMAN]
O96009	13.33	5	4	420	45.4	6.61	246.37	Napsin-A OS=Homo sapiens GN=NAPSA PE=1 SV=1 - [NAPSA_HUMAN]
P08473	15.73	8	8	750	85.5	5.73	503.40	Neprilysin OS=Homo sapiens GN=MME PE=1 SV=2 - [NEP_HUMAN]
Q99574	2.44	1	1	410	46.4	4.91	79.31	Neuroserpin OS=Homo sapiens GN=SERPINI1 PE=1 SV=1 - [NEUS_HUMAN]
P59665	9.57	1	1	94	10.2	6.99	98.17	Neutrophil defensin 1 OS=Homo sapiens GN=DEFA1 PE=1 SV=1 - [DEF1_HUMAN]
P10153	10.56	2	1	161	18.3	8.73	167.89	Non-secretory ribonuclease OS=Homo sapiens GN=RNASE2 PE=1 SV=2 - [RNAS2_HUMAN]
P61970	29.13	2	2	127	14.5	5.38	112.01	Nuclear transport factor 2 OS=Homo sapiens GN=NUTF2 PE=1 SV=1 - [NTF2_HUMAN]
P50897	9.15	2	2	306	34.2	6.52	102.00	Palmitoyl-protein thioesterase 1 OS=Homo sapiens GN=PPT1 PE=1 SV=1 - [PPT1_HUMAN]
P04746	10.37	4	4	511	57.7	7.05	330.28	Pancreatic alpha-amylase OS=Homo sapiens GN=AMY2A PE=1 SV=2 - [AMYP_HUMAN]
O95497	10.33	4	4	513	57.0	5.55	197.26	Pantetheinase OS=Homo sapiens GN=VNN1 PE=1 SV=2 - [VNN1_HUMAN]
Q9UBV8	3.87	1	1	284	30.4	6.54	115.32	Peflin OS=Homo sapiens GN=PEF1 PE=1 SV=1 - [PEF1_HUMAN]
O75594	39.29	6	4	196	21.7	8.59	338.88	Peptidoglycan recognition protein 1 OS=Homo sapiens GN=PGLYRP1 PE=1 SV=1 - [PGRP1_HUMAN]
P30041	11.61	2	2	224	25.0	6.38	99.97	Peroxiredoxin-6 OS=Homo sapiens GN=PRDX6 PE=1 SV=3 -

								[PRDX6_HUMAN]
P04180	3.18	1	1	440	49.5	6.11	80.92	Phosphatidylcholine-sterol acyltransferase OS=Homo sapiens GN=LCAT PE=1 SV=1 - [LCAT_HUMAN]
P30086	18.18	2	2	187	21.0	7.53	143.21	Phosphatidylethanolamine-binding protein 1 OS=Homo sapiens GN=PEBP1 PE=1 SV=3 - [PEBP1_HUMAN]
Q96FE7	5.70	1	1	263	28.2	5.01	96.38	Phosphoinositide-3-kinase-interacting protein 1 OS=Homo sapiens GN=PIK3IP1 PE=1 SV=2 - [P3IP1_HUMAN]
P05155	15.40	10	6	500	55.1	6.55	434.63	Plasma protease C1 inhibitor OS=Homo sapiens GN=SERPING1 PE=1 SV=2 - [IC1_HUMAN]
P05154	19.21	9	7	406	45.7	9.26	440.96	Plasma serine protease inhibitor OS=Homo sapiens GN=SERPINA5 PE=1 SV=2 - [IPSP_HUMAN]
P13796	3.83	2	2	627	70.2	5.33	91.37	Plastin-2 OS=Homo sapiens GN=LCP1 PE=1 SV=5 - [PLSL_HUMAN]
P01833	29.45	16	15	764	83.2	5.74	881.01	Polymeric immunoglobulin receptor OS=Homo sapiens GN=PIGR PE=1 SV=4 - [PIGR_HUMAN]
P0CG48	21.02	1	1	685	77.0	7.66	124.57	Polyubiquitin-C OS=Homo sapiens GN=UBC PE=1 SV=1 - [UBC_HUMAN]
P07602	11.64	8	5	524	58.1	5.17	331.77	Proactivator polypeptide OS=Homo sapiens GN=PSAP PE=1 SV=2 - [SAP_HUMAN]
Q9H3G5	2.52	1	1	476	54.1	5.62	71.60	Probable serine carboxypeptidase CPVL OS=Homo sapiens GN=CPVL PE=1 SV=2 - [CPVL_HUMAN]
P01133	20.30	19	17	1207	133.9	5.85	1093.64	Pro-epidermal growth factor OS=Homo sapiens GN=EGF PE=1 SV=2 - [EGF_HUMAN]
Q8WUM4	3.23	2	2	868	96.0	6.52	136.52	Programmed cell death 6-interacting protein OS=Homo sapiens GN=PDCD6IP PE=1 SV=1 - [PDC6I_HUMAN]
O75340	26.18	3	3	191	21.9	5.40	147.99	Programmed cell death protein 6 OS=Homo sapiens GN=PDCD6 PE=1 SV=1 - [PDCD6_HUMAN]
P12273	22.60	4	3	146	16.6	8.05	252.60	Prolactin-inducible protein OS=Homo sapiens GN=PIP PE=1 SV=1 - [PIP_HUMAN]
P41222	24.74	7	4	190	21.0	7.80	348.74	Prostaglandin-H2 D-isomerase OS=Homo sapiens GN=PTGDS PE=1 SV=1 - [PTGDS_HUMAN]
Q16651	11.66	3	3	343	36.4	5.85	103.73	Prostasin OS=Homo sapiens GN=PRSS8 PE=1 SV=1 - [PRSS8_HUMAN]
O43653	21.95	2	2	123	12.9	5.29	154.17	Prostate stem cell antigen OS=Homo sapiens GN=PSCA PE=1 SV=1 - [PSCA_HUMAN]
P15309	20.47	6	6	386	44.5	6.24	299.53	Prostatic acid phosphatase OS=Homo sapiens GN=ACPP PE=1 SV=3 - [PPAP_HUMAN]
Q06323	4.82	1	1	249	28.7	6.02	72.70	Proteasome activator complex subunit 1 OS=Homo sapiens GN=PSME1 PE=1 SV=1 - [PSME1_HUMAN]
P02760	43.18	22	12	352	39.0	6.25	1147.02	Protein AMBP OS=Homo sapiens GN=AMBP PE=1 SV=1 - [AMBP_HUMAN]

P80370	2.09	1	1	383	41.3	5.67	63.89	Protein delta homolog 1 OS=Homo sapiens GN=DLK1 PE=1 SV=3 - [DLK1_HUMAN]
P31949	15.24	1	1	105	11.7	7.12	90.24	Protein S100-A11 OS=Homo sapiens GN=S100A11 PE=1 SV=2 - [S10AB_HUMAN]
Q96FQ6	10.68	1	1	103	11.8	6.79	66.20	Protein S100-A16 OS=Homo sapiens GN=S100A16 PE=1 SV=1 - [S10AG_HUMAN]
P06703	8.89	1	1	90	10.2	5.48	84.12	Protein S100-A6 OS=Homo sapiens GN=S100A6 PE=1 SV=1 - [S10A6_HUMAN]
P31151	10.89	1	1	101	11.5	6.77	60.94	Protein S100-A7 OS=Homo sapiens GN=S100A7 PE=1 SV=4 - [S10A7_HUMAN]
P05109	11.83	1	1	93	10.8	7.03	91.60	Protein S100-A8 OS=Homo sapiens GN=S100A8 PE=1 SV=1 - [S10A8_HUMAN]
P06702	56.14	6	5	114	13.2	6.13	270.52	Protein S100-A9 OS=Homo sapiens GN=S100A9 PE=1 SV=1 - [S10A9_HUMAN]
P25815	13.68	1	1	95	10.4	4.88	74.10	Protein S100-P OS=Homo sapiens GN=S100P PE=1 SV=2 - [S100P_HUMAN]
P00734	11.09	7	5	622	70.0	5.90	343.90	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2 - [THRB_HUMAN]
Q12913	0.90	1	1	1337	145.9	5.58	59.90	Receptor-type tyrosine-protein phosphatase eta OS=Homo sapiens GN=PTPRJ PE=1 SV=3 - [PTPRJ_HUMAN]
Q13332	1.23	2	2	1948	217.0	6.51	73.14	Receptor-type tyrosine-protein phosphatase S OS=Homo sapiens GN=PTPRS PE=1 SV=2 - [PTPRS_HUMAN]
O75787	2.57	1	1	350	39.0	6.10	84.90	Renin receptor OS=Homo sapiens GN=ATP6AP2 PE=1 SV=2 - [RENR_HUMAN]
Q9HD89	24.07	2	2	108	11.4	6.86	108.16	Resistin OS=Homo sapiens GN=RETN PE=2 SV=1 - [RETN_HUMAN]
Q9HB40	7.96	4	3	452	50.8	5.81	226.38	Retinoid-inducible serine carboxypeptidase OS=Homo sapiens GN=SCPEP1 PE=1 SV=1 - [RISC_HUMAN]
P02753	11.44	1	1	201	23.0	6.07	68.77	Retinol-binding protein 4 OS=Homo sapiens GN=RBP4 PE=1 SV=3 - [RET4_HUMAN]
P07998	43.59	3	3	156	17.6	8.79	145.18	Ribonuclease pancreatic OS=Homo sapiens GN=RNASE1 PE=1 SV=4 - [RNAS1_HUMAN]
Q8WVN6	14.52	3	2	248	27.0	7.43	206.16	Secreted and transmembrane protein 1 OS=Homo sapiens GN=SECTM1 PE=1 SV=2 - [SCTM1_HUMAN]
P55000	24.27	1	1	103	11.2	5.33	83.12	Secreted Ly-6/uPAR-related protein 1 OS=Homo sapiens GN=SLURP1 PE=1 SV=2 - [SLUR1_HUMAN]
Q86SR0	45.36	2	2	97	10.2	6.62	121.44	Secreted Ly-6/uPAR-related protein 2 OS=Homo sapiens GN=SLURP2 PE=2 SV=1 - [SLUR2_HUMAN]
P04279	7.36	2	2	462	52.1	9.29	99.57	Semenogelin-1 OS=Homo sapiens GN=SEMG1 PE=1 SV=2 - [SEMG1_HUMAN]
P02787	5.73	3	3	698	77.0	7.12	206.86	Serotransferrin OS=Homo sapiens GN=TF PE=1 SV=2 - [TRFE_HUMAN]
P02768	62.40	53	33	609	69.3	6.28	2536.79	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2 -

								[ALBU_HUMAN]
Q9HAT2	5.16	2	2	523	58.3	7.33	116.86	Sialate O-acetylesterase OS=Homo sapiens GN=SIAE PE=1 SV=1 - [SIAE_HUMAN]
Q99519	3.37	1	1	415	45.4	5.88	72.82	Sialidase-1 OS=Homo sapiens GN=NEU1 PE=1 SV=1 - [NEUR1_HUMAN]
P30626	9.09	2	2	198	21.7	5.59	104.80	Sorcin OS=Homo sapiens GN=SRI PE=1 SV=1 - [SORCN_HUMAN]
Q9UGT4	8.39	3	3	822	90.1	6.28	140.38	Sushi domain-containing protein 2 OS=Homo sapiens GN=SUSD2 PE=1 SV=1 - [SUSD2_HUMAN]
O60635	5.39	1	1	241	26.3	5.25	70.21	Tetraspanin-1 OS=Homo sapiens GN=TSPAN1 PE=1 SV=2 - [TSN1_HUMAN]
P10599	20.95	2	2	105	11.7	4.92	124.20	Thioredoxin OS=Homo sapiens GN=TXN PE=1 SV=3 - [THIO_HUMAN]
P07204	3.48	1	1	575	60.3	4.92	77.65	Thrombomodulin OS=Homo sapiens GN=THBD PE=1 SV=2 - [TRBM_HUMAN]
P04216	18.01	2	2	161	17.9	8.73	136.44	Thy-1 membrane glycoprotein OS=Homo sapiens GN=THY1 PE=1 SV=2 - [THY1_HUMAN]
P05543	6.02	2	2	415	46.3	6.30	113.65	Thyroxine-binding globulin OS=Homo sapiens GN=SERPINA7 PE=1 SV=2 - [THBG_HUMAN]
P20062	3.75	1	1	427	47.5	7.01	89.75	Transcobalamin-2 OS=Homo sapiens GN=TCN2 PE=1 SV=3 - [TCO2_HUMAN]
P02766	47.62	4	4	147	15.9	5.76	178.78	Transthyretin OS=Homo sapiens GN=TTR PE=1 SV=1 - [TTHY_HUMAN]
O14773	10.48	3	3	563	61.2	6.48	136.25	Tripeptidyl-peptidase 1 OS=Homo sapiens GN=TPP1 PE=1 SV=2 - [TPP1_HUMAN]
Q969Z4	2.56	2	2	430	46.1	8.35	127.20	Tumor necrosis factor receptor superfamily member 19L OS=Homo sapiens GN=RELT PE=1 SV=1 - [TR19L_HUMAN]
P30530	3.36	2	2	894	98.3	5.43	112.99	Tyrosine-protein kinase receptor UFO OS=Homo sapiens GN=AXL PE=1 SV=3 - [UFO_HUMAN]
Q6UX73	11.69	3	3	402	45.4	6.19	113.58	UPF0764 protein C16orf89 OS=Homo sapiens GN=C16orf89 PE=2 SV=2 - [CP089_HUMAN]
P07911	30.16	27	14	640	69.7	5.24	1573.47	Uromodulin OS=Homo sapiens GN=UMOD PE=1 SV=1 - [UROM_HUMAN]
P11684	12.09	2	2	91	10.0	5.06	100.09	Uteroglobin OS=Homo sapiens GN=SCGB1A1 PE=1 SV=1 - [UTER_HUMAN]
Q9NP79	6.51	1	1	307	33.9	6.29	78.64	Vacuolar protein sorting-associated protein VTA1 homolog OS=Homo sapiens GN=VTA1 PE=1 SV=1 - [VTA1_HUMAN]
Q6EMK4	7.88	5	4	673	71.7	7.39	299.15	Vasorin OS=Homo sapiens GN=VASN PE=1 SV=1 - [VASN_HUMAN]
Q12907	5.90	2	2	356	40.2	6.95	102.94	Vesicular integral-membrane protein VIP36 OS=Homo sapiens GN=LMAN2 PE=1 SV=1 - [LMAN2_HUMAN]
P22891	11.25	3	3	400	44.7	5.97	233.11	Vitamin K-dependent protein Z OS=Homo sapiens GN=PROZ PE=1 SV=2 - [PROZ_HUMAN]
Q7Z5L0	32.67	4	4	202	21.5	5.07	231.32	Vitelline membrane outer layer protein 1 homolog OS=Homo

								sapiens GN=VMO1 PE=1 SV=1 - [VMO1_HUMAN]
P04004	10.88	4	4	478	54.3	5.80	192.37	Vitronectin OS=Homo sapiens GN=VTN PE=1 SV=1 - [VTNC_HUMAN]
Q14508	38.71	4	3	124	13.0	4.84	239.68	WAP four-disulfide core domain protein 2 OS=Homo sapiens GN=WFDC2 PE=1 SV=2 - [WFDC2_HUMAN]
O43895	19.58	9	9	674	75.6	6.04	592.18	Xaa-Pro aminopeptidase 2 OS=Homo sapiens GN=XPNPEP2 PE=1 SV=3 - [XPP2_HUMAN]
P25311	19.13	5	5	298	34.2	6.05	225.11	Zinc-alpha-2-glycoprotein OS=Homo sapiens GN=AZGP1 PE=1 SV=2 - [ZA2G_HUMAN]
Q96DA0	25.96	4	4	208	22.7	7.39	271.33	Zymogen granule protein 16 homolog B OS=Homo sapiens GN=ZG16B PE=1 SV=3 - [ZG16B_HUMAN]

CHAPTER 3

PROTEOMIC ANALYSIS OF URINARY MEMBRANE VESICLES ISOLATED BY A MODIFIED DIFFERENTIAL CENTRIFUGATION METHOD

3.1 Introduction

Building on the previous chapter, it was considered that after defining a method for enrichment of urinary membrane vesicles and removal of soluble protein contaminants, the obvious next step is the proteomic characterization of the vesicles enriched. Exosomes and other types of vesicles present in urine are an important source of potentially valuable biomarkers and a thorough knowledge of their proteome is an absolute requirement. There was one previous study with large-scale proteomic profiling of the exosomes/membrane vesicles isolated from urine using the differential centrifugation method and removal of soluble contaminants like THP using DTT treatment followed by a second centrifugation (Gonzales *et al.,* 2009). Based on this approach Gonzalez *et al.,* (2009) were able to identify 1,132 proteins altogether unambiguously, including 14 phosphoproteins. Another very recent study published has expanded the proteome coverage of urinary exosomes/membrane vesicles with up to 3,280 proteins identified (Wang *et al.*, 2011). This study compared two methods of protein digestion one being trifluoroethanol solution phase digestion and the other being 'in-gel' digestion. It was concluded that both the digestion protocols were not susbtantially different from each other in terms of proteins identified. Furthermore, a new generation of high efficiency ion-trap instrumentation (LTQ velos, Thermo Fisher Scientific, USA) was compared to the older LTQ. It was found that Velos performed better than LTQ in increasing proteome coverage. With our method which is efficient in removing high contaminant proteins from the exosomes/membrane vesicles, it is hoped that more low abundance proteins are identified increasing the proteome coverage of these vesicles from urine. However, we did not take a multidimensional separation approach for resolving peptides. This would mean a lesser number of identifications compared to some of the other studies. However, here, we want to develop an approach which can be applied to a large number of samples such as encountered in a clinical setting. A multi-dimensional approach

applied to a large number of samples would not work. Our approach has a significantly reduced number of steps and it can be expanded in future to include quantitative analysis of the membrane vesicular proteome. Here we report the proteomic characterisation of the high-speed pellet (exosomal fraction, P200,000g) and low speed pellet (urine sediment and possibly large size membrane vesicles, P18,000g) prepared using an enrichment protocol established in our lab, as discussed in the previous chapter. The high speed pellet proteomic identifications are compared to most of the studies reported in the literature and low speed pellet identification (due to the lack of proteomic studies on large membrane vesicles which pellet down at low speed) is compared between the two methods of enrichment (DTT and CHAPS treatment of P18,000g).

3.2 Material and methods

3.2.1 Isolation of CHAPS detergent resistant nano vesicles

A schematic representation of the methodology used to isolate vesicles is shown in Figure 3.1. The detailed method used for isolation of membrane vesicles is reported in methods section of Chapter 2 (Section 2.2.2).

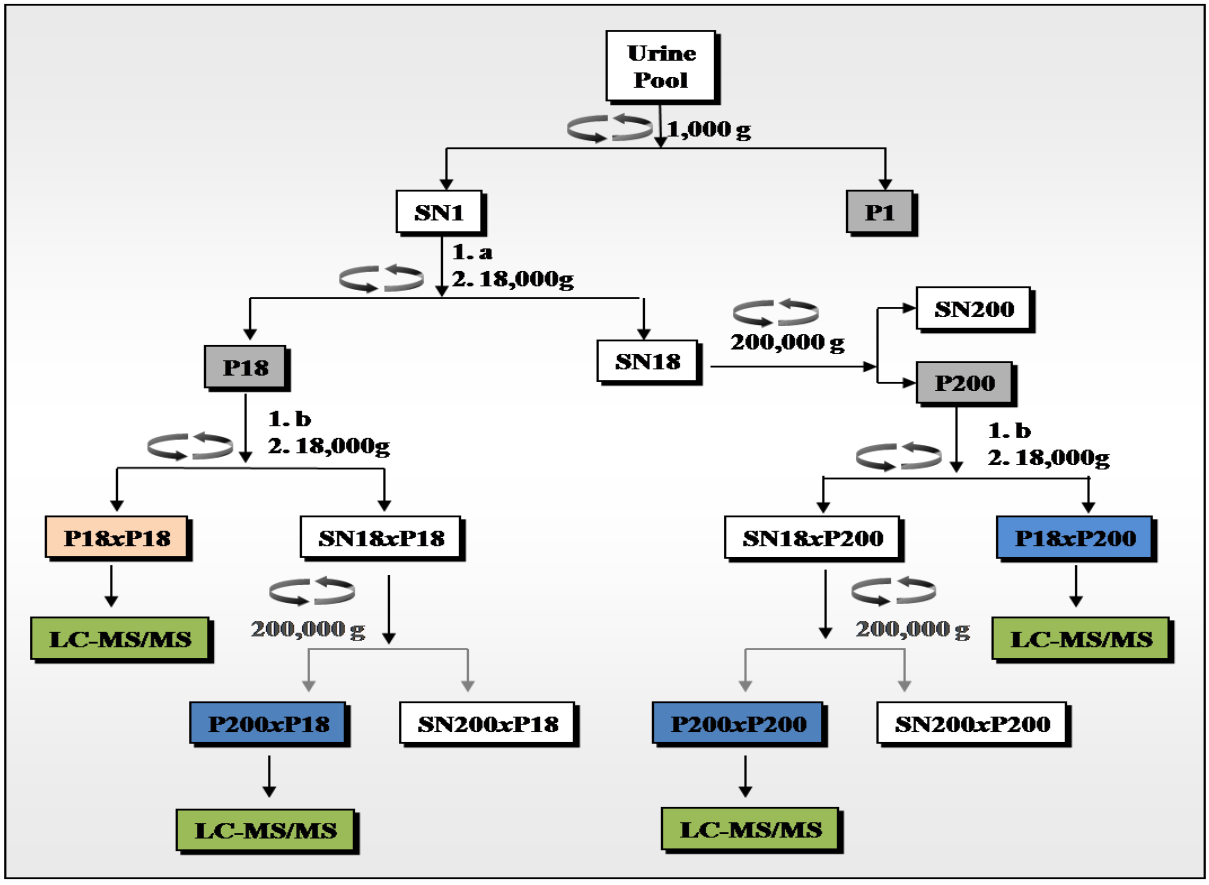

Figure 3.1: Workflow for isolation of nano vesicles. The samples highlighted in blue were high speed pellets (P200,000g) selected for proteomic analysis using mass spectrometry and samples highlighted in orange were low speed pellet (P18,000g) selected for proteomic analysis. a: Dialysis and concentration, b: DTT/CHAPS treatment, x: DTT/CHAPS.

3.2.2 LC_MS/MS analysis and Data Analysis

This was performed as described in methods section of Chapter 2 (Sections 2.2.5 and 2.2.6).

3.2.3 Bioinformatic analysis

Bioinformatic analysis was performed using IPA software (Ingenuity systems, Mountain View, CA). Guest trial access to IPA was obtained. All the Tables were made manually and gene ontology figures were created in the MS Excel program. David bioinformatic resources 6.7 (Huang, Sherman & Lempicki, 2009a; Huang, Sherman & Lempicki, 2009b) and Blast2Go software (Conesa *et al.,* 2005) were also used for bioinformatic analysis as indicated in sections of this chapter (section 3.3.1 and sub-sections therein).

3.3 Results

The workflow used for enrichment of exosomes/microveicles is shown in Figure 3.1 and the fractions selected for proteomic analysis using Orbitrap LTQ are highlighted. There are mainly two types of vesicles enriched one being in high-speed pellet (Blue in Figure 3.1) and other in low speed pellet (Orange in Figure 3.1). High speed pellet contains smaller vesicles (exosomes/ectosomes) and low speed pellet most likely consist of bigger sized vesicles (membrane vesicles). Although some small vesicles may be trapped in THP polymers and pellet down at low speed but, as discussed in the previous chapter, CHAPS or DTT treatment helps reduce this problem of interference.

3.3.1 Large-scale proteomic analysis of the high speed pellet

We were able to identify 437 proteins unambiguously in the high speed pellet. The proteins identified are listed in Supplementary Table S3.1.

3.3.1.1 Comparison with other exosome/microvesicle and urine proteomics studies

There have been a few proteomic studies characterising urinary exosomes (Gonzales *et al.,* 2009) (Wang *et al.,* 2011) and exosome like vesicles (Hogan *et al.,* 2009) for their protein content. We compare, here, proteomic identifications of our analysis with that of these studies as well as with the largest proteomic study on whole urine (Marimuthu *et al.,* 2011). This comparison validates the identification of proteins in our analysis. We also compare our identifications with the the largest database of human exosomal proteins purified from any source (Exocarta) (Mathivanan *et al.,* 2012).

Figure 3.2 compares our proteomic identifications (P200,000g) with Pistikun *et al.,* (Pisitkun, Shen & Knepper, 2004) and Gonzalez *et al.,* (Gonzales *et al.,* 2009) (both studies have been combined here non-redundantly and given the name Knepper), Hogan *et al.,* (Hogan *et al.,* 2009), and Marimuthu *et al.,* (Marimuthu *et al.,* 2011) The first two studies (combined as

Knepper) are on urinary exosomes isolated using the differential centrifugation protocol. The second study (Hogan) is on urinary 'exosome-like' vesicles which were fractionated further by a sucrose density gradient using differential centrifugation-derived exosomal pellets as the starting material and segregated based on the prevalence of three biomakers among the fractions (aquaporin-2, polycystin-1, and podocin). The last one (Marimuthu) is a large-scale proteomic study on whole urine.

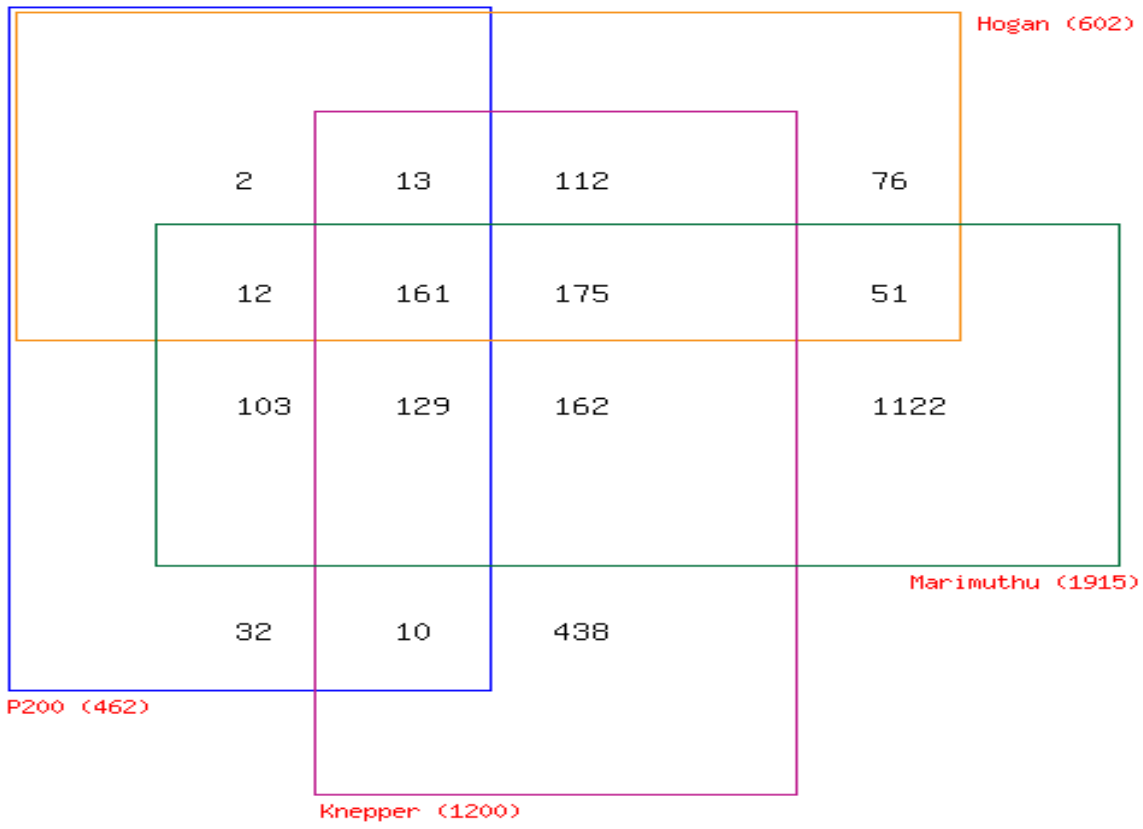

Figure 3.2: Comparison of various urinary exosomes (Knepper) (Gonzales *et al.*, 2009; Pisitkun, Shen & Knepper, 2004) and 'exosome-like' vesicles (Hogan) (Hogan *et al.*, 2009) and whole urine (Marimuthu) (Marimuthu *et al.*, 2011) with our protein list. To enable comparison the various identifiers reported in the studies (Entrez Id, uniprot accessions, gene name, Uniprot Id) were converted to unigene identifiers using a batch retrieval service of Uniprot (http://www.uniprot.org/).

Hundred and sixty one proteins are common to all four studies while different numbers come up when comparing the two studies, as can be seen in Figure 3.2,. Hogan and Knepper are studies of urinary exosomes and 'exosome-like' vesicles and, when compared with our study, 135 proteins (~30%) in our study are new. Out of these 135, 103 have been identified in urine previously (Marimuthu) while 32 are reported here for the first time in urine.

Very recently another study has been published on urinary exosomes which has identified 3,280 proteins confidently (Wang *et al.*, 2011). Comparison of our protein list to this study is shown in Figure 3.3.

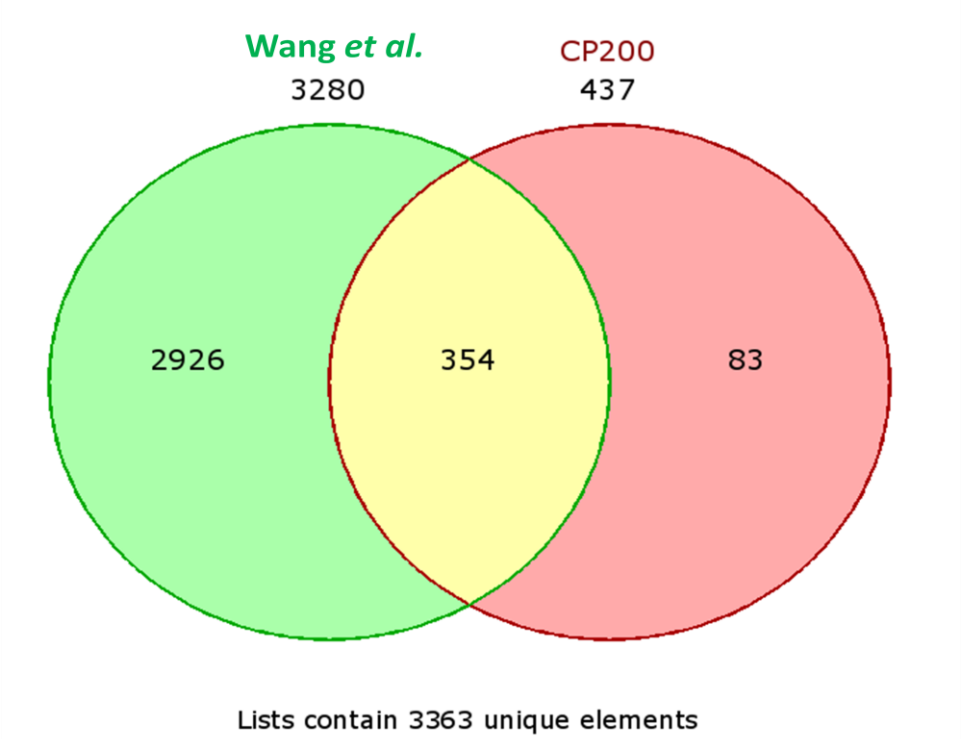

Figure 3.3: Venn diagram comparing our protein list to Wang *et al.*, (Wang *et al.*, 2011), which is the largest study on urinary exosomes to date.

The protein list in our study and Wang *et al.*, (Wang *et al.*, 2011) was directly compared as both had Uniprot accession as identifiers so no conversion to other gene identifiers was needed. 83 proteins (19%) were unique to our study.

When our protein list was compared to Exocarta (Mathivanan *et al.*, 2012), which is the largest database for exosomal proteins from any cell type or body fluid, 86 proteins (~20%) were found to be unique to our study.

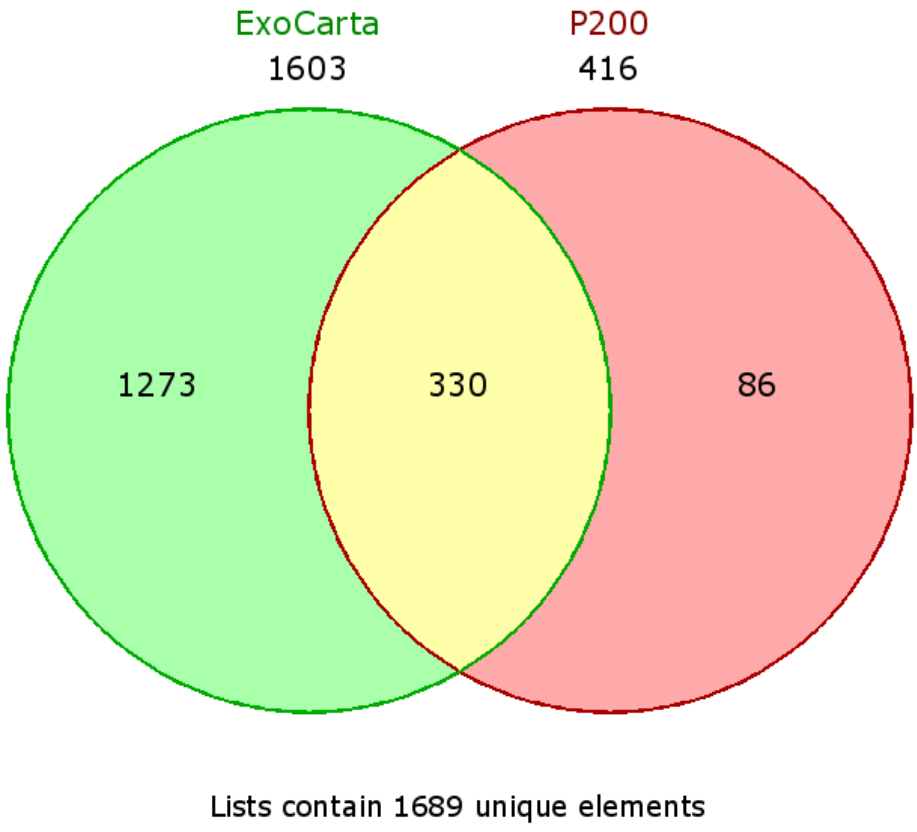

Figure 3.4: Comparison between our protein list and that of Exocarta (Mathivanan *et al.*, 2012). Our protein list was converted to Entrez Id identifiers to enable comparison to Exocarta protein list (Exocarta provides Entrez Ids as identifiers of proteins).

The proteins which are novel in term of urinary membrane vesicles association are shown in Table 3.1.

Table 3.1: Presents the proteins which have not previously been shown to be part of any type of urinary membrane vesicles, including exosomes.

Gene Name	UniGene	SwissProt (name)	SwissProt (acc)
TUBB4	Hs.110837	TBB4_HUMAN RIB2_HUMAN	P04350 P04844
APOB	Hs.120759	AGTR2_HUMAN	P50052
KRT17	Hs.2785		
THOC7	Hs.288151		
LDHAL6B	Hs.307052	LDH6B_HUMAN	Q9BYZ2
DEFB1	Hs.32949	BD01_HUMAN	P60022
FCN3	Hs.333383	P4HA2_HUMAN FCN3_HUMAN	O15460 O75636
SERPINB12	Hs.348541	SPB12_HUMAN	Q96P63
CFH	Hs.363396	CFAH_HUMAN	P08603
C8B	Hs.391835	CO8B_HUMAN	P07358
CNDP1	Hs.400613	CNDP1_HUMAN	Q96KN2
ITIH1	Hs.420257	ITIH1_HUMAN	P19827
KRT6A	Hs.433845	K2C6A_HUMAN K2C6C_HUMAN	P02538 P48666
HIST1H2BK	Hs.437275	H2B1C_HUMAN H2B1K_HUMAN	P62807 O60814
PTGDS	Hs.446429		
HLA-C	Hs.449621	1C07_HUMAN	P10321
IFITM2	Hs.458414	IFM2_HUMAN	Q01629
C1QC	Hs.467753		
IGHG1	Hs.510635	IGHD_HUMAN	P01880
GCNT2	Hs.519884	GCNT2_HUMAN	Q06430
UBC	Hs.520348		
ORM2	Hs.522356	A1AG2_HUMAN	P19652
MUC5B	Hs.523395	MUC5B_HUMAN	Q9HC84
CEL	Hs.533258	CEL_HUMAN	P19835
C4A	Hs.534847		
PSMAL	Hs.645352	PSMAL_HUMAN	Q9HBA9
AMY2A	Hs.654437	AMYP_HUMAN	P04746
PRSS3	Hs.654513		
KRT13	Hs.654550	K1C13_HUMAN	P13646
KRT15	Hs.654570	K1C15_HUMAN	P19012
KRT4	Hs.654610	K2C4_HUMAN	P19013
KRT16	Hs.655160	K1C16_HUMAN	P08779
LRG1	Hs.655559	A2GL_HUMAN	P02750
KRT72	Hs.662013		
TUBA4B	Hs.664469		
HIST1H2BK	Hs.689116	H2B1C_HUMAN H2B1K_HUMAN	P62807 O60814
CFB	Hs.69771	CFAB_HUMAN	P00751
	Hs.708950		
	Hs.719954		
	Hs.720022		
	Hs.724927		
C7	Hs.78065	CO7_HUMAN	P10643

C1QB	Hs.8986		

Mucin 5B (MUC5B) was shown not to be expressed in normal kidney and renal cell carcinoma by histochemistry and northern blot analysis. This becomes the first report suggesting that MUC5B might be expressed in kidney by at least some of the cell types. To the best of our knowledge MUC5B has not been reported to part of exosomes either (source: Exocarta).

3.3.1.2 Disease associations

Identified proteins were classified using DAVID Bioinformatics Resources 6.7, National Institute of Allergy and Infectious Diseases (NIAID), NIH (Huang, Sherman & Lempicki, 2009a; Huang, Sherman & Lempicki, 2009b). These 437 proteins mapped to 412 unique DAVID Ids and various types of analysis were performed on these proteins using DAVID. Out of these 412, 166 proteins are related to human diseases on the basis of their presence in OMIM database (Online Mendelian Inheritance in Man, a database of human genes and genetic disorders, http://www.ncbi.nlm.nih.gov/omim). These proteins are listed in supplemntary Table S3.2. The presence of these proteins suggests that urinary exosomes and other membrane vesicles are a rich source of biomarkers and can be applied for clinical biomarker analysis. Moreover, our identifications also contain 66 proteins whose gene have been found to contain genetic associations with cardiovascular diseases (Supplementary Table S3.3) and 23 proteins with genetic association to various renal diseases (Supplementary Table S3.4).

3.3.1.3 Sequence features

Total proteins identified were run on DAVID Bioinformatics Resources 6.7 for enrichment analysis for various sequence features. There were 192 different categories for various sequence features including, but not limited to, types of glycosylation, signal peptide,

mutagenesis sites, protein variants and many types of repeats. Some of the major ones are listed in Table 3.2.

Table 3.2: Lists the major sequence features found in proteins identified in the high speed pellet (P200,000xg).

Term	Count	% of total
signal peptide	214	51.94175
disulfide bond	165	40.04854
glycosylation site:N-linked (GlcNAc...)	222	53.8835
glycosylation site:O-linked (GalNAc...)	22	5.339806
glycosylation site:N-linked (GlcNAc...) (complex)	5	1.213592
glycosylation site:C-linked (Man)	7	1.699029
glycosylation site:N-linked (Glc) (glycation)	7	1.699029
sequence variant	306	74.27184
mutagenesis site	80	19.41748

Two hundred and fourteen proteins, corresponding to 48.9% of the total proteins identified were classified as having signal peptides. Signal peptide is found in proteins which are known to be secreted. Forty % of proteins were shown to employ disulfide bonds to generate tertiary structure. Fifty three % proteins were found to be N-glycosylated while only 5.3 % were reported to be O-glycosylated. This is expected as N-linked glycosylation is much more common compared to O-linked glycosylation. Among the O-glycosylated proteins, inter-alpha (globulin) inhibitor H4 is known to be involved in susceptibility to hypercholesterolemia. Among the N-glycosylated proteins, 5 proteins were found to possess complex type N-glycans. These proteins were alpha-1-microglobulin/bikunin precursor, angiotensin I converting enzyme 1, carboxyl ester lipase and inter-alpha (globulin) inhibitors H1 and H2. Another 7 proteins were found which are known to be glycated including albumin, CD59, complement factor B, apoliporpoteins A-I and E and haemoglobin alpha and

beta chains. 309 proteins were found to have sequence variants and 80 proteins had one or more mutagenesis sites.

3.3.1.4 Gene ontology/annotations

IPA software (Ingenuity systems, USA) was used for annotating the protein list and graph representations were generated manually using Microsoft (MS) Excel program. Figure 3.5 presents the major categories of proteins annotated using IPA in cellular component analysis. This classifies the protein list into sub-categories according to their localization in cell. As can be seen in figure 3.5, 36% of proteins belong to extracellular space category. Extracellular proteins if associated to the vesicle surface should be removed by DTT/CHAPS treatment and accordingly some extracellular proteins were found in CHAPS and DTT SN200,000 (Figure 2.7, Chapter 2). However, extracellular proteins which are endocytosed by the cells might be released with exosomes entrapped in the lumen. Such proteins would be found in the P200,000g regardless of CHAPS/DTT treatment. Thirty eight % proteins are cytoplasmic and 23% belong to plasma membrane category. Only 2 % proteins were found to be nuclear while 1% proteins were annotated as unknown. When the list of proteins was annotated according to the molecular function of proteins (Figure 3.6), 45% of proteins were found to be enzymes. Twenty one % and 18% were found to be peptidases and transporters, respectively. Four % transmembrane receptors were present while transcription/translation regulators and phosphatases were 3% each. Two % each of G-protein coupled-receptors, kinases and growth factors were also found.

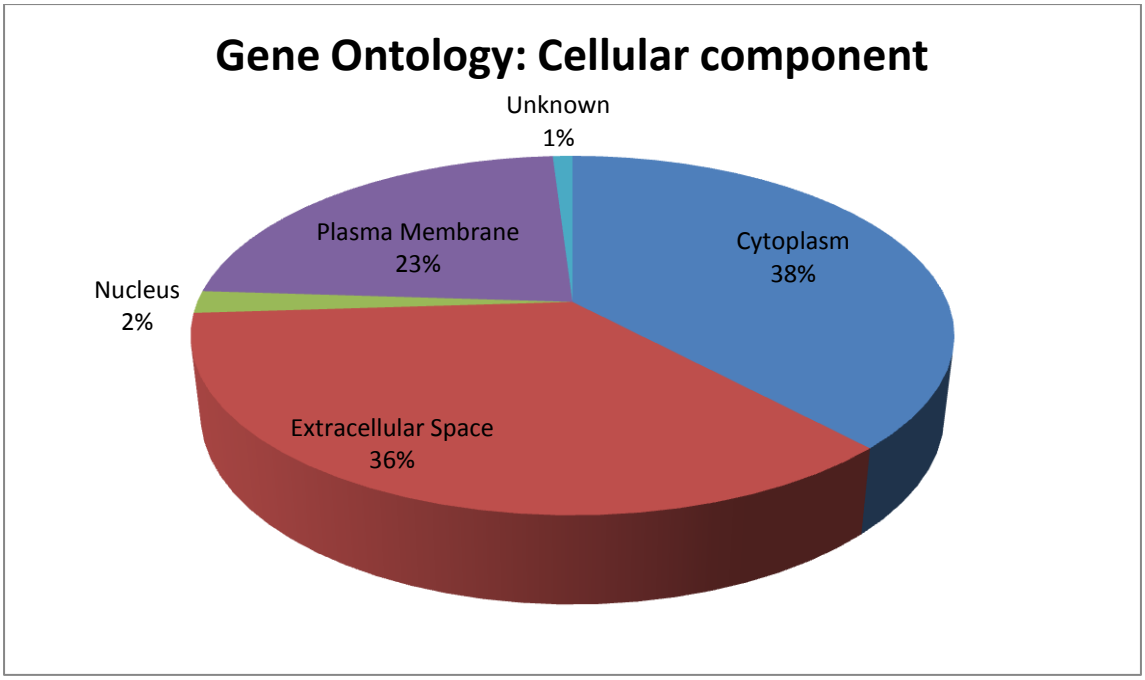

Figure 3.5: Annotation of our protein list using IPA software according to the cellular component to which they belong.

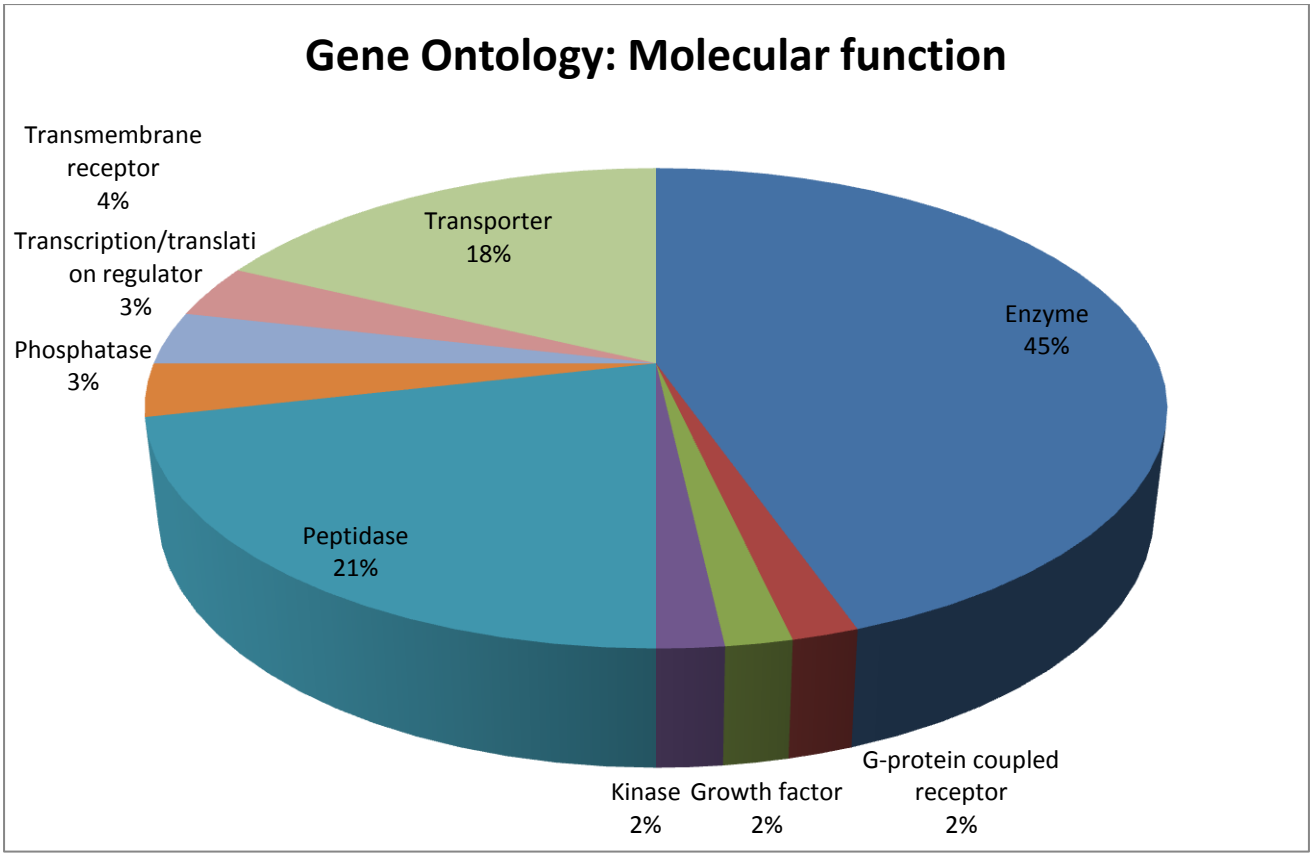

Figure 3.6: Annotation of our protein list using IPA software according to their molecular function.

Our protein list was annotated according to biological processes, in which these proteins are involved, using Blast2GO software (Conesa *et al.*, 2005) and 'cut-off' value was kept high (70 proteins) so only major biological processes are enriched (Figure 3.7).

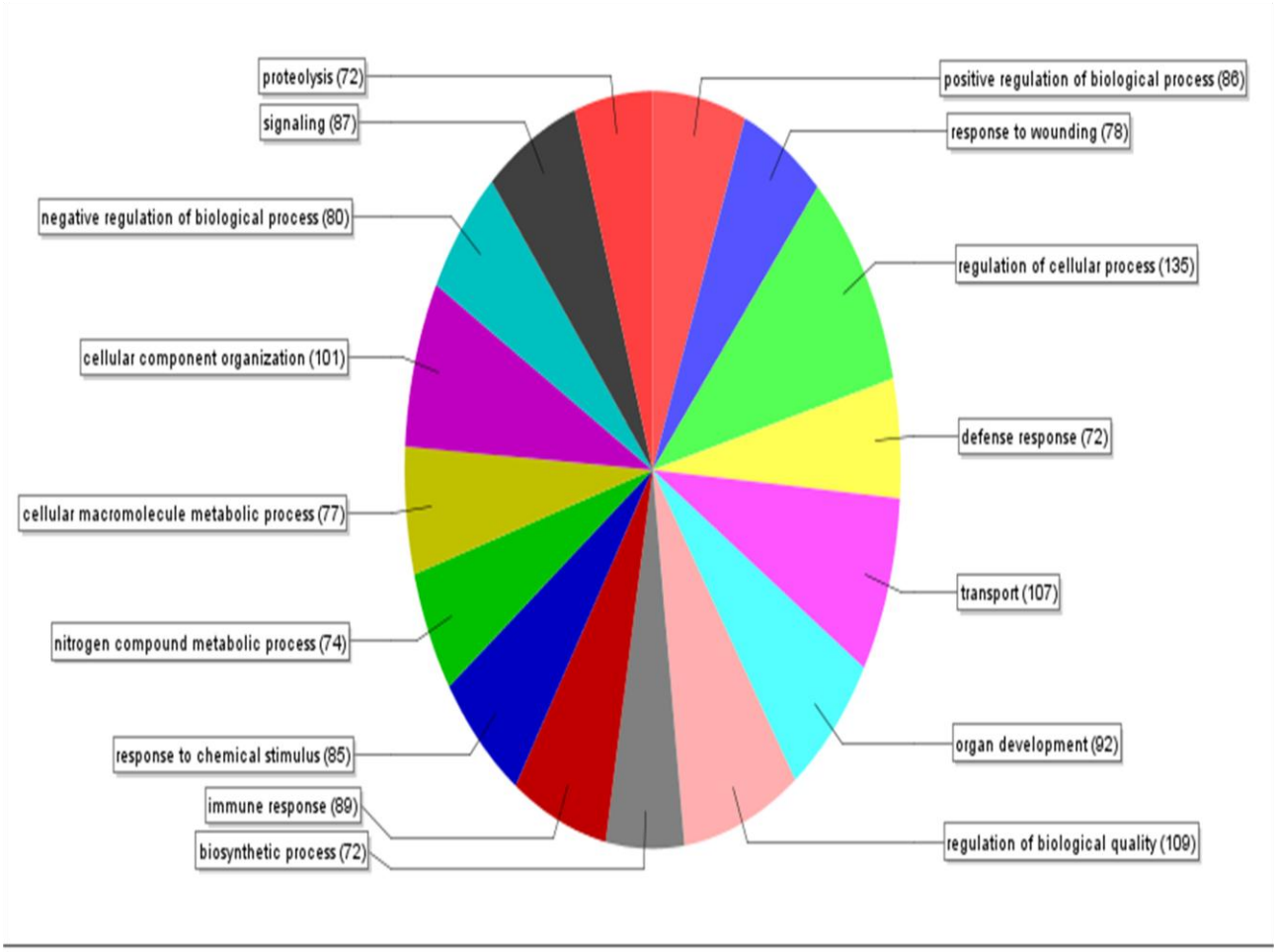

Figure 3.7: Annotation of our protein list using Blast2Go software (Conesa *et al.*, 2005) according to the biological processes in which they are involved.

As can be seen in Figure 3.7, the most number of proteins present in our list are the regulatory proteins regulating cellular processes (135 proteins) with 86 proteins involved in positive regulation of biological processes and 80 being the negative regulators of biological processes. Another significant biological process is the immune function with 89 proteins involved in immune responses being present and 72 proteins involved in defence responses. Cellular homeostasis is another important category as 78 proteins involved in response to wounding are also present. Seventy two proteins involved in proteolysis also demonstrate the importance of these vesicles in homeostasis. One hundred and seven transporters and 87

signalling proteins are present suggesting these vesicles have the complete machinery to impart functionality to their target cells.

Blast2go takes protein sequences in FASTA format as its input and there is no option of retrieving proteins which are classified into a particular category. For this reason, all the identified proteins in the high-speed pellet (437), which mapped to 412 DAVID Id, were annotated again using DAVID Bioinformatics Resources 6.7 into three categories: 1. Molecular function, 2. Biological process, and 3. Cellular components. Major classes falling in these categories were retrieved in the form of gene lists to discuss more about them. Some of the lists which seem interesting cast further light on importance of these vesicles are presented here.

Exosomes are released into urine when multivesicular body (MVB) fuse to the apical plasma membrane of cells lining the urinary drainage system. Therefore, identification of endosomal protein would suggest MVB origin of the vesicles we have identified. DAVID annotation was performed on our protein list and proteins belonging to endosomes were retrieved (Table 3.3). Class E vacuolar protein-sorting (VPS) proteins and some associated proteins were found in addition to tumor susceptibility gene 101 (Tsg101) and CD63 which are exosomal markers. Cubilin and megalin which are involved in endocytosis of proteins and cofactors into the endosomal system and they would be expected in exosomes as well.

Table 3.3: Proteins identified in our list which according to DAVID annotation belong to endosomes.

ID	Gene Name
Q6UXG3	CD300 molecule-like family member g
P08962	CD63 molecule
Q9H223	EH-domain containing 4
Q9NP79	Vps20-associated 1 homolog (S. cerevisiae)
P12821	angiotensin I converting enzyme (peptidyl-dipeptidase A) 1

P07355	annexin A2 pseudogene 3; annexin A2; annexin A2 pseudogene 1
P04114	apolipoprotein B (including Ag (x) antigen)
P15289	arylsulfatase A
Q7LBR1	chromatin modifying protein 1B
O43633	chromatin modifying protein 2A
Q9H444	chromatin modifying protein 4B
O60494	cubilin (intrinsic factor-cobalamin receptor)
P02649	hypothetical LOC100129500; apolipoprotein E
P98164	low density lipoprotein-related protein 2
P11279	lysosomal-associated membrane protein 1
P13473	lysosomal-associated membrane protein 2
P02787	Transferring
Q99816	tumor susceptibility gene 101
Q9UN37	vacuolar protein sorting 4 homolog A (S. cerevisiae)
O75351	vacuolar protein sorting 4 homolog B (S. cerevisiae)

Exosomes are thought to be derived from the apical part of plasma membrane in kidney. The DAVID annotation classified some proteins as belonging to the apical part. They are presented in Table 3.4. These proteins found in the exosomal pellet would agree with the theory that exosomes are released from the apical part of the plasma membrane in urinary tract.

Table 3.4: Proteins annotated to be from apical part of the plasma membrane.

ID	Gene Name
P05026	ATPase, Na+/K+ transporting, beta 1 polypeptide
Q9UBD6	Rh family, C glycoprotein
P14550	aldo-keto reductase family 1, member A1 (aldehyde reductase)
P08133	annexin A6
P29972	aquaporin 1 (Colton blood group)
P43251	Biotinidase
P00918	carbonic anhydrase II
P07858	cathepsin B
P60953	cell division cycle 42 (GTP binding protein, 25kDa); cell division cycle 42 pseudogene 2
O60494	cubilin (intrinsic factor-cobalamin receptor)
P27487	dipeptidyl-peptidase 4
P02751	fibronectin 1
Q07075	glutamyl aminopeptidase (aminopeptidase A)
P15311	hypothetical protein LOC100129652; ezrin

P98164	low density lipoprotein-related protein 2
P26038	Moesin
P15941	mucin 1, cell surface associated
P27105	phosphatidylethanolamine binding protein 1
O43490	prominin 1
P35241	Radixin
P55017	solute carrier family 12 (sodium/chloride transporters), member 3
Q13621	solute carrier family 12 (sodium/potassium/chloride transporters), member 1
P07911	Uromodulin

Although some proteins from the basolateral part of the plasma membrane were also found (Table 3.5).

Table 3.5: Proteins annotated by DAVID to be part of basolateral membrane.

ID	Gene Name
P08183	ATP-binding cassette, sub-family B (MDR/TAP), member 1
P05026	ATPase, Na+/K+ transporting, beta 1 polypeptide
Q9UBD6	Rh family, C glycoprotein
P31151	S100 calcium binding protein A7
P15291	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 1
P12814	actinin, alpha 1
P04083	annexin A1
P29972	aquaporin 1 (Colton blood group)
P00918	carbonic anhydrase II
P15924	Desmoplakin
P26038	Moesin
O00159	myosin IC
O00560	syndecan binding protein (syntenin)
P02787	Transferring

The expression of these proteins (Table 3.5) is not restricted to basolateral part so the theory of apical generation of exosomes cannot be excluded. For example, Aquaporin 1 is also found on the apical side of the plasma membrane (Source: Uniprot).

Some proteins of the brush border membrane of proximal tubuli were found in our identifications (Table 3.6) suggesting that most if not all parts of the urinary drainage system release these vesicles.

Table 3.6: Proteins annotated to brush border/microvillus by DAVID.

ID	Gene Name
P15291	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 1
O60494	cubilin (intrinsic factor-cobalamin receptor)
Q7Z4W1	dicarbonyl/L-xylulose reductase
Q07075	glutamyl aminopeptidase (aminopeptidase A)
P98164	low density lipoprotein-related protein 2
O00159	myosin IC
O43490	prominin 1
P31639	solute carrier family 5 (sodium/glucose cotransporter), member 2
P35241	Radixin
P26038	Moesin
P00918	carbonic anhydrase II
P15311	hypothetical protein LOC100129652; ezrin

The whole membrane attack complex subunits of complement system were found in our identifications. This raises questions about constitutive secretion over stimulus-mediated secretion.

Table 3.7: Proteins annotated to be complement components by DAVID.

ID	Gene Name
P10909	Clusterin
P02746	complement component 1, q subcomponent, B chain
P02747	complement component 1, q subcomponent, C chain
Q9NZP8	complement component 1, r subcomponent-like
P04003	complement component 4 binding protein, alpha
P0C0L4	complement component 4A (Rodgers blood group)
P0C0L5	complement component 4B (Chido blood group)
P01031	complement component 5
P10643	complement component 7
P07357	complement component 8, alpha polypeptide
P07358	complement component 8, beta polypeptide

P02748	complement component 9
P00751	complement factor B
P08603	complement factor H
P05156	complement factor I
Q15485	ficolin (collagen/fibrinogen domain containing lectin) 2 (hucolin)
O75636	ficolin (collagen/fibrinogen domain containing) 3 (Hakata antigen)
P04264	keratin 1
O00187	mannan-binding lectin serine peptidase 2
P05155	serpin peptidase inhibitor, clade G (C1 inhibitor), member 1
P01024	similar to Complement C3 precursor; complement component 3; hypothetical protein LOC100133511

Another interesting feature of our protein list was some proteins which bind unfolded proteins. In other words these proteins act as chaperones aiding protein folding and also acting in the clearance of misfolded proteins. These proteins are presented in Table 3.8. HSP70 variants and HSP90 are present as well as cyclophilin A, B and C.

Table 3.8: Unfolded protein binding proteins annotated by DAVID.

ID	Gene Name
P02743	amyloid P component, serum
P08107	heat shock 70kDa protein 1A; heat shock 70kDa protein 1B
P54652	heat shock 70kDa protein 2
P11142	heat shock 70kDa protein 8
P08238	heat shock protein 90kDa alpha (cytosolic), class B member 1
P23284	peptidylprolyl isomerase B (cyclophilin B)
P45877	peptidylprolyl isomerase C (cyclophilin C)
P62937	similar to TRIMCyp; peptidylprolyl isomerase A (cyclophilin A); peptidylprolyl isomerase A (cyclophilin A)-like 3
P68371	tubulin, beta 2C

Another important class of proteins were identified in our high speed pellet as having GTPase activity. These include many of the G-proteins, RAS oncogene family members and translation elongation factor 1.

Table 3.9: Proteins annotated as having GTPase activity by DAVID

ID	Gene Name
Q9H223	EH-domain containing 4
P63092	GNAS complex locus
P61224	RAP1B, member of RAS oncogene family
P60953	cell division cycle 42 (GTP binding protein, 25kDa); cell division cycle 42 pseudogene 2
P68104	eukaryotic translation elongation factor 1 alpha-like 7; eukaryotic translation elongation factor 1 alpha-like 3; similar to eukaryotic translation elongation factor 1 alpha 1; eukaryotic translation elongation factor 1 alpha 1
P29992	guanine nucleotide binding protein (G protein), alpha 11 (Gq class)
P63096	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 1
P62873	guanine nucleotide binding protein (G protein), beta polypeptide 1
P62879	guanine nucleotide binding protein (G protein), beta polypeptide 2
P50148	guanine nucleotide binding protein (G protein), q polypeptide
P61586	ras homolog gene family, member A
Q9BQE3	tubulin, alpha 1c
P68366	tubulin, alpha 4a
Q9H853	tubulin, alpha 4b (pseudogene)
P68371	tubulin, beta 2C
P04350	tubulin, beta 4

3.3.1.5 Comparison of High speed pellet with HDL and LDL particles

Lipoprotein particles, particularly the high-density lipoproteins (HDL), share some biophysical characteristics with urinary exosomes. For example, HDL have a density of 1.063-1.125 g/mL although its diameter ranges from 7-12nm (Gordon *et al.,* 2010), which is smaller than exosomes. HDL particles have been identified in urine previously (Gomo, Henderson & Myrick, 1988). Owing to a similar density, HDL might sediment with exosomes when exosomes are prepared from urine using ultracentrifugation. LDL and VLDL would be expected to remain in solution due to their lower densities, although the presence of dense LDL particles cannot be ruled out. Proteomic studies on HDL particles have to deal with the possibility of contamination of lipoprotein preparations with soluble abundant proteins. For this reason, one recent HDL proteomics study was selected in which size exclusion chromatography and affinity chromatography were used to purify HDLs which

were subsequently analysed by MS for their proteomic content ("the HDL proteome") (Gordon *et al.,* 2010). Lipid binding proteins identified in this study were compared to our protein list (Figure 3.8).

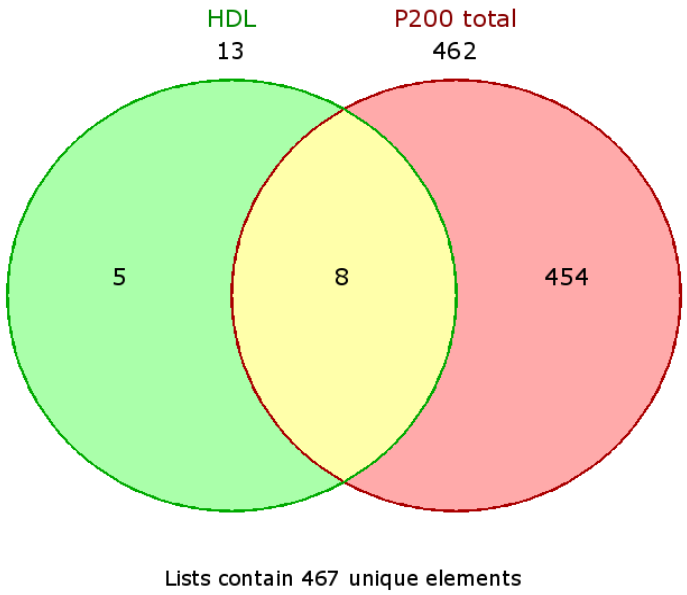

Figure 3.8: Comparison of HDL-associated proteins (Gordon *et al.,* 2010) with proteins of high speed pellet (P200,000g).

Out of the 12 proteins which were converted to 13 Unigene identifications for comparison, apolipoprotein C1 (APOC1), apolipoprotein L1 (APOL1), apolipoprotein A2 (APOA2) and serum amyloid A (SAA) were not found in our study. In the Figure 5 proteins are shown but only four were absent. This was due to the conversion of proteins to different identification Ids for comparison. All other proteins were present in our protein list. APOA1, APOA4, APOE, complement C3, PON1, complement C4-B and cholesteryl ester transfer protein (CETP), which have been identified previously to be HDL-associated proteins (Vaisar *et al.,* 2007) are also present in our protein list. The presence of full length APO B-100 was suggested by the peptides identified in our analysis (Figure 3.9) which also suggests the presence of LDL.

EEEMLENVSLVCPKDATRFKHLRKYTYNYEAESSSGVPGTADSRSATRINCKVELEVPQLCSFILKTSQCTLKEVYGFNPEGKALLKKTKNSEEFAAAMSRYELKLAI
PEGKQVFLYPEKDEPTYILNIKRGIISALLVPPETEEAKQVLFLDTVYGNCSTHFTVKTRKGNVATEISTERDLGQCDRFKPIRTGISPLALIKGMTRPLSTLISSSQSCQ
YTLDAKRKHVAEAICKEQHLFLPFSYKNKYGMVAQVTQTLKLEDTPKINSRFFGEGTKKMGLAFESTKSTSPPKQAEAVLKTLQELKKLTISEQNIQRANLFNKLVT
ELRGLSDEAVTSLLPQLIEVSSPITLQALVQCGQPQCSTHILQWLKRVHANPLLIDVVTYLVALIPEPSAQQLREIFNMARDQRSRATLYALSHAVNNYHKTNPTGT
QELLDIANYLMEQIQDDCTGDEDYTYLILRVIGNMGQTMEQLTPELKSSILKCVQSTKPSLMIQKAAIQALRKMEPKDKDQEVLLQTFLDDASPGDKRLAAYLML
MRSPSQADINKIVQILPWEQNEQVKNFVASHIANILNSEELDIQDLKKLVKEALKESQLPTVMDFRKFSRNYQLYKSVSLPSLDPASAKIEGNLIFDPNNYLPKESM
LKTTLTAFGFASADLIEIGLEGKGFEPTLEALFGKQGFFPDSVNKALYWVNGQVPDGVSKVLVDHFGYTKDDKHEQDMVNGIMLSVEKLIKDLKSKEVPEARAYL
RILGEELGFASLHDLQLLGKLLLMGARTLQGIPQMIGEVIRKGSKNDFFLHYIFMENAFELPTGAGLQLQISSSGVIAPGAKAGVKLEVANMQAELVAKPSVSVEFV
TNMGIIIPDFARSGVQMNTNFFHESGLEAHVALKAGKLKFIIPSPKRPVKLLSGGNTLHLVSTTKTEVIPPLIENRQSWSVCKQVFPGLNYCTSGAYSNASSTDSASY
YPLTGDTRLELELRPTGEIEQYSVSATYELQREDRALVDTLKFVTQAEGAKQTEATMTFKYNRQSMTLSSEVQIPDFDVDLGTILRVNDESTEGKTSYRLTLDIQNK
KITEVALMGHLSCDTKEERKIKGVISIPRLQAEARSEILAHWSPAKLLLQMDSSATAYGSTVSKRVAWHYDEEKIEFEWNTGTNVDTKKMTSNFPVDLSDYPKSLH
MYANRLLDHRVPQTDMTFRHVGSKLIVAMSSWLQKASGSLPYTQTLQDHLNSLKEFNLQNMGLPDFHIPENLFLKSDGRVKYTLNKNSLKIEIPLPFGGKSSRDL
KMLETVRTPALHFKSVGFHLPSREFQVPTFTIPKLYQLQVPLLGVLDLSTNVYSNLYNWSASYSGGNTSTDHFSLRARYHMKADSVVDLLSYNVQGSGETTYDHK
NTFTLSYDGSLRHKFLDSNIKFSHVEKLGNNPVSKGLLIFDASSSWGPQMSASVHLDSKKKQHLFVKEVKIDGQFRVSSFYAKGTYGLSCQRDPNTGRLNGESNLR
FNSSYLQGTNQITGRYEDGTLSLTSTSDLQSGIIKNTASLKYENYELTLKSDTNGKYKNFATSNKMDMTFSKQNALLRSEYQADYESLRFFSLLSGSLNSHGLELNAD
ILGTDKINSGAHKATLRIGQDGISTSATTNLKCSLLVLENELNAELGLSGASMKLTTNGRFREHNAKFSLDGKAALTELSLGSAYQAMILGVDSKNIFNFKVSQEGLK
LSNDMMGSYAEMKFDHTNSLNIAGLSLDFSSKLDNIYSSDKFYKQTVNLQLQPYSLVTTLNSDLKYNALDLTNNGKLRLEPLKLHVAGNLKGAYQNNEIKHIYAISS
AALSASYKADTVAKVQGVEFSHRLNTDIAGLASAIDMSTNYNSDSLHFSNVFRSVMAPFTMTIDAHTNGNGKLALWGEHTGQLYSKFLLKAEPLAFTFSHDYKG
STSHHLVSRKSISAALEHKVSALLTPAEQTGTWKLKTQFNNNEYSQDLDAYNTKDKIGVELTGRTLADLTLLDSPIKVPLLLSEPINIIDALEMRDAVEKPQEFTIVAF
VKYDKNQDVHSINLPFFETLQEYFERNRQTIIVVLENVQRNLKHINIDQFVRKYRAALGKLPQQANDYLNSFNWERQVSHAKEKLTALTKKYRITENDIQIALDDA
KINFNEKLSQLQTYMIQFDQYIKDSYDLHDLKIAIANIIDEIIEKLKSLDEHYHIRVNLVKTIHDLHLFIENIDFNKSGSSTASWIQNVDTKYQIRIQIQEKLQQLKRHIQ
NIDIQHLAGKLKQHIEAIDVRVLLDQLGTTISFERINDILEHVKHFVINLIGDFEVAEKINAFRAKVHELIERYEVDQQIQVLMDKLVELAHQYKLKETIQKLSNVLQQ
VKIKDYFEKLVGFIDDAVKKLNELSFKTFIEDVNKFLDMLIKKLKSFDYHQFVDETNDKIREVTQRLNGEIQALELPQKAEALKLFLEETKATVAVYLESLQDTKITLIIN
WLQEALSSASLAHMKAKFRETLEDTRDRMYQMDIQQELQRYLSLVGQVYSTLVTYISDWWTLAAKNLTDFAEQYSIQDWAKRMKALVEQGFTVPEIKTILGTM
PAFEVSLQALQKATFQTPDFIVPLTDLRIPSVQINFKDLKNIKIPSRFSTPEFTILNTFHIPSFTIDFVEMKVKIIRTIDQMLNSELQWPVPDIYLRDLKVEDIPLARITLP
DFRLPEIAIPEFIIPTLNLNDFQVPDLHIPEFQLPHISHTIEVPTFGKLYSILKIQSPLFTLDANADIGNGTTSANEAGIAASITAKGESKLEVLNFDFQANAQLSNPKINP
LALKESVKFSSKYLRTEHGSEMLFFGNAIEGKSNTVASLHTEKNTLELSNGVIVKINNQLTLDSNTKYFHKLNIPKLDFSSQADLRNEIKTLLKAGHIAWTSSGKGSW
KWACPRFSDEGTHESQISFTIEGPLTSFGLSNKINSKHLRVNQNLVYESGSLNFSKLEIQSQVDSQHVGHSVLTAKGMALFGEGKAEFTGRHDAHLNGKVIGTLKN
SLFFSAQPFEITASTNNEGNLKVRFPLRLTGKIDFLNNYALFLSPSAQQASWQVSARFNQYKYNQNFSAGNNENIMEAHVGINGEANLDFLNIPLTIPEMRLPYTII
TTPPLKDFSLWEKTGLKEFLKTTKQSFDLSVKAQYKKNKHRHSITNPLAVLCEFISQSIKSFDRHFEKNRNNALDFVTKSYNETKIKFDKYKAEKSHDELPRTFQIPGY
TVPVVNVEVSPFTIEMSAFGYVFPKAVSMPSFSILGSDVRVPSYTLILPSLELPVLHVPRNLKLSLPDFKELCTISHIFIPAMGNITYDFSFKSSVITLNTNAELFNQSDI
VAHLLSSSSSVIDALQYKLEGTTRLTRKRGLKLATALSLSNKFVEGSHNSTVSLTTKNMEVSVATTTKAQIPILRMNFKQELNGNTKSKPTVSSSMEFKYDFNSSML
YSTAKGAVDHKLSLESLTSYFSIESSTKGDVKGSVLSREYSGTIASEANTYLNSKSTRSSVKLQGTSKIDDIWNLEVKENFAGEATLQRIYSLWEHSTKNHLQLEGLFF
TNGEHTSKATLELSPWQMSALVQVHASQPSSFHDFPDLGQEVALNANTKNQKIRWKNEVRIHSGSFQSQVELSNDQEKAHLDIAGSLEGHLRFLKNIILPVYDKS
LWDFLKLDVTTSIGRRQHLRVSTAFVYTKNPNGYSFSIPVKVLADKFIIPGLKLNDLNSVLVMPTFHVPFTDLQVPSCKLDFREIQIYKKLRTSSFALNLPTLPEVKFPE
VDVLTKYSQPEDSLIPFFEITVPESQLTVSQFTLPKSVSDGIAALDLNAVANKIADFELPTIIVPEQTIEIPSIKFSVPAGIVIPSFQALTARFEVDSPVYNATWSASLKNK
ADYVETVLDSTCSSTVQFLEYELNVLGTHKIEDGTLASKTKGTFAHRDFSAEYEEDGKYEGLQEWEGKAHLNIKSPAFTDLHLRYQKDKKGISTSAASPAVGTVGM
DMDEDDDFSKWNFYYSPQSSPDKKLTIFKTELRVRESDEETQIKVNWEEEAASGLLTSLKDNVPKATGVLYDYVNKYHWEHTGLTLREVSSKLRRNLQNNAEWV
YQGAIRQIDDIDVRFQKAASGTTGTYQEWKDKAQNLYQELLTQEGQASFQGLKDNVFDGLVRVTQEFHMKVKHLIDSLIDFLNFPRFQFPGKPGIYTREELCTM
FIREVGTVLSQVYSKVHNGSEILFSYFQDLVITLPFELRKHKLIDVISMYRELLKDLSKEAQEVFKAIQSLKTTEVLRNLQDLLQFIFQLIEDNIKQLKEMKFTYLINYIQD
EINTIFSDYIPYVFKLLKENLCLNLHKFNEFIQNELQEASQELQQIHQYIMALREEYFDPSIVGWTVKYYELEEKIVSLIKNLLVALKDFHSEYIVSASNFTSQLSSQVEQ
FLHRNIQEYLSILTDPDGKGKEKIAELSATAQEIIKSQAIATKKIISDYHQQFRYKLQDFSDQLSDYYEKFIAESKRLIDLSIQNYHTFLIYITELLKKLQSTTVMNPYMKL
APGELTIILSEAEDASLLSFMQGYMKHATKTAKDALSSVQESQVAQQARGWVTDGFSSLKDYWSTVKDKFSEFWDLDPEVRPTSAVAA

Figure 3.9: Sequence of full length APO B-100. Amino acids highlighted in green are the peptides identified in our study and red amino acids are lipid binding amino acids.

3.3.2 Low speed pellet proteomic analysis

Proteins pelleting down at low speed (18,000g) after DTT or CHAPS treatment were also identified. It is known from previous studies that some populations of exosomes pellet down at low speed pellet owing to their entrapment by THP polymers. However, DTT treatment releases these vesicles which then pellet down at high speed (P200,000g) (Fernandez-Llama

et al., 2010). Therefore, DTT-treated P18,000g, theoretically, would be expected to have only larger vesicles. Seventy five proteins were identified in DTT and CHAPS-treated P18,000g (pelleted again after treatment fo the crude pellet P18**DTT/CHAPS**P18). A complete non-redundant list of proteins identified in low speed pellet (P18,000g) is presented in Supplementary Table S3.5. DTT treatment is the method suggested by some previous studies while CHAPS treatment was established by us (Chapter 2). When these two methods were compared to each other in terms of protein identifications, our method (CHAPS) yielded more proteins while most (66%) of the proteins identified in DTT method were common to our method.

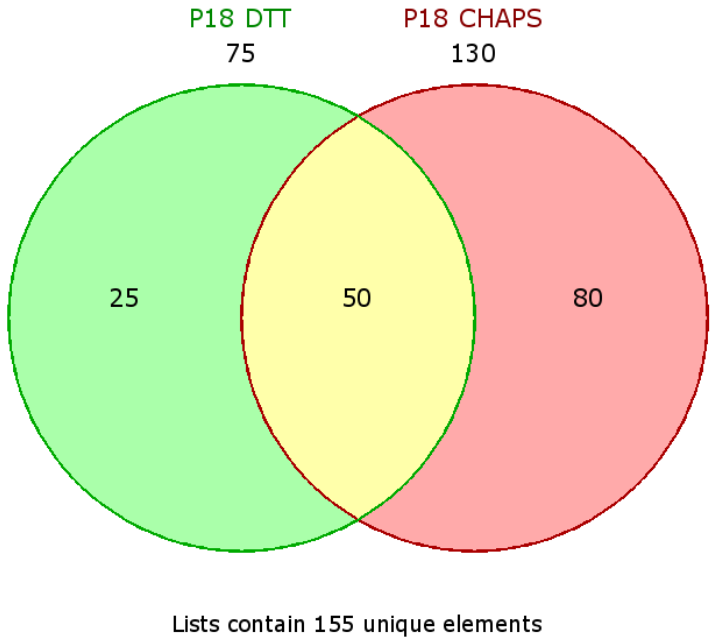

Figure 3.10: DTT versus CHAPS treatment in terms of protein identification for low speed pellet (18,000g).

Sixty six % of DTT-treated pellet proteins are common to CHAPS pellet while only 38% of CHAPS are common to DTT method (Figure 3.10). This suggests that CHAPS is not only better in terms of removing contamination of soluble proteins but this also enriches for proteins not otherwise identified in the pellet when it is treated with DTT.

Proteins identified in P18,000, after treatment with either DTT or CHAPS separately (DTT and CHAPS treated P18,000g; P18DP18 and P18CP18), were combined together non-redundantly and compared to P200,000g (Figure 3.11).

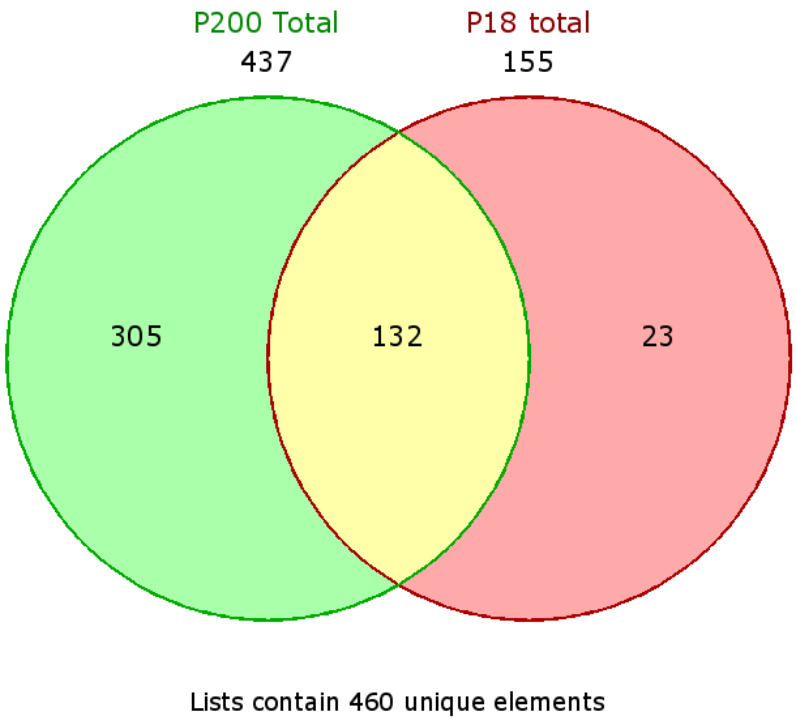

Figure 3.11: Comparison of total proteins of low speed pellet (P18DTT and P18CHAPS combined) with High speed pellet (P200,000g).

Eighty five % proteins in P18,000g are common to P200,000g as can be seen in Figure 3.11. Out of 23 proteins unique to the low speed pellet (P18,000g), 14 have been previously identified in a large-scale proteomic study of urinary exosomes (Wang *et al.*, 2011).

3.4 Discussion

Proteomic analyses of the high and low speed pellets were carried out and a large number of proteins were identified. Out of the 437 proteins identified in the high-speed pellet (P200,000g or exosomal fraction), 86 proteins have never been described in urinary exosomes. Exocarta is the largest database of exosomal proteins from any source (Mathivanan *et al.,* 2012) and 83 proteins identified in our study are new and not reported in Exocarta. However, as compared to the largest proteomics study on whole urine (Marimuthu *et al.,* 2011) most of our proteins have been identified in urine except the 32 proteins which are new to our study. Thirty nine proteins are, for the first time, shown to be part of urinary membrane vesicles and they have not been reported previously in urinary exosomes and 'exosome-like' vesicles study.

One hundred and sixty-six proteins have been found in the OMIM database associated with various diseases with 66 proteins having genetic associations to cardiovascular diseases and 23 with genetic associations to renal diseases. This highlights the potential importance of these vesicles in clinical practise. Two hundred and twenty two proteins in our list were found to be N-glycosylated, as annotated by DAVID bioinformatics resources 6.7, and 22 were O-glycosylated. Glycosylation patterns of proteins are important variables in disease processes which most likely reflect the respective pathology in the form of aberrant or altered glycosylation. Interestingly, membrane vesicles have a distinct glycan signature which is different from the parent cells secreting them (Escrevente *et al.,* 2011). These membrane vesicles originate from specialised microdomains of the plasma membrane of the cell suggesting that glycosylation plays a part in sorting these proteins to such microdomains. As an example, special processing of prion protein is involved in targeting it to membrane vesicles and only a certain glycoform is preferentially secreted as part of membrane vesicles (Vella *et al.,* 2007) over other forms. This further suggests a role for glycosylation in sorting

different proteins to specific membrane vesicle groups. Additionally, 7 proteins including albumin and haemoglobin alpha and beta chains which are known to be glycated are present in our high-speed pellet protein list. Advanced glycation end (AGE) products are known to be formed characteristically in diabetes and diabetic nephropathy. They seem to act as stimuli to induce production of extracellular matrix and inhibit its degradation. Furthermore, AGE modification of matrix proteins disrupts matrix-matrix and matrix-cell interactions leading to a profibrotic state (Silbiger *et al.,* 1993; Krishnamurti *et al.,* 1997; Mott *et al.,* 1997; Raabe *et al.,* 1998). It remains to be established if these proteins are secreted in their glycated forms in membrane vesicles or not. If glycated forms can be found in membrane vesicles it could work as a diagnostic modality to determine concentration of glycated proteins in association with progression of various diseases. Twenty two proteins were annotated as O-glycosylated in our high-speed pellet protein list. It has been reported that metabolic flux through the hexosamine biosynthetic pathway is increased in the presence of high circulating glucose levels as in diabetes. Furthermore, it induces the expression of genes necessary for development of diabetic nephropathy and O-glycosylation was found to be necessary for the expression of these genes (Goldberg *et al.,* 2006). Twenty proteins were annotated to be part of endosomes in our high-speed pellet protein list. Class E vacuolar protein-sorting (VPS) proteins are involved in MVB biogenesis and Tsg101 and CD63 help sort proteins to MVB. This provides evidence that high speed pellets contain mostly vesicles originating in MVB which are also known as urinary exosomes. It is believed that exosomes are secreted following the fusion of MVB to the apical part of plasma membranes in the urinary drainage system (Pisitkun, Shen & Knepper; 2004). Twenty three proteins, which are markers of apical part of the plasma membrane, were found in our study confirming this hypothesis. However, 14 proteins were also annotated to be part of basolateral aspect of the membrane while many of these proteins like ATPase Na+/K+ transporting, beta1 polypeptide and Rh family C

glycoprotein are common to the apical part as well. The expression of these proteins, annotated as basolateral proteins, is not strictly basolateral and typically strict basolateral makrer proteins like gp58 and E-cadherin (Manninen *et al.*, 2005) were not found in our protein list. Therefore, our results support the hypothesis about generation of exosomes from the the apical part of the plasma membrane.

HDL particles have been identified in urine previously (Gomo, Henderson & Myrick, 1988) and they share some biophysical characteristics with exosomes like a similar density range. The traditional method of exosome isolation uses ultracentrifugation and membrane vesicles, like exosomes, pellet down depending on their density and mass. Therefore, it is not unexpected if this method also pellets down HDL particles along with membrane vesicles. A number of HDL-associated proteins are present in our identifications but some like APOC1, APOL1, APOA2 and SAA are absent as well. However, a point to remember is that CHAPS, which is used in our enrichment protocol, has been reported to solubilise HDL and release apolipoproteins at roughly 10mM final concentration along with other detergents like cholate (Shiflett *et al.,* 2005). We have used 1% (w/v) CHAPS which is roughly equivalent to 16mM suggesting that HDL, if present, may be solubilised to release apolipoproteins. Identification of these apolipoproteins in our high speed pellet would then mean that either these proteins are secreted as part of membrane vesicles, or, upon their release from HDL, they bind to exosome or other membrane vesicles. Alternatively, CHAPS, without the presence of other detergents, may not have any solubilisation effect on HDL and thus these particles would be present in our nano vesicle fraction. The presence of dense LDL particle can also not be ruled out as peptides covering the full length of APO B-100 are present in our identifications.

Exosomes are thought to be secreted constitutively from various cells as part of routine cellular transport processes, whereas ectosomes are mostly released upon receipt of a stimulus such as complement attack (Morgan, Dankert & Esser, 1987). Complement attack

leads to the formation of a membrane attack complex (MAC) on the cell surface which, if not disposed of, can lead to cell death (Mayer 1961). Cells dispose of MAC by either releasing ectosomes consisiting of MAC (Morgan, Dankert & Esser, 1987) or by internalising MAC into MVBs and then secreting it as part of exosomes, as shown in glomerular epithelial cells in kidney (Kerjaschki *et al.,* 1989). Our protein list in the high-speed pellet (P200,000g) contains all subunits of MAC and it can only be speculated as to whether it is exosomes or ectosomes which delivered them to urine. However, HDL has also been found previously to be associated with MAC (Choimiura *et al.,* 1993) and presence of HDL in our vesicle fraction carrying MAC cannot be ruled out.

Our protein list also consisted of a number of proteins having chaperone activity which bind to unfolded proteins including Hsp70, 90 and clycophilin A, B and C. This suggests that target cells taking up exosomes (and these chaperones with them), would be better equipped to deal with stress induced by unfolded proteins although it remains to be determined in detail. Moreover, cyclophilin A is known to be packaged in HIV-1 virion (Gitti *et al.,* 1996) and it has been suggested that HIV-1 uses the same machinery as membrane vesicles for budding (Krishnamoorthy *et al.,* 2009).

The low speed pellet (P18,000g) of urine has never been subjected to proteomic analysis before. For the first time we have processed the pellet using two methods described in Chapter one namely, DTT and CHAPS treatment. It was clear that CHAPS treatment helped us to identify more proteins in the low speed pellet compared to DTT treatment (130 as compared to 75 proteins identified). This is possibly due to the more efficient removal of interference of the highly abundant protein THP. However, when we combined the proteins recovered with these two methods and compared it to the high-speed pellet (P200,000g) only 23 proteins were found to be unique to the low speed pellet. Out of these 23, 14 have been previously identified in the high speed pellet (P200,000g) (Wang *et al.,* 2011). This suggests

that due to either THP entrapment of membrane vesicles (Fernandez-Llama *et al.*, 2010) or other physicochemical factors, a population of exosomes, and other vesicles of low density, pellet down at low speed (P18,000g). Therefore, complete recovery of all the vesciles, present in urine with sizes similar to exosomes, at high-speed pellet (P200,000g) was obtained even following both treatment methods (DTT and CHAPS), although it can definitely be improved using these two methods.

3.5 Reference

Choimiura, N., Sakamoto, T., Tobe, T., Nakano, Y. & Tomita, M. (1993), "the role of HDL consisting of SP-40, APO A-I, and lipids in the formation of sMAC of complement", *JOURNAL OF BIOCHEMISTRY*, vol. 113, no. 4, pp. 484-487.

Conesa, A., Gotz, S., Garcia-Gomez, J., Terol, J., Talon, M. & Robles, M. (2005), "Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research", *BIOINFORMATICS*, vol. 21, no. 18, pp. 3674-3676.

Escrevente, C., Keller, S., Altevogt, P. & Costa, J. (2011), "Interaction and uptake of exosomes by ovarian cancer cells", *BMC CANCER*, vol. 11, pp. 108.

Fernandez-Llama, P., Khositseth, S., Gonzales, P.A., Star, R.A., Pisitkun, T. & Knepper, M.A. (2010), "Tamm-Horsfall protein and urinary exosome isolation", *JOURNAL OF HYPERTENSION*, vol. 28, no. A, pp. E164.

Gitti, R.K., Lee, B.M., Walker, J., Summers, M.F., Yoo, S. & Sundquist, W.I. (1996), "Structure of the Amino-Terminal Core Domain of the HIV-1 Capsid Protein", *SCIENCE*, vol. 273, no. 5272, pp. 231-235.

Goldberg, H., Whiteside, C., Hart, G. & Fantus, I. (2006), "Posttranslational, reversible O-glycosylation is stimulated by high glucose and mediates plasminogen activator inhibitor-1 gene expression and Sp1 transcriptional activity in glomerular mesangial cells", *ENDOCRINOLOGY*, vol. 147, no. 1, pp. 222-231.

Gomo, Z.A., Henderson, L.O. & Myrick, J.E. (1988), "High-density lipoprotein apolipoproteins in urine: I. Characterization in normal subjects and in patients with proteinuria.", *Clinical Chemistry*, vol. 34, no. 9, pp. 1775-1780.

Gonzales, P.A., Pisitkun, T., Hoffert, J.D., Tchapyjnikov, D., Star, R.A., Kleta, R., Wang, N.S. & Knepper, M.A. (2009), "Large-Scale Proteomics and Phosphoproteomics of Urinary Exosomes", *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY*, vol. 20, no. 2, pp. 363-379.

Gordon, S.M., Deng, J., Lu, L.J. & Davidson, W.S. (2010), "Proteomic Characterization of Human Plasma High Density Lipoprotein Fractionated by Gel Filtration Chromatography", *JOURNAL OF PROTEOME RESEARCH*, vol. 9, no. 10, pp. 5239-5249.

Gordon, S., Durairaj, A., Lu, J. & Davidson, W. (2010), "High-Density Lipoprotein Proteomics: Identifying New Drug Targets and Biomarkers by Understanding Functionality", *CURRENT CARDIOVASCULAR RISK REPORTS*, vol. 4, no. 1, pp. 1-8.

Hogan, M.C., Manganelli, L., Woollard, J.R., Masyuk, A.I., Masyuk, T.V., Tammachote, R., Huang, B.Q., Leontovich, A.A., Beito, T.G., Madden, B.J., Charlesworth, M.C., Torres, V.E., LaRusso, N.F., Harris, P.C. & Ward, C.J. (2009), "Characterization of PKD Protein-Positive Exosome-Like Vesicles", *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY*, vol. 20, no. 2, pp. 278-288.

Huang, D.W., Sherman, B.T. & Lempicki, R.A. (2009 a), "Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists", *NUCLEIC ACIDS RESEARCH*, vol. 37, no. 1, pp. 1-13.

Huang, D.W., Sherman, B.T. & Lempicki, R.A. (2009 b), "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources", *NATURE PROTOCOLS*, vol. 4, no. 1, pp. 44-57.

Kerjaschki, D., Schulze, M., Binder, S., Kain, R., Ojha, P., Susani, M., Horvat, R., Baker, P. & Couser, W. (1989), "Transcellular transport and membrane insertion of the C5b-9 membrane attack complex of complement by glomerular epithelial cells in experimental membranous nephropathy", *THE JOURNAL OF IMMUNOLOGY*, vol. 143, no. 2, pp. 546-552.

Krishnamoorthy, L., Bess, J.,Julian W., Preston, A.B., Nagashima, K. & Mahal, L.K. (2009), "HIV-1 and microvesicles from T cells share a common glycome, arguing for a common origin", *NATURE CHEMICAL BIOLOGY*, vol. 5, no. 4, pp. 244-250.

Krishnamurti, U., Rondeau, E., Sraer, J., Michael, A. & Tsilibary, E. (1997), "Alterations in human glomerular epithelial cells interacting with nonenzymatically glycosylated matrix", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 272, no. 44, pp. 27966-27970.

Manninen, A., Verkade, P., Le Lay, S., Torkko, J., Kasper, M., FÃ¼llekrug, J. & Simons, K. November 15, (2005), "Caveolin-1 Is Not Essential for Biosynthetic Apical Membrane Transport", *MOLECULAR AND CELLULAR BIOLOGY*, vol. 25, no. 22, pp. 10087-10096.

Marimuthu, A., O Meally, R.N., Chaerkady, R., Subbannayya, Y., Nanjappa, V., Kumar, P., Kelkar, D.S., Pinto, S.M., Sharma, R., Renuse, S., Goel, R., Christopher, R., Delanghe, B., Cole, R.N., Harsha, H.C. & Pandey, A. (2011), "A Comprehensive Map of the Human Urinary Proteome", *JOURNAL OF PROTEOME RESEARCH*, vol. 10, no. 6, pp. 2734-2743.

Mathivanan, S., Fahner, C.J., Reid, G.E. & Simpson, R.J. (2012), "ExoCarta 2012: database of exosomal proteins, RNA and lipids", *NUCLEIC ACIDS RESEARCH*, vol. 40, no. D1, pp. D1241-D1244.

Mayer, M.M. (1961), "On the Destruction of Erythrocytes and Other Cells by Antibody and Complement", *CANCER RESEARCH*, vol. 21, no. 9, pp. 1262-1269.

Morgan, B., Dankert, J. & Esser, A. (1987), "Recovery of human-neutrophils from complement attack - removal of the membrane attack complex by endocytosis and exocytosis", *JOURNAL OF IMMUNOLOGY*, vol. 138, no. 1, pp. 246-253.

Mott, J., Khalifah, R., Nagase, H., Shield, C., Hudson, J. & Hudson, B. (1997), "Nonenzymatic glycation of type IV collagen and matrix metalloproteinase susceptibility", *KIDNEY INTERNATIONAL*, vol. 52, no. 5, pp. 1302-1312.

Pisitkun, T., Shen, R. & Knepper, M. (2004), "Identification and proteomic profiling of exosomes in human urine", *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, vol. 101, no. 36, pp. 13368-13373.

Raabe, H., Hopner, J., Notbohm, H., Sinnecker, G., Kruse, K. & Muller, P. (1998), "Biochemical and biophysical alterations of the 7S and NC1 domain of collagen IV from human diabetic kidneys", *DIABETOLOGIA*, vol. 41, no. 9, pp. 1073-1079.

Shiflett, A., Bishop, J., Pahwa, A. & Hajduk, S. (2005), "Human high density lipoproteins are platforms for the assembly of multi-component innate immune complexes", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 280, no. 38, pp. 32578-32585.

Silbiger, S., Crowley, S., Shan, Z., Brownlee, M., Satriano, J. & Schlondorff, D. (1993), "Nonenzymatic glycation of mesangial matrix and prolonged exposure of mesangial matrix to elevated glucose reduces collagen-synthesis and proteoglycan charge", *KIDNEY INTERNATIONAL*, vol. 43, no. 4, pp. 853-864.

Vaisar, T., Pennathur, S., Green, P.S., Gharib, S.A., Hoofnagle, A.N., Cheung, M.C., Byun, J., Vuletic, S., Kassim, S., Singh, P., Chea, H., Knopp, R.H., Brunzell, J., Geary, R., Chait, A., Zhao, X., Elkon, K., Marcovina, S., Ridker, P., Oram, J.F. & Heinecke, J.W. (2007), "Shotgun proteomics implicates protease inhibition and complement activation in the antiinflammatory properties of HDL", *JOURNAL OF CLINICAL INVESTIGATION*, vol. 117, no. 3, pp. 746-756.

Vella, L.J., Sharples, R.A., Lawson, V.A., Masters, C.L., Cappai, R. & Hill, A.F. (2007), "Packaging of prions into exosomes is associated with a novel pathway of PrP processing", *JOURNAL OF PATHOLOGY*, vol. 211, no. 5, pp. 582-590.

Wang, Z., Hill, S., Luther, J.M., Hachey, D.L. & Schey, K.L. (2011), "Proteomic analysis of urine exosomes by multidimensional protein identification technology (MudPIT)", *PROTEOMICS*, vol. 12, no. 2, pp. 329-338.

Supplementary table S3.1: Complete non-redundant list of proteins identified in the high speed pellet (P200,000g). Uniprot accession and gene names are given.

UNIPROT_ACCESSION	Gene name
P07339	cathepsin D
P19440	gamma-glutamyltransferase light chain 3; gamma-glutamyltransferase 4 pseudogene; gamma-glutamyltransferase 2; gamma-glutamyltransferase 1; gamma-glutamyltransferase light chain 5 pseudogene
P63104	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide
P04279	semenogelin I
O00560	syndecan binding protein (syntenin)
Q9NR99	matrix-remodelling associated 5
P15941	mucin 1, cell surface associated
Q14624	inter-alpha (globulin) inhibitor H4 (plasma Kallikrein-sensitive glycoprotein)
P33908	mannosidase, alpha, class 1A, member 1
P02766	Transthyretin
Q9H6S3	EPS8-like 2
P01011	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3
P50148	guanine nucleotide binding protein (G protein), q polypeptide
P05546	serpin peptidase inhibitor, clade D (heparin cofactor), member 1
O14773	tripeptidyl peptidase I
Q9H444	chromatin modifying protein 4B
O00391	quiescin Q6 sulfhydryl oxidase 1
Q9NQ84	G protein-coupled receptor, family C, group 5, member C
Q92896	golgi apparatus protein 1
Q9Y287	integral membrane protein 2B
Q8TCD5	5', 3'-nucleotidase, cytosolic
Q99816	tumor susceptibility gene 101
Q8IV08	phospholipase D family, member 3
P05543	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 7
Q8WWI5	solute carrier family 44, member 1
Q9UBD6	Rh family, C glycoprotein
P31949	S100 calcium binding protein A11; S100 calcium binding protein A11 pseudogene
P08582	antigen p97 (melanoma associated) identified by monoclonal antibodies 133.2 and 96.5
P69905	hemoglobin, alpha 2; hemoglobin, alpha 1
P62937	similar to TRIMCyp; peptidylprolyl isomerase A (cyclophilin A); peptidylprolyl isomerase A (cyclophilin A)-like 3
Q9HAT2	sialic acid acetylesterase
Q96DG6	carboxymethylenebutenolidase homolog (Pseudomonas)
P02675	fibrinogen beta chain
Q6V0I7	FAT tumor suppressor homolog 4 (Drosophila)
P53990	similar to CG10103; KIAA0174
Q03154	aminoacylase 1
P13987	CD59 molecule, complement regulatory protein
P34896	serine hydroxymethyltransferase 1 (soluble)
Q96KP4	CNDP dipeptidase 2 (metallopeptidase M20 family)

P19827	inter-alpha (globulin) inhibitor H1
Q9UI12	ATPase, H+ transporting, lysosomal 50/57kDa, V1 subunit H
Q5D862	filaggrin family member 2
Q13510	N-acylsphingosine amidohydrolase (acid ceramidase) 1
P68371	tubulin, beta 2C
P22732	solute carrier family 2 (facilitated glucose/fructose transporter), member 5
P01008	serpin peptidase inhibitor, clade C (antithrombin), member 1
P30039	phenazine biosynthesis-like protein domain containing
P07998	ribonuclease, RNase A family, 1 (pancreatic)
P29622	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 4
O00159	myosin IC
Q12913	protein tyrosine phosphatase, receptor type, J
O43866	CD5 molecule-like
O15162	phospholipid scramblase 1
P01024	similar to Complement C3 precursor; complement component 3; hypothetical protein LOC100133511
P35030	protease, serine, 3
Q00610	clathrin, heavy chain (Hc)
P61224	RAP1B, member of RAS oncogene family
Q9Y644	RFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase
P29508	serpin peptidase inhibitor, clade B (ovalbumin), member 3
P29992	guanine nucleotide binding protein (G protein), alpha 11 (Gq class)
Q9UHL4	dipeptidyl-peptidase 7
Q12805	EGF-containing fibulin-like extracellular matrix protein 1
P11597	cholesteryl ester transfer protein, plasma
Q9H3G5	carboxypeptidase, vitellogenic-like
P98160	heparan sulfate proteoglycan 2
Q96IU4	abhydrolase domain containing 14B
P10253	glucosidase, alpha; acid
P41222	prostaglandin D2 synthase, hematopoietic; prostaglandin D2 synthase 21kDa (brain)
Q12907	lectin, mannose-binding 2
P21810	Biglycan
P25815	S100 calcium binding protein P
P07858	cathepsin B
P12277	creatine kinase, brain
Q7LBR1	chromatin modifying protein 1B
P01861	immunoglobulin heavy constant gamma 4 (G4m marker)
P10909	Clusterin
P50395	GDP dissociation inhibitor 2
P20073	annexin A7
P08183	ATP-binding cassette, sub-family B (MDR/TAP), member 1
P31946	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, beta polypeptide
P10153	ribonuclease, RNase A family, 2 (liver, eosinophil-derived neurotoxin)
P01857	immunoglobulin heavy constant gamma 1 (G1m marker); immunoglobulin heavy constant mu; immunoglobulin heavy variable 3-7; immunoglobulin heavy constant gamma 3 (G3m marker); immunoglobulin heavy variable 3-11 (gene/pseudogene); immunoglobulin heavy variable 4-31; immunoglobulin heavy locus

P01871	immunoglobulin heavy constant gamma 1 (G1m marker); immunoglobulin heavy constant mu; immunoglobulin heavy variable 3-7; immunoglobulin heavy constant gamma 3 (G3m marker); immunoglobulin heavy variable 3-11 (gene/pseudogene); immunoglobulin heavy variable 4-31; immunoglobulin heavy locus
P01860	immunoglobulin heavy constant gamma 1 (G1m marker); immunoglobulin heavy constant mu; immunoglobulin heavy variable 3-7; immunoglobulin heavy constant gamma 3 (G3m marker); immunoglobulin heavy variable 3-11 (gene/pseudogene); immunoglobulin heavy variable 4-31; immunoglobulin heavy locus
P51654	glypican 3
O96009	napsin A aspartic peptidase
P12109	collagen, type VI, alpha 1
P17174	glutamic-oxaloacetic transaminase 1, soluble (aspartate aminotransferase 1)
P09525	annexin A4
P08236	glucuronidase, beta
P01133	epidermal growth factor (beta-urogastrone)
P19801	amiloride binding protein 1 (amine oxidase (copper-containing))
P12035	keratin 3
P02748	complement component 9
P49721	proteasome (prosome, macropain) subunit, beta type, 2
P05026	ATPase, Na+/K+ transporting, beta 1 polypeptide
P15311	hypothetical protein LOC100129652; ezrin
P15924	Desmoplakin
P68104	eukaryotic translation elongation factor 1 alpha-like 7; eukaryotic translation elongation factor 1 alpha-like 3; similar to eukaryotic translation elongation factor 1 alpha 1; eukaryotic translation elongation factor 1 alpha 1
P11279	lysosomal-associated membrane protein 1
P35527	keratin 9
P02765	alpha-2-HS-glycoprotein
Q5SZK8	FRAS1 related extracellular matrix protein 2
Q7Z7M0	multiple EGF-like-domains 8
P04406	glyceraldehyde-3-phosphate dehydrogenase-like 6; hypothetical protein LOC100133042; glyceraldehyde-3-phosphate dehydrogenase
P36955	serpin peptidase inhibitor, clade F (alpha-2 antiplasmin, pigment epithelium derived factor), member 1
P05154	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 5
Q99519	sialidase 1 (lysosomal sialidase)
O75636	ficolin (collagen/fibrinogen domain containing) 3 (Hakata antigen)
Q8WW52	family with sequence similarity 151, member A
Q86YZ3	Hornerin
P05109	S100 calcium binding protein A8
P30041	peroxiredoxin 6
P15289	arylsulfatase A
Q9HBJ8	transmembrane protein 27
P23526	Adenosylhomocysteinase
P14618	similar to Pyruvate kinase, isozymes M1/M2 (Pyruvate kinase muscle isozyme) (Cytosolic thyroid hormone-binding protein) (CTHBP) (THBP1); pyruvate kinase, muscle
P63261	actin, gamma 1
P01042	kininogen 1
P01764	immunoglobulin heavy variable 3-23; immunoglobulin heavy variable group
P04196	histidine-rich glycoprotein
P16278	galactosidase, beta 1
O75340	aryl-hydrocarbon receptor repressor; programmed cell death 6

Q9C0H2	tweety homolog 3 (Drosophila)
Q8N0V5	glucosaminyl (N-acetyl) transferase 2, I-branching enzyme (I blood group)
P09467	fructose-1,6-bisphosphatase 1
P21266	glutathione S-transferase mu 3 (brain)
Q53GD3	solute carrier family 44, member 4
P68871	hemoglobin, beta
O95865	dimethylarginine dimethylaminohydrolase 2
P60022	defensin, beta 1
P01593	similar to hCG1642538
P00966	argininosuccinate synthetase 1
P08962	CD63 molecule
P00751	complement factor B
P12273	prolactin-induced protein
P61626	lysozyme (renal amyloidosis)
P04746	amylase, alpha 2A (pancreatic)
P01023	alpha-2-macroglobulin
P23284	peptidylprolyl isomerase B (cyclophilin B)
P39059	collagen, type XV, alpha 1
P24855	deoxyribonuclease I
Q6EMK4	Vasorin
P11117	acid phosphatase 2, lysosomal
Q9Y6W3	calpain 7
P34059	galactosamine (N-acetyl)-6-sulfate sulfatase
P19835	carboxyl ester lipase (bile salt-stimulated lipase)
Q9Y6R7	Fc fragment of IgG binding protein; similar to IgGFc-binding protein precursor (FcgammaBP) (Fcgamma-binding protein antigen)
P02753	retinol binding protein 4, plasma
Q7Z4W1	dicarbonyl/L-xylulose reductase
P60033	CD81 molecule
Q08380	lectin, galactoside-binding, soluble, 3 binding protein
P08603	complement factor H
Q7Z794	keratin 77
Q9H1C7	chromosome 5 open reading frame 32
P35858	insulin-like growth factor binding protein, acid labile subunit
P01876	immunoglobulin heavy constant alpha 1
P02649	hypothetical LOC100129500; apolipoprotein E
Q04695	keratin 17; keratin 17 pseudogene 3
P02774	group-specific component (vitamin D binding protein)
P14384	carboxypeptidase M
O60814	histone cluster 1, H2bk
P04350	tubulin, beta 4
O95336	6-phosphogluconolactonase
P01880	immunoglobulin heavy constant delta
P12821	angiotensin I converting enzyme (peptidyl-dipeptidase A) 1
P00747	Plasminogen

P02746	complement component 1, q subcomponent, B chain
Q6UXG3	CD300 molecule-like family member g
P00450	ceruloplasmin (ferroxidase)
P06280	galactosidase, alpha
Q96FQ6	S100 calcium binding protein A16
Q9H853	tubulin, alpha 4b (pseudogene)
P00558	phosphoglycerate kinase 1
P15309	acid phosphatase, prostate
P35908	keratin 2
O00322	uroplakin 1A
P00918	carbonic anhydrase II
P04264	keratin 1
P02533	keratin 14
Q8IWA5	solute carrier family 44, member 2
P08294	superoxide dismutase 3, extracellular
P02788	Lactotransferrin
P62805	histone cluster 1, H4l; histone cluster 1, H4k; histone cluster 4, H4; histone cluster 1, H4h; histone cluster 1, H4j; histone cluster 1, H4i; histone cluster 1, H4d; histone cluster 1, H4c; histone cluster 1, H4f; histone cluster 1, H4e; histone cluster 1, H4b; histone cluster 1, H4a; histone cluster 2, H4a; histone cluster 2, H4b
P60709	actin, beta
Q8N271	prominin 2
P48061	chemokine (C-X-C motif) ligand 12 (stromal cell-derived factor 1)
P05156	complement factor I
P02751	fibronectin 1
P08238	heat shock protein 90kDa alpha (cytosolic), class B member 1
P06396	gelsolin (amyloidosis, Finnish type)
P54802	N-acetylglucosaminidase, alpha-
P00749	plasminogen activator, urokinase
Q5VW32	chromosome 1 open reading frame 58
Q99835	smoothened homolog (Drosophila)
P60953	cell division cycle 42 (GTP binding protein, 25kDa); cell division cycle 42 pseudogene 2
Q9NS93	transmembrane 7 superfamily member 3
Q86UD1	OAF homolog (Drosophila)
Q6UX06	olfactomedin 4
Q6UX73	chromosome 16 open reading frame 89
P01031	complement component 5
P19823	inter-alpha (globulin) inhibitor H2
P05155	serpin peptidase inhibitor, clade G (C1 inhibitor), member 1
P55083	microfibrillar-associated protein 4
P42785	prolylcarboxypeptidase (angiotensinase C)
O75882	Attractin
P62879	guanine nucleotide binding protein (G protein), beta polypeptide 2
P02749	apolipoprotein H (beta-2-glycoprotein I)
Q9NP79	Vps20-associated 1 homolog (S. cerevisiae)
P05090	apolipoprotein D
P22352	glutathione peroxidase 3 (plasma)

P00734	coagulation factor II (thrombin)
P04217	alpha-1-B glycoprotein
P19652	orosomucoid 2
P02760	alpha-1-microglobulin/bikunin precursor
P29972	aquaporin 1 (Colton blood group)
P02679	fibrinogen gamma chain
P08195	solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2
P01009	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1
P08185	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 6
P50995	annexin A11
P03951	coagulation factor XI
P06312	immunoglobulin kappa variable 4-1
P08133	annexin A6
P51884	Lumican
Q9Y646	plasma glutamate carboxypeptidase
P15291	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 1
Q99988	growth differentiation factor 15
P08779	keratin 16; keratin type 16-like
Q12794	hyaluronoglucosaminidase 1
O75594	peptidoglycan recognition protein 1
P07357	complement component 8, alpha polypeptide
P08697	serpin peptidase inhibitor, clade F (alpha-2 antiplasmin, pigment epithelium derived factor), member 2
P07996	thrombospondin 1
P07195	lactate dehydrogenase B
Q93088	betaine-homocysteine methyltransferase
Q9BQE3	tubulin, alpha 1c
O43895	X-prolyl aminopeptidase (aminopeptidase P) 2, membrane-bound
P01877	immunoglobulin heavy constant alpha 2 (A2m marker)
P35241	Radixin
O75787	ATPase, H+ transporting, lysosomal accessory protein 2
P06870	kallikrein 1
P11142	heat shock 70kDa protein 8
P10599	Thioredoxin
P21399	aconitase 1, soluble
O75351	vacuolar protein sorting 4 homolog B (S. cerevisiae)
P55017	solute carrier family 12 (sodium/chloride transporters), member 3
P04083	annexin A1
Q9HC84	mucin 5B, oligomeric mucus/gel-forming
Q96FN5	kinesin family member 12
P02763	orosomucoid 1
P16444	dipeptidase 1 (renal)
P13647	keratin 5
P53634	cathepsin C
O75955	flotillin 1

P02750	leucine-rich alpha-2-glycoprotein 1
Q86T13	C-type lectin domain family 14, member A
P68366	tubulin, alpha 4a
P07355	annexin A2 pseudogene 3; annexin A2; annexin A2 pseudogene 1
P02743	amyloid P component, serum
Q96RW7	hemicentin 1
P04003	complement component 4 binding protein, alpha
P22891	protein Z, vitamin K-dependent plasma glycoprotein
Q92673	sortilin-related receptor, L(DLR class) A repeats-containing
P00738	haptoglobin-related protein; haptoglobin
Q14393	similar to growth arrest-specific 6; growth arrest-specific 6
P14923	junction plakoglobin
P02790	Hemopexin
P27487	dipeptidyl-peptidase 4
P0C0L4	complement component 4A (Rodgers blood group)
Q14CN4	keratin 72
Q3LXA3	dihydroxyacetone kinase 2 homolog (S. cerevisiae)
P62258	similar to 14-3-3 protein epsilon (14-3-3E) (Mitochondrial import stimulation factor L subunit) (MSF L); tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, epsilon polypeptide
Q9BYF1	angiotensin I converting enzyme (peptidyl-dipeptidase A) 2
O43653	prostate stem cell antigen
P09211	glutathione S-transferase pi 1
P61970	nuclear transport factor 2
P27105	Stomatin
P01591	immunoglobulin J polypeptide, linker protein for immunoglobulin alpha and mu polypeptides
Q9UBV8	penta-EF-hand domain containing 1
Q8WVN6	secreted and transmembrane 1
P31639	solute carrier family 5 (sodium/glucose cotransporter), member 2
Q9UN37	vacuolar protein sorting 4 homolog A (S. cerevisiae)
P61026	RAB10, member RAS oncogene family
P06744	glucose phosphate isomerise
P06733	enolase 1, (alpha)
Q92820	gamma-glutamyl hydrolase (conjugase, folylpolygammaglutamyl hydrolase)
P12429	annexin A3
Q02383	semenogelin II
P0C0L5	complement component 4B (Chido blood group)
P10643	complement component 7
P02792	similar to ferritin, light polypeptide; ferritin, light polypeptide
P01034	cystatin C
P12814	actinin, alpha 1
Q02413	desmoglein 1
P06727	apolipoprotein A-IV
P12111	collagen, type VI, alpha 3
Q14314	fibrinogen-like 2
Q14019	coactosin-like 1 (Dictyostelium)

P01833	polymeric immunoglobulin receptor
P08758	annexin A5
Q15485	ficolin (collagen/fibrinogen domain containing lectin) 2 (hucolin)
O43451	maltase-glucoamylase (alpha-glucosidase)
Q14108	scavenger receptor class B, member 2
P04114	apolipoprotein B (including Ag(x) antigen)
Q6W4X9	mucin 6, oligomeric mucus/gel-forming
O00187	mannan-binding lectin serine peptidase 2
Q8TF66	leucine rich repeat containing 15
P02747	complement component 1, q subcomponent, C chain
P06702	S100 calcium binding protein A9
P18428	lipopolysaccharide binding protein
Q7Z5L0	vitelline membrane outer layer 1 homolog (chicken)
P07358	complement component 8, beta polypeptide
P21926	CD9 molecule
O60637	tetraspanin 3
P54652	heat shock 70kDa protein 2
P04004	Vitronectin
P08107	heat shock 70kDa protein 1A; heat shock 70kDa protein 1B
P21796	voltage-dependent anion channel 1; similar to voltage-dependent anion channel 1
P14550	aldo-keto reductase family 1, member A1 (aldehyde reductase)
P10619	cathepsin A
Q7RTS7	keratin 74
O43490	prominin 1
P09543	2',3'-cyclic nucleotide 3' phosphodiesterase
P40925	malate dehydrogenase 1, NAD (soluble)
O60494	cubilin (intrinsic factor-cobalamin receptor)
Q9UGT4	sushi domain containing 2
Q8NFJ5	G protein-coupled receptor, family C, group 5, member A
P81605	Dermcidin
P05164	Myeloperoxidase
Q8N1N4	keratin 78
P31944	caspase 14, apoptosis-related cysteine peptidase
Q96SA4	serine incorporator 2
P04066	fucosidase, alpha-L- 1, tissue
P22792	carboxypeptidase N, polypeptide 2
P02671	fibrinogen alpha chain
Q07075	glutamyl aminopeptidase (aminopeptidase A)
P19961	amylase, alpha 2B (pancreatic)
O43633	chromatin modifying protein 2A
P13164	interferon induced transmembrane protein 1 (9-27)
Q96DA0	zymogen granule protein 16 homolog B (rat)
Q495M3	solute carrier family 36 (proton/amino acid symporter), member 2
Q96KN2	carnosine dipeptidase 1 (metallopeptidase M20 family)

P02768	Albumin
P01859	immunoglobulin heavy constant gamma 2 (G2m marker)
Q9HB40	serine carboxypeptidase 1
Q8WUM4	programmed cell death 6 interacting protein
P07911	Uromodulin
P15144	alanyl (membrane) aminopeptidase
Q16706	mannosidase, alpha, class 2A, member 1
P01834	similar to hCG26659; immunoglobulin kappa constant; similar to Ig kappa chain V-I region HK102 precursor
P19971	thymidine phosphorylase
Q9H223	EH-domain containing 4
Q16769	glutaminyl-peptide cyclotransferase
P07288	kallikrein-related peptidase 3
Q9HBA9	folate hydrolase 1B
P45877	peptidylprolyl isomerase C (cyclophilin C)
P13646	keratin 13
P02647	apolipoprotein A-I
P13473	lysosomal-associated membrane protein 2
Q9Y2S2	crystallin, lambda 1
P25787	proteasome (prosome, macropain) subunit, alpha type, 2
Q9UKU6	thyrotropin-releasing hormone degrading enzyme
P59665	defensin, alpha 1
P16152	carbonyl reductase 1
Q9UIQ6	leucyl/cystinyl aminopeptidase
P08571	CD14 molecule
P19012	keratin 15
P43251	Biotinidase
P01019	angiotensinogen (serpin peptidase inhibitor, clade A, member 8)
P52209	phosphogluconate dehydrogenase
P60174	TPI1 pseudogene; triosephosphate isomerase 1
P00338	lactate dehydrogenase A
P61586	ras homolog gene family, member A
P63092	GNAS complex locus
P02538	keratin 6A
P53801	pituitary tumor-transforming 1 interacting protein
P27169	paraoxonase 1
P54793	arylsulfatase F
P31151	S100 calcium binding protein A7
Q6I9Y2	THO complex 7 homolog (Drosophila)
P63096	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 1
Q16270	insulin-like growth factor binding protein 7
P51688	N-sulfoglucosamine sulfohydrolase
P14543	nidogen 1
P08473	membrane metallo-endopeptidase
P01617	hypothetical LOC440786

P04745	amylase, alpha 1A (salivary); amylase, alpha 1B (salivary); amylase, alpha 1C (salivary)
P27105	phosphatidylethanolamine binding protein 1
P13645	keratin 10
Q6UWR7	ectonucleotide pyrophosphatase/phosphodiesterase 6
P62873	guanine nucleotide binding protein (G protein), beta polypeptide 1
O00182	lectin, galactoside-binding, soluble, 9
P26038	Moesin
P49221	transglutaminase 4 (prostate)
P17050	N-acetylgalactosaminidase, alpha-
Q9BRK3	matrix-remodelling associated 8
P05062	aldolase B, fructose-bisphosphate
Q9HD89	Resistin
Q96P63	serpin peptidase inhibitor, clade B (ovalbumin), member 12
P08263	glutathione S-transferase alpha 1
P19013	keratin 4
P02787	Transferring
P04259	keratin 6B
Q9NZP8	complement component 1, r subcomponent-like
Q8WZ75	roundabout homolog 4, magic roundabout (Drosophila)
P02794	ferritin, heavy polypeptide 1; ferritin, heavy polypeptide-like 16; similar to ferritin, heavy polypeptide 1; ferritin, heavy polypeptide-like 3 pseudogene
Q9H8L6	multimerin 2
P25311	alpha-2-glycoprotein 1, zinc-binding pseudogene 1; alpha-2-glycoprotein 1, zinc-binding
Q9UKU9	angiopoietin-like 2
Q13621	solute carrier family 12 (sodium/potassium/chloride transporters), member 1
Q9BYZ2	lactate dehydrogenase A-like 6B
O00468	Agrin
P98164	low density lipoprotein-related protein 2
P01768	Ig heavy chain V-III region CAM
P0CG04	Ig lambda-1 chain C regions
P04220	Ig mu heavy chain disease protein
P01613	Ig kappa chain V-I region Ni
P01770	Ig heavy chain V-III region NIE
P01779	Ig heavy chain V-III region TUR
P01621	Ig kappa chain V-III region NG9 (Fragment)
P04208	Ig lambda chain V-I region WAH
P01769	Ig heavy chain V-III region GA
P01620	Ig kappa chain V-III region SIE
P01598	Ig kappa chain V-I region EU
P01714	Ig lambda chain V-III region SH
P0CG05	Ig lambda-2 chain C regions
P80748	Ig lambda chain V-III region LOI
P01777	Ig heavy chain V-III region TEI
P01766	Ig heavy chain V-III region BRO
P0CG48	Polyubiquitin-C

P01762	Ig heavy chain V-III region TRO
P01765	Ig heavy chain V-III region TIL
P04207	Ig kappa chain V-III region CLL (Rheumatoid factor)
P01781	Ig heavy chain V-III region GAL
P01619	Ig kappa chain V-III region B6
P01616	Ig kappa chain V-II region MIL
P01625	Ig kappa chain V-IV region Len

Supplementary table S3.2: The following genes were annotated and classified as being related to human diseases as evidenced by their OMIM IDs mapped using DAVID Bioinformatics Resources 6.7, National Institute of Allergy and Infectious Diseases (NIAID), NIH. Uniprot Id, gene name and OMIM disease names are given in the table.

ID	Gene Name	OMIM_DISEASE
P08183	ATP-binding cassette, sub-family B (MDR/TAP), member 1	Colchicine resistance,Inflammatory bowel disease 13, susceptibility to,
O75787	ATPase, H+ transporting, lysosomal accessory protein 2	Mental retardation, X-linked, with epilepsy,
P05026	ATPase, Na+/K+ transporting, beta 1 polypeptide	Blood pressure regulation QTL,
P13987	CD59 molecule, complement regulatory protein	CD59 deficiency,
Q12805	EGF-containing fibulin-like extracellular matrix protein 1	Doyne honeycomb degeneration of retina,Genome-wide association analysis identifies 20 loci that influence adult height,Many sequence variants affecting diversity of adult human height,
Q5SZK8	FRAS1 related extracellular matrix protein 2	Fraser syndrome,
P63092	GNAS complex locus	Acromegaly,McCune-Albright syndrome,Osseous heteroplasia, progressive,Pituitary ACTH secreting adenoma,Pituitary ACTH secreting adenoma, somatic,Pituitary ACTH secreting adenoma, somatic 219090,Prolonged bleeding time, brachydactyly and mental retardation,Pseudohypoparathyroidism Ia,Pseudohypoparathyroidism Ib,Pseudohypoparathyroidism, type Ia,Pseudohypoparathyroidism, type Ib,Pseudopseudohypoparathyroidism,Somatotrophinoma,
P17050	N-acetylgalactosaminidase, alpha-	Kanzaki disease,Schindler disease, type I,Schindler disease, type III,
P54802	N-acetylglucosaminidase, alpha-	mucopolysaccharidosis type IIIB,Sanfilippo syndrome, type B,
Q13510	N-acylsphingosine amidohydrolase (acid ceramidase) 1	Farber lipogranulomatosis,
P51688	N-sulfoglucosamine sulfohydrolase	Sanfilippo syndrome, type A,
P60174	TPI1 pseudogene; triosephosphate isomerase 1	Hemolytic anemia due to triosephosphate isomerase deficiency,
P15291	UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 1	Congenital disorder of glycosylation, type IId,
O43895	X-prolyl aminopeptidase (aminopeptidase P) 2, membrane-bound	Angioedema induced by ACE inhibitors, susceptibility to,
P60709	actin, beta	Dystonia, juvenile-onset,
P63261	actin, gamma 1	Deafness, autosomal dominant 20/26,
P23526	Adenosylhomocysteinase	Hypermethioninemia with deficiency of S-adenosylhomocysteine hydrolase,
P02768	Albumin	Analbuminemia,Dysalbuminemic hyperthyroxinemia,Dysalbuminemic hyperzincemia,
P05062	aldolase B, fructose-bisphosphate	Fructose intolerance,
P01023	alpha-2-macroglobulin	Alzheimer disease, susceptibility to,Emphysema due to alpha-2-macroglobulin deficiency,

Q03154	aminoacylase 1	Aminoacylase 1 deficiency,
P02743	amyloid P component, serum	?Amyloidosis, secondary, susceptibility to,
P12821	angiotensin I converting enzyme (peptidyl-dipeptidase A) 1	Alzheimer disease, susceptibility to,Angiotensin I-converting enzyme, benign serum increase,Diabetic nephropathy, susceptibility to,Microvascular complications of diabetes 3,Myocardial infarction, susceptibility to,Renal tubular dysgenesis,SARS, progression of,
P01019	angiotensinogen (serpin peptidase inhibitor, clade A, member 8)	Hypertension, essential, susceptibility to,Preeclampsia, susceptibility to,Renal tubular dysgenesis,
P02647	apolipoprotein A-I	A null mutation in human APOC3 confers a favorable plasma lipid profile and apparent cardioprotection,Amyloidosis, 3 or more types,ApoA-I and apoC-III deficiency, combined,Common variants at 30 loci contribute to polygenic dyslipidemia,Corneal clouding, autosomal recessive,Genome-wide association study identifies genes for biomarkers of cardiovascular disease: serum urate and dyslipidemia,Genome-wide scan identifies variation in MLXIPL associated with plasma triglycerides,Hypertriglyceridemia, one form,Hypoalphalipoproteinemia,Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts,Newly identified loci that influence lipid concentrations and risk of coronary artery disease,Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans,
P06727	apolipoprotein A-IV	A null mutation in human APOC3 confers a favorable plasma lipid profile and apparent cardioprotection,Common variants at 30 loci contribute to polygenic dyslipidemia,Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts,Newly identified loci that influence lipid concentrations and risk of coronary artery disease,Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans,
P04114	apolipoprotein B (including Ag(x) antigen)	Common variants at 30 loci contribute to polygenic dyslipidemia,Genome-wide association analysis of metabolic traits in a birth cohort from a founder population,Hypercholesterolemia, due to ligand-defective apo B,Hypobetalipoproteinemia,Hypobetalipoproteinemia, normotriglyceridemic,LDL-cholesterol concentrations: a genome-wide association study,Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts,Newly identified loci that influence lipid concentrations and risk of coronary artery disease,Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans,
P02749	apolipoprotein H (beta-2-glycoprotein I)	Apolipoprotein H deficiency,
P29972	aquaporin 1 (Colton blood group)	Aquaporin-1 deficiency,Blood group, Colton,
P00966	argininosuccinate synthetase 1	Citrullinemia,
P15289	arylsulfatase A	Metachromatic leukodystrophy,
P43251	Biotinidase	Biotinidase deficiency,
P00918	carbonic anhydrase II	Osteopetrosis, autosomal recessive 3, with renal tubular acidosis,Renal tubular acidosis-osteopetrosis syndrome,
P19835	carboxyl ester lipase (bile salt-stimulated lipase)	Maturity-onset diabetes of the young, type VIII,
Q9H3G5	carboxypeptidase, vitellogenic-like	Genome-wide association with diabetes-related traits in the Framingham Heart Study,
P10619	cathepsin A	Galactosialidosis,
P53634	cathepsin C	Haim-Munk syndrome,Papillon-Lefevre syndrome,Periodontitis, juvenile,
P07339	cathepsin D	Ceroid lipofuscinosis, neuronal, 10,
P00450	ceruloplasmin (ferroxidase)	Cerebellar ataxia,Hemosiderosis, systemic, due to aceruloplasminemia,Hypoceruloplasminemia, hereditary,
P48061	chemokine (C-X-C motif) ligand 12 (stromal cell-derived factor 1)	AIDS, resistance to,Genomewide association analysis of coronary artery disease,

P11597	cholesteryl ester transfer protein, plasma	CETP deficiency,Common genetic variation near MC4R is associated with waist circumference and insulin resistance,Common variants at 30 loci contribute to polygenic dyslipidemia,Genome-wide association analysis of metabolic traits in a birth cohort from a founder population,High density lipoprotein cholesterol level QTL 10,Hyperalphalipoproteinemia,Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts,Longevity, exceptional,Newly identified loci that influence lipid concentrations and risk of coronary artery disease,Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans,
Q9H444	chromatin modifying protein 4B	Cataract, posterior polar, 3,Cataract, posterior polar-3,
P00734	coagulation factor II (thrombin)	Dysprothrombinemia,Hyperprothrombinemia,Hypoprothrombinemia,
P03951	coagulation factor XI	Factor XI deficiency, autosomal dominant,Factor XI deficiency, autosomal recessive,
P12109	collagen, type VI, alpha 1	Bethlem myopathy,Ossification of the posterior longitudinal spinal ligaments,Ullrich congenital muscular dystrophy,
P12111	collagen, type VI, alpha 3	Bethlem myopathy,Ullrich congenital muscular dystrophy,
P02746	complement component 1, q subcomponent, B chain	C1q deficiency, type B,
P02747	complement component 1, q subcomponent, C chain	C1q deficiency, type C,
P0C0L4	complement component 4A (Rodgers blood group)	Blood group, Rodgers,C4 deficiency,Systemic lupus erythematosus, susceptibility to or protection against,
P0C0L5	complement component 4B (Chido blood group)	C4 deficiency,
P01031	complement component 5	C5 deficiency,Liver fibrosis, susceptibility to,
P10643	complement component 7	C7 deficiency,
P07357	complement component 8, alpha polypeptide	C8 deficiency, type I,
P07358	complement component 8, beta polypeptide	C8 deficiency, type II,
P02748	complement component 9	C9 deficiency,C9 deficiency with dermatomyositis,
P00751	complement factor B	Macular degeneration, age-related, reduced risk of,
P08603	complement factor H	Basal laminar drusen,Complement factor H deficiency,Complement factor H polymorphism in age-related macular degeneration,Factor H and factor H-like 1,Hemolytic-uremic syndrome,Macular degeneration, age-related, 4,Membranoproliferative glomerulonephritis with CFH deficiency,Myocardial infarction, susceptibility to,
P05156	complement factor I	C3b inactivator deficiency,Complement factor I deficiency,
O60494	cubilin (intrinsic factor-cobalamin receptor)	Megaloblastic anemia-1, Finnish type,
P01034	cystatin C	A genome-wide association for kidney function and endocrine-related traits in the NHLBI's Framingham Heart Study,Cerebral amyloid angiopathy,Macular degeneration, age-related, 11,
P24855	deoxyribonuclease I	Systemic lupus erythematosus, susceptibility to,
P81605	Dermcidin	Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes,
Q02413	desmoglein 1	Keratosis palmoplantaris striata I,
P15924	Desmoplakin	Arrhythmogenic right ventricular dysplasia 8,Dilated cardiomyopathy with woolly hair and keratoderma,Epidermolysis bullosa, lethal acantholytic,Keratosis palmoplantaris striata II,Skin fragility-woolly hair syndrome,
P06733	enolase 1, (alpha)	Enolase deficiency,
P01133	epidermal growth factor (beta-urogastrone)	Hypomagnesemia 4, renal,
P02794	ferritin, heavy polypeptide 1; ferritin, heavy polypeptide-like 16; similar to ferritin, heavy polypeptide 1; ferritin, heavy polypeptide-like 3 pseudogene	Iron overload, autosomal dominant,

P02671	fibrinogen alpha chain	Afibrinogenemia, congenital,Amyloidosis, hereditary renal,Dysfibrinogenemia, alpha type, causing bleeding diathesis,Dysfibrinogenemia, alpha type, causing recurrent thrombosis,
P02675	fibrinogen beta chain	Afibrinogenemia, congenital,Dysfibrinogenemia, beta type,Thrombophilia, dysfibrinogenemic,
P02679	fibrinogen gamma chain	Dysfibrinogenemia, gamma type,Hypofibrinogenemia, gamma type,Thrombophilia, dysfibrinogenemic,
P02751	fibronectin 1	Glomerulopathy with fibronectin deposits 2,Glomerulopathy, fibronectin,
P09467	fructose-1,6-bisphosphatase 1	Fructose-1,6-bidphosphatase deficiency,Fructose-bisphosphatase deficiency,
P04066	fucosidase, alpha-L- 1, tissue	Fucosidosis,
P34059	galactosamine (N-acetyl)-6-sulfate sulfatase	Mucopolysaccharidosis IVA,
P06280	galactosidase, alpha	Fabry disease,Fabry disease, cardiac variant,
P16278	galactosidase, beta 1	GM1-gangliosidosis,GM1-gangliosidosis, type I,GM1-gangliosidosis, type II,GM1-gangliosidosis, type III,Morquio syndrome B,Mucopolysaccharidosis IVB,
P19440	gamma-glutamyltransferase light chain 3; gamma-glutamyltransferase 4 pseudogene; gamma-glutamyltransferase 2; gamma-glutamyltransferase 1; gamma-glutamyltransferase light chain 5 pseudogene	A Genome-Wide Association Study Identifies Protein Quantitative Trait Loci (pQTLs),Gamma-glutamyltransferase, familial high serum,Glutathionuria,Population-based genome-wide association studies reveal six loci influencing plasma levels of liver enzymes,
P06396	gelsolin (amyloidosis, Finnish type)	Amyloidosis, Finnish type,
Q8N0V5	glucosaminyl (N-acetyl) transferase 2, I-branching enzyme (I blood group)	Adult i phenotype with congenital cataract,Adult i phenotype without cataract,Blood group, Ii,Juvenile congenital cataract,
P06744	glucose phosphate isomerise	Hemolytic anemia due to glucosephosphate isomerase deficiency,Hydrops fetalis, one form,
P10253	glucosidase, alpha; acid	Glycogen storage disease II,
P08236	glucuronidase, beta	Mucopolysaccharidosis VII,
Q07075	glutamyl aminopeptidase (aminopeptidase A)	Variants conferring risk of atrial fibrillation on chromosome 4q25,
P51654	glypican 3	Simpson-Golabi-Behmel syndrome, type 1,Wilms tumor, somatic,
Q92896	golgi apparatus protein 1	A pilot genome-wide association study of early-onset breast cancer,
P02774	group-specific component (vitamin D binding protein)	Graves disease, susceptibility to, 3,
P50148	guanine nucleotide binding protein (G protein), q polypeptide	Bleeding diathesis due to GNAQ deficiency,
P00738	haptoglobin-related protein; haptoglobin	Anhaptoglobinemia,Hypohaptoglobinemia,
Q96RW7	hemicentin 1	Macular degeneration, age-related, 1,
P69905	hemoglobin, alpha 2; hemoglobin, alpha 1	Erythremias, alpha-,Erythrocytosis,Heinz body anemia,Heinz body anemias, alpha-,Hemoglobin H disease,Hypochromic microcytic anemia,Methemoglobinemias, alpha-,Thalassemia, alpha-,Thalassemias, alpha-,
P68871	hemoglobin, beta	A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15,Erythremias, beta-,Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of beta-thalassemia,Heinz body anemias, beta-,Hereditary persistence of fetal hemoglobin,HPFH, deletion type,Methemoglobinemias, beta-,Sickle cell anemia,Thalassemia-beta, dominant inclusion-body,Thalassemias, beta-,
P98160	heparan sulfate proteoglycan 2	Dyssegmental dysplasia, Silverman-Handmaker type,Schwartz-Jampel syndrome, type 1,
P04196	histidine-rich glycoprotein	?Thrombophilia due to elevated HRG,Thrombophilia due to HRG deficiency,
Q12794	hyaluronoglucosaminidase 1	Mucopolysaccharidosis type IX,
P02649	hypothetical LOC100129500; apolipoprotein E	A genome-wide association study for late-ons*et al.,*zheimer's disease using DNA pooling,A high-density whole-genome association study reveals that APOE is the major susceptibility gene for sporadic late-ons*et al.,*zheimer's disease,Alzheimer disease-2,Candidate single-nucleotide polymorphisms from a

		genomewide association study of Alzheimer disease,Common SNPs in HMGCR in Micronesians and Whites Associated With LDL-Cholesterol Levels Affect Alternative Splicing of Exon13,Common variants at 30 loci contribute to polygenic dyslipidemia,Hyperlipoproteinemia, type III,Lipoprotein glomerulopathy,Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts,Loci Related to Metabolic-Syndrome Pathways Including LEPR, HNF1A, IL6R, and GCKR Associate with Plasma C-Reactive Protein: The Women's Genome Health Study,Macular degeneration, age-related,Myocardial infarction susceptibility,Newly identified loci that influence lipid concentrations and risk of coronary artery disease,Polymorphisms of the HNF1A Gene Encoding Hepatocyte Nuclear Factor-1 Alpha are Associated with C-Reactive Protein,Sea-blue histiocyte disease,Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans,Sorl1 as an Alzheimer's disease predisposition gene?,
P01857, P01871, P01860	immunoglobulin heavy constant gamma 1 (G1m marker); immunoglobulin heavy constant mu; immunoglobulin heavy variable 3-7; immunoglobulin heavy constant gamma 3 (G3m marker); immunoglobulin heavy variable 3-11 (gene/pseudogene); immunoglobulin heavy variable 4-31; immunoglobulin heavy locus	Agammaglobulinemia,
P01859	immunoglobulin heavy constant gamma 2 (G2m marker)	IgG2 deficiency, selective,
P35858	insulin-like growth factor binding protein, acid labile subunit	Acid-labile subunit, deficiency of,
Q9Y287	integral membrane protein 2B	Dementia, familial British,Dementia, familial Danish,
Q14624	inter-alpha (globulin) inhibitor H4 (plasma Kallikrein-sensitive glycoprotein)	Hypercholesterolemia, susceptibility to,
P14923	junction plakoglobin	Arrhythmogenic right ventricular dysplasia, familial, 12,Naxos disease,
P06870	kallikrein 1	Kallikrein, decreased urinary activity of,
P07288	kallikrein-related peptidase 3	Multiple newly identified loci associated with prostate cancer susceptibility,
P04264	keratin 1	Cyclic ichthyosis with epidermolytic hyperkeratosis,Epidermolytic hyperkeratosis,Ichthyosis histrix, Curth-Macklin type,Keratosis palmoplantaria striata,Keratosis palmoplantaris striata III,Unna-Thost disease, nonepidermolytic,
P13645	keratin 10	Epidermolytic hyperkeratosis,Ichthyosis, cyclic, with epidermolytic hyperkeratosis,Keratosis palmaris et plantaris,Nevus, epidermal, epidermolytic hyperkeratotic type,
P13646	keratin 13	White sponge nevus,
P02533	keratin 14	Dermatopathia pigmentosa reticularis,Epidermolysis bullosa simplex, Dowling-Meara type,Epidermolysis bullosa simplex, Koebner type,Epidermolysis bullosa simplex, Koebner, Dowling-Meara, and Weber-Cockayne types, 131900, 131760,Epidermolysis bullosa simplex, recessive,Epidermolysis bullosa simplex, Weber-Cockayne type,Naegeli-Franceschetti-Jadassohn syndrome,
P08779	keratin 16; keratin type 16-like	Pachyonychia congenita, Jadassohn-Lewandowsky type,Palmoplantar keratoderma, nonepidermolytic,Palmoplantar verrucous nevus, unilateral,
Q04695	keratin 17; keratin 17 pseudogene 3	Pachyonychia congenita, Jackson-Lawler type,Steatocystoma multiplex,
P35908	keratin 2	Ichthyosis bullosa of Siemens,
P12035	keratin 3	Meesmann corneal dystrophy,
P19013	keratin 4	White sponge nevus,
P13647	keratin 5	Dowling-Degos disease,Epidermolysis bullosa simplex with migratory circinate erythema,Epidermolysis bullosa simplex with mottled pigmentation,Epidermolysis bullosa simplex, Dowling-Meara type,Epidermolysis bullosa simplex, Koebner type,Epidermolysis bullosa simplex, Koebner, Dowling-Meara,

		and Weber-Cockayne types, 131900, 131760,Epidermolysis bullosa simplex, Weber-Cockayne type,
P02538	keratin 6A	Pachyonychia congenita, Jadassohn-Lewandowsky type,
P04259	keratin 6B	Pachyonychia congenita, Jackson-Lawler type,
P35527	keratin 9	Epidermolytic palmoplantar keratoderma,
P01042	kininogen 1	Fitzgerald factor deficiency,High molecular weight kininogen deficiency,High-molecular-weight kininogen deficiency,Kininogen deficiency,
P00338	lactate dehydrogenase A	Exertional myoglobinuria due to deficiency of LDH-A,
P07195	lactate dehydrogenase B	Lactate dehydrogenase-B deficiency,
P98164	low density lipoprotein-related protein 2	Donnai-Barrow syndrome,
P13473	lysosomal-associated membrane protein 2	Glycogen storage disease IIb,
P61626	lysozyme (renal amyloidosis)	Amyloidosis, renal,Genomic association analysis suggests chromosome 12 locus influencing antihypertensive response to thiazide diuretic,Many sequence variants affecting diversity of adult human height,
O00187	mannan-binding lectin serine peptidase 2	MASP2 deficiency,
P08473	membrane metallo-endopeptidase	Membranous glomerulonephritis, antenatal,Neutral endopeptidase deficiency,
P05164	Myeloperoxidase	Alzheimer disease, susceptibility to,Lung cancer, protection against, in smokers,Myeloperoxidase deficiency,
P27169	paraoxonase 1	Coronary artery disease, susceptibility to,Coronary artery spasm, susceptibility to,Genome-wide association scan identifies candidate polymorphisms associated with differential response to anti-TNF treatment in Rheumatoid Arthritis,Organophosphate poisoning, sensitivity to,
P00558	phosphoglycerate kinase 1	Myoglobinuria/hemolysis due to PGK deficiency,Phosphoglycerate kinase 1 deficiency,
P00747	Plasminogen	Conjunctivitis, ligneous,Plasminogen deficiency, types I and II,Plasminogen Tochigi disease,Thrombophilia, dysplasminogenemic,
P00749	plasminogen activator, urokinase	Alzheimer disease, late-onset, susceptibility to,
P01833	polymeric immunoglobulin receptor	IgA nephropathy, susceptibility to,
O43490	prominin 1	Cone-rod dystrophy 12,Macular dystrophy 2, Bull's eye,Macular dystrophy, retinal, 2,Retinal degeneration, autosomal recessive, prominin-related,Retinitis pigmentosa-41,Stargardt disease 4,
P22891	protein Z, vitamin K-dependent plasma glycoprotein	Genome-wide association and linkage analyses of hemostatic factors and hematological phenotypes in the Framingham Heart Study,
Q12913	protein tyrosine phosphatase, receptor type, J	Colon cancer, somatic,
P35241	Radixin	Deafness, autosomal recessive, 24,
Q9HD89	Resistin	Diabetes mellitus, noninsulin-dependent, susceptibility to,Hypertension, insulin resistance-related, susceptibility to,
P02753	retinol binding protein 4, plasma	Retinol binding protein, deficiency of,
Q14108	scavenger receptor class B, member 2	Action myoclonus-renal failure syndrome,
P01009	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1	Emphysema,Emphysema-cirrhosis,Hemorrhagic diathesis due to 'antithrombin' Pittsburgh,Hemorrhagic diathesis due to `antithrombin' Pittsburgh,Pulmonary disease, chronic obstructive, susceptibility to,
P01011	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3	Alpha-1-antichymotrypsin deficiency,Cerebrovascular disease, occlusive,
P05154	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 5	Protein C inhibitor deficiency,
P08185	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 6	Corticosteroid-binding globulin deficiency,Transcortin deficiency,
P05543	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 7	Thyroxine-binding globulin deficiency,
P01008	serpin peptidase inhibitor, clade C (antithrombin), member 1	Antithrombin III deficiency,
P05546	serpin peptidase inhibitor, clade D (heparin cofactor), member 1	Thrombophilia due to heparin cofactor II deficiency,

P08697	serpin peptidase inhibitor, clade F (alpha-2 antiplasmin, pigment epithelium derived factor), member 2	Plasmin inhibitor deficiency,
P05155	serpin peptidase inhibitor, clade G (C1 inhibitor), member 1	Angioedema, hereditary,Angioedema, hereditary, types I and II,Complement component 4, partial deficiency of,
Q99519	sialidase 1 (lysosomal sialidase)	Sialidosis, type I,Sialidosis, type II,
P62258	similar to 14-3-3 protein epsilon (14-3-3E) (Mitochondrial import stimulation factor L subunit) (MSF L); tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, epsilon polypeptide	Miller-Dieker lissencephaly,
P53990	similar to CG10103; KIAA0174	Conduct disorder and ADHD: Evaluation of conduct problems as a categorical and quantitative trait in the international multicentre ADHD genetics study,
P01024	similar to Complement C3 precursor; complement component 3; hypothetical protein LOC100133511	C3 deficiency,Macular degeneration, age-related, 9,
P02792	similar to ferritin, light polypeptide; ferritin, light polypeptide	Basal ganglia disease, adult-onset,Hyperferritinemia-cataract syndrome,
P01834	similar to hCG26659; immunoglobulin kappa constant; similar to Ig kappa chain V-I region HK102 precursor	Kappa light chain deficiency,
Q99835	smoothened homolog (Drosophila)	Basal cell carcinoma, somatic,
P55017	solute carrier family 12 (sodium/chloride transporters), member 3	Gitelman syndrome,
Q13621	solute carrier family 12 (sodium/potassium/chloride transporters), member 1	Bartter syndrome, type 1,
P31639	solute carrier family 5 (sodium/glucose cotransporter), member 2	Renal glucosuria,
Q92673	sortilin-related receptor, L(DLR class) A repeats-containing	Alzheimer disease, pathogenesis, association with,
P08294	superoxide dismutase 3, extracellular	Superoxide dismutase, elevated extracellular,
P19971	thymidine phosphorylase	Mitochondrial neurogastrointestinal encephalomyopathy syndrome,
P02787	Transferring	Atransferrinemia,Iron deficiency anemia, susceptibility to,Variants in TF and HFE explain approximately 40% of genetic variation in serum-transferrin levels,
P02766	Transthyretin	Amyloid neuropathy, familial, several allelic types,Amyloid polyneuropathy, several types,Amyloidosis, senile systemic,Carpal tunnel syndrome, familial,Dystransthyretinemic hyperthyroxinemia,
O14773	tripeptidyl peptidase I	Ceroid-lipofuscinosis, neuronal 2, classic late infantile,
P68366	tubulin, alpha 4a	Many sequence variants affecting diversity of adult human height,
Q99816	tumor susceptibility gene 101	Breast cancer,
P63104	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide	Conduct disorder and ADHD: Evaluation of conduct problems as a categorical and quantitative trait in the international multicentre ADHD genetics study,
P07911	Uromodulin	Glomerulocystic kidney disease with hyperuricemia and isosthenuria,Hyperuricemic nephropathy, familial juvenile,Medullary cystic kidney disease 2,Medullary cystic kidney disease 2 (autosomal dominant),

Supplementary table S3.3: The following genes were annotated and classified as having genetic association to cardiovascular diseases using DAVID Bioinformatics resources 6.7, National Institute of Allergy and Infectious Diseases (NIAID), NIH. Uniprot Ids and gene names are given.

ID	Gene Name
P09543	2',3'-cyclic nucleotide 3' phosphodiesterase
P08183	ATP-binding cassette, sub-family B (MDR/TAP), member 1
P08571	CD14 molecule
P63092	GNAS complex locus
O43895	X-prolyl aminopeptidase (aminopeptidase P) 2, membrane-bound
P23526	Adenosylhomocysteinase
P02765	alpha-2-HS-glycoprotein
P01023	alpha-2-macroglobulin
P12821	angiotensin I converting enzyme (peptidyl-dipeptidase A) 1
Q9BYF1	angiotensin I converting enzyme (peptidyl-dipeptidase A) 2
P01019	angiotensinogen (serpin peptidase inhibitor, clade A, member 8)
P08758	annexin A5
P02647	apolipoprotein A-I
P06727	apolipoprotein A-IV
P04114	apolipoprotein B (including Ag(x) antigen)
P02749	apolipoprotein H (beta-2-glycoprotein I)
Q93088	betaine-homocysteine methyltransferase
Q96KN2	carnosine dipeptidase 1 (metallopeptidase M20 family)
P53634	cathepsin C
P48061	chemokine (C-X-C motif) ligand 12 (stromal cell-derived factor 1)
P11597	cholesteryl ester transfer protein, plasma
P10909	Clusterin
P00734	coagulation factor II (thrombin)
P03951	coagulation factor XI
P0C0L4	complement component 4A (Rodgers blood group)
P0C0L5	complement component 4B (Chido blood group)
P08603	complement factor H
P01034	cystatin C
P24855	deoxyribonuclease I
O95865	dimethylarginine dimethylaminohydrolase 2
P01133	epidermal growth factor (beta-urogastrone)
P02671	fibrinogen alpha chain
P02675	fibrinogen beta chain
P02679	fibrinogen gamma chain
P02751	fibronectin 1
P06280	galactosidase, alpha
Q16769	glutaminyl-peptide cyclotransferase
P08263	glutathione S-transferase alpha 1
P09211	glutathione S-transferase pi 1

P22352	glutathione peroxidase 3 (plasma)
P00738	haptoglobin-related protein; haptoglobin
P08107	heat shock 70kDa protein 1A; heat shock 70kDa protein 1B
P68871	hemoglobin, beta
P98160	heparan sulfate proteoglycan 2
P02649	hypothetical LOC100129500; apolipoprotein E
P06870	kallikrein 1
P18428	lipopolysaccharide binding protein
P05164	Myeloperoxidase
P27169	paraoxonase 1
P00747	Plasminogen
P00749	plasminogen activator, urokinase
P41222	prostaglandin D2 synthase, hematopoietic; prostaglandin D2 synthase 21kDa (brain)
P22891	protein Z, vitamin K-dependent plasma glycoprotein
Q9HD89	Resistin
P34896	serine hydroxymethyltransferase 1 (soluble)
P01009	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1
P01011	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3
P01008	serpin peptidase inhibitor, clade C (antithrombin), member 1
Q14393	similar to growth arrest-specific 6; growth arrest-specific 6
P55017	solute carrier family 12 (sodium/chloride transporters), member 3
Q13621	solute carrier family 12 (sodium/potassium/chloride transporters), member 1
P08294	superoxide dismutase 3, extracellular
P07996	thrombospondin 1
P02787	Transferring
P02766	Transthyretin
P07911	Uromodulin

Supplementary table S3.4: The following genes were annotated and classified as having genetic association to renal diseases using DAVID Bioinformatics resources 6.7, National Institute of Allergy and Infectious Diseases (NIAID), NIH. Uniprot Ids and gene names are given.

ID	Gene Name
P08183	ATP-binding cassette, sub-family B (MDR/TAP), member 1
P08571	CD14 molecule
P12821	angiotensin I converting enzyme (peptidyl-dipeptidase A) 1
Q9BYF1	angiotensin I converting enzyme (peptidyl-dipeptidase A) 2
P01019	angiotensinogen (serpin peptidase inhibitor, clade A, member 8)
P04114	apolipoprotein B (including Ag(x) antigen)
Q96KN2	carnosine dipeptidase 1 (metallopeptidase M20 family)
P11597	cholesteryl ester transfer protein, plasma
P08603	complement factor H

P02675	fibrinogen beta chain
P00738	haptoglobin-related protein; haptoglobin
P08107	heat shock 70kDa protein 1A; heat shock 70kDa protein 1B
P11142	heat shock 70kDa protein 8
P02649	hypothetical LOC100129500; apolipoprotein E
P01042	kininogen 1
P02788	Lactotransferrin
P05164	Myeloperoxidase
P27169	paraoxonase 1
P01833	polymeric immunoglobulin receptor
P01009	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1
Q99519	sialidase 1 (lysosomal sialidase)
P55017	solute carrier family 12 (sodium/chloride transporters), member 3
P08294	superoxide dismutase 3, extracellular

Supplementary table S3.5: Complete non-redundant list of proteins identified in the low speed pellet (P18,000g) using both DTT and CHAPS treatment methods. Uniprot accessions and gene names are given.

UNIPROT_ACCESSION	Gene name
Q6UX06	olfactomedin 4
P63104	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide
O76031	ClpX caseinolytic peptidase X homolog (E. coli)
P04114	apolipoprotein B (including Ag(x) antigen)
P04279	semenogelin I
P05109	S100 calcium binding protein A8
Q86YZ3	Hornerin
O00560	syndecan binding protein (syntenin)
P23526	Adenosylhomocysteinase
O95837	guanine nucleotide binding protein (G protein), alpha 14
P02747	complement component 1, q subcomponent, C chain
P06702	S100 calcium binding protein A9
P15941	mucin 1, cell surface associated
P63261	actin, gamma 1
P01042	kininogen 1
Q14624	inter-alpha (globulin) inhibitor H4 (plasma Kallikrein-sensitive glycoprotein)
P62879	guanine nucleotide binding protein (G protein), beta polypeptide 2
P53675	clathrin, heavy chain-like 1
P13929	enolase 3 (beta, muscle)
P04196	histidine-rich glycoprotein
P05090	apolipoprotein D
P01011	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3
P00734	coagulation factor II (thrombin)

Q7Z5L0	vitelline membrane outer layer 1 homolog (chicken)
P16278	galactosidase, beta 1
Q9C0H2	tweety homolog 3 (Drosophila)
P02760	alpha-1-microglobulin/bikunin precursor
P50148	guanine nucleotide binding protein (G protein), q polypeptide
P29972	aquaporin 1 (Colton blood group)
P02652	apolipoprotein A-II
P02679	fibrinogen gamma chain
P04004	Vitronectin
P36543	ATPase, H+ transporting, lysosomal 31kDa, V1 subunit E1
P01009	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1
P21796	voltage-dependent anion channel 1; similar to voltage-dependent anion channel 1
P50995	annexin A11
P11678	eosinophil peroxidise
P08133	annexin A6
P21266	glutathione S-transferase mu 3 (brain)
O43490	prominin 1
Q53GD3	solute carrier family 44, member 4
P68871	hemoglobin, beta
Q8N2U0	chromosome 17 open reading frame 61
Q9HCY8	S100 calcium binding protein A14
P31949	S100 calcium binding protein A11; S100 calcium binding protein A11 pseudogene
O00299	chloride intracellular channel 1
P08779	keratin 16; keratin type 16-like
P01593	similar to hCG1642538
P62937	similar to TRIMCyp; peptidylprolyl isomerase A (cyclophilin A); peptidylprolyl isomerase A (cyclophilin A)-like 3
O75594	peptidoglycan recognition protein 1
Q9NSB4	keratin 82
P07195	lactate dehydrogenase B
Q8N474	secreted frizzled-related protein 1
P02768	Albumin
Q9HB40	serine carboxypeptidase 1
Q9BQE3	tubulin, alpha 1c
Q8WUM4	programmed cell death 6 interacting protein
P07911	Uromodulin
O43895	X-prolyl aminopeptidase (aminopeptidase P) 2, membrane-bound
P15144	alanyl (membrane) aminopeptidase
P01834	similar to hCG26659; immunoglobulin kappa constant; similar to Ig kappa chain V-I region HK102 precursor
P35241	Radixin
P13987	CD59 molecule, complement regulatory protein
P39059	collagen, type XV, alpha 1
Q6EMK4	Vasorin
P55017	solute carrier family 12 (sodium/chloride transporters), member 3
P04083	annexin A1

P13646	keratin 13
Q5D862	filaggrin family member 2
Q9Y277	voltage-dependent anion channel 3
Q13510	N-acylsphingosine amidohydrolase (acid ceramidase) 1
P16444	dipeptidase 1 (renal)
P13647	keratin 5
P02647	apolipoprotein A-I
O00159	myosin IC
P07355	annexin A2 pseudogene 3; annexin A2; annexin A2 pseudogene 1
Q14390	gamma-glutamyltransferase light chain 2
P02743	amyloid P component, serum
Q08380	lectin, galactoside-binding, soluble, 3 binding protein
P04003	complement component 4 binding protein, alpha
Q7Z794	keratin 77
Q9H1C7	chromosome 5 open reading frame 32
P01024	similar to Complement C3 precursor; complement component 3; hypothetical protein LOC100133511
P19012	keratin 15
P01876	immunoglobulin heavy constant alpha 1
P02649	hypothetical LOC100129500; apolipoprotein E
P00338	lactate dehydrogenase A
P08727	keratin 19
P02538	keratin 6A
P36873	protein phosphatase 1, catalytic subunit, gamma isoform
O60814	histone cluster 1, H2bk
P04350	tubulin, beta 4
O60635	tetraspanin 1
P54793	arylsulfatase F
Q9Y512	sorting and assembly machinery component 50 homolog (S. cerevisiae)
O43653	prostate stem cell antigen
P09211	glutathione S-transferase pi 1
P01591	immunoglobulin J polypeptide, linker protein for immunoglobulin alpha and mu polypeptides
P12277	creatine kinase, brain
Q96FQ6	S100 calcium binding protein A16
P08473	membrane metallo-endopeptidase
Q8WVN6	secreted and transmembrane 1
Q9H853	tubulin, alpha 4b (pseudogene)
P10909	Clusterin
P00558	phosphoglycerate kinase 1
P15309	acid phosphatase, prostate
P13645	keratin 10
P20073	annexin A7
P10153	ribonuclease, RNase A family, 2 (liver, eosinophil-derived neurotoxin)
Q6UWR7	ectonucleotide pyrophosphatase/phosphodiesterase 6
P31946	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, beta polypeptide

P61026	RAB10, member RAS oncogene family
P01857	immunoglobulin heavy constant gamma 1 (G1m marker); immunoglobulin heavy constant mu; immunoglobulin heavy variable 3-7; immunoglobulin heavy constant gamma 3 (G3m marker); immunoglobulin heavy variable 3-11 (gene/pseudogene); immunoglobulin heavy variable 4-31; immunoglobulin heavy locus
P62873	guanine nucleotide binding protein (G protein), beta polypeptide 1
P35908	keratin 2
P26038	Moesin
O96009	napsin A aspartic peptidase
P12109	collagen, type VI, alpha 1
P09525	annexin A4
Q92820	gamma-glutamyl hydrolase (conjugase, folylpolygammaglutamyl hydrolase)
P01133	epidermal growth factor (beta-urogastrone)
Q02383	semenogelin II
P00918	carbonic anhydrase II
P00748	coagulation factor XII (Hageman factor)
P12035	keratin 3
P04264	keratin 1
Q9HD89	Resistin
P02533	keratin 14
P02788	Lactotransferrin
P15924	Desmoplakin
P15311	hypothetical protein LOC100129652; ezrin
P60709	actin, beta
P48061	chemokine (C-X-C motif) ligand 12 (stromal cell-derived factor 1)
P02751	fibronectin 1
P35527	keratin 9
P02787	Transferring
P19013	keratin 4
Q6ZVX7	non-specific cytotoxic cell receptor protein 1 homolog (zebrafish)
P06727	apolipoprotein A-IV
P04259	keratin 6B
P54802	N-acetylglucosaminidase, alpha-
P04406	glyceraldehyde-3-phosphate dehydrogenase-like 6; hypothetical protein LOC100133042; glyceraldehyde-3-phosphate dehydrogenase
P00749	plasminogen activator, urokinase
P01833	polymeric immunoglobulin receptor
P08758	annexin A5
P05154	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 5
Q12931	TNF receptor-associated protein 1
P60953	cell division cycle 42 (GTP binding protein, 25kDa); cell division cycle 42 pseudogene 2
Q13621	solute carrier family 12 (sodium/potassium/chloride transporters), member 1
A6NIZ1	Ras-related protein Rap-1b-like protein
A6NGU5	Putative gamma-glutamyltranspeptidase 3
P01620	Ig kappa chain V-III region SIE
P0CG04	Ig lambda-1 chain C regions
P0CG05	Ig lambda-2 chain C regions

P04220	Ig mu heavy chain disease protein (BOT)

CHAPTER 4

A LIPID AFFINITY-BASED NOVEL METHOD FOR ISOLATION OF URINARY MEMBRANE VESICLES AND THEIR SUBSEQUENT CHARACTERIZATION

4.1 Introduction

Human urine contains various types of vesicles including exosomes (40-100nm) (Pisitkun, Shen & Knepper 2004) and bigger vesicles (66-187nm) which reportedly contain polycystin-1, aquaporin-2 and podocin (Hogan *et al.,* 2009) as well as podocalyxin-positive 125nm membrane particles (Hara *et al.,* 2010). These vesicles can potentially provide information about the pathophysiological state of the entire urinary tract. The traditional method for urinary exosome isolation was the differential centrifugation where a low speed centrifugation (17,000g) was adopted to remove whole cells, tubular casts and membrane fragments. This is followed by ultracentrifugation (200,000g) to pellet down urinary exosomes and other similar sized vesicles (Pisitkun, Shen & Knepper, 2004). However, the problem associated with this method is contamination of the 200,000g pellet with soluble or aggregated high-abundance proteins like Tamm-Horsfall glycoprotein (THP) and albumin. It was found that, entrapped in THP polymers, exosomes were also precipitating with low speed pellet (P17,000g) (Fernandez-Llama *et al.,* 2010). To remove contamination of aggregated proteins one solution has been proposed which includes treatment of low speed (P17,000g) and high speed pellets (P200,000g) with DTT followed by a second centrifugation at low and high speed as above. This procedure increases the exosomal yield but it creates a new problem as THP is now present in high amounts in the high speed pellet. This may interfere with the further analysis of exosomal pellet by masking low abundance proteins. Moreover, the random re-oxidation subsequent to DTT treatment would result in multiple proteins losing their activity (See Chapter 2). If an activity-based analysis is the ultimate aim of the membrane vesicle isolation, this loss in activity would render it unsuccessful.

Another method which can be employed for vesicle isolation is sucrose density gradient centrifugation which has been used mainly to isolate exosomes from immune cells and cell culture but also on urinary exosome separations (Welton *et al.,* 2010; Keller *et al.,* 2011). The

sucrose density gradient ultracentrifugation method is more sensitive and yields the purest form of exosomes with minimal contamination with high abundance proteins. It is recognised that both these ultracentrifugation-based methods are very labour intensive and require extensive instrumentation and long processing times making them unsuitable for clinical settings. The fact that these isolation methods are labour intensive involving manual techniques makes them impractical when a large number of samples have to be handled. Alternatives have been proposed including the use of a nanomembrane concentrator (pore size 13nm and MWCO 100kDa) to enrich for the exosomal fraction from urine (Cheruvanky *et al.,* 2007). Advantages of the method are the short processing time and the use of inexpensive instrumentation like a table-top centrifuge. The 100 kDa 'cut-off' would mean that high molecular weight proteins and protein complexes not part of exosomes would be enriched as well and would interfere with further analysis masking low abundance proteins in exosomes. Many proteins, like THP and albumin, remain in polymeric forms (Atmeh, Shabsoug 1997; Fernandez-Llama *et al.,* 2010) that would be enriched with exosomes as well. It was already shown that nanomembrane ultrafiltration enriched soluble proteins, like albumin and α-1-antitrypsin in large amounts along with exosomes, when applied on nephrotic urine (Rood *et al.,* 2010) which limited the detection of microvesicular proteins. This study also used an alternative method which showed that loading crude ultracentrifugation pellet onto a size exclusion chromatography columns yielded three fractions (HMW, LMW and >10kDa). The high molecular weight fraction (HMW) showed the presence of exosomal markers while the low molecular weight (LMW) fraction had only high abundance proteins (Rood *et al.,* 2010). The LMW fraction was present only in nephrotic urine and absent in normal urine suggesting that large amounts of high abundance proteins interfere with the isolation and subsequent analysis of exosomes in nephrotic urine. Another method has been proposed involving the commercial reagent ExoQuick-TC (System

biosciences, Mountain view, CA) (Alvarez *et al.*, 2012). This method is simple and easy to perform and involves precipitatation of the exosomes from minimally processed urine. However, it was found in the study that this reagent also precipitates the THP with the exosomal fraction making it no different from other methods.

Taken together, most of the methods used in various studies are either labour intensive, require sophisticated instrumentation or they result in high abundance protein contamination in the vesicular fraction. Most of the methods mentioned above rely on biophysical parameters to separate membrane vesicles from urine which are shared by some of the contaminants as well. Therefore, there is need and scope for development of membrane vesicle isolation methods from the urine which are simple, robust and can be applied to a large number of samples without the need for additional sophisticated instrumentation. We hypothesized that a method based on affinity or biomolecular recognition should lead to a higher purity vesicle-containing fraction. The obvious choice for affinity, namely using immuno-affinity chromatography requires antibodies which are often expensive and, due to sensitivity to harsh elution conditions, their re-use is problematic if not completely impossible. Here we have developed a lipid affinity-based method for isolation of urinary microvesicles. Based on available information we have selected five peptides which bind to various lipids (phosphatidylserine, phosphatidylcholine, cholesterol, galactosylceramide and sphingomyelin) and immobilized synthetic biotinylated peptides on streptavidin agarose which was applied to ‘cell-free’ urine to enrich for microvesicles from urine.

4.2 Material and methods

Urine collection has been described in chapter 2.

4.2.1 Peptide-affinity chromatography

Peptides were appropriately designed and thereafter obtained from Genscript USA Inc, NJ. Five peptides were selected based on their physicochemical properties (Table 4.1). Three of them showed strong affinity to phosphatidylserine (PS), cholesterol (CL) and sphingomyelin (SM) while two others bind phosphatidylcholine (PC). Two peptides binding PC were inserted in a single sequence separated by three glycine residues. This would be expected to increase avidity of the resulting peptides having two binding sites for PC.

Table 4.1: List of all five peptides employed in the study. Amino acids in red denote glycine which was employed as a spacer to reduce steric hindrance because of the proximity of the peptide with the agarose beads. N-terminal glycine was biotinylated for capturing the peptide on streptavidin-agarose which was subsequently employed for affinity chromatography to isolate membrane vesicles from urine.

Peptide	Sequence	Lipid binder	Comments	Reference
P1	Biotin-GGGGGFNFRLKAGQKIRFG	Phosphatidylserine (PS)	Naturally Found in PS-decarboxylase, protein kinase C	(Igarashi *et al.*, 1995)
P2	Biotin-GGGGGKRESGGGFREL	Phoshpatidylcholine	Synthetic peptides; Can bind HDL	(Navab *et al.*, 2005)
P3	Biotin-GGGGATVLNYYVWRDNS	Cholesterol	Naturally found in Peripheral-type benzodiazepine receptor	(Li *et al.*, 2001)
P4	Biotin-GGGGKQHTVTTTTKGENFTETDVKMMER	Galactosylceramide, sphingomyelin	Naturally found in Prion protein, can potentially bind Lipid rafts	(Mahfoud *et al.*, 2002)

Four or five glycines were added at the N-terminal and the first glycine in all the peptides was biotinylated. This biotin was used for immobilising the peptides on streptavidin-agarose (Genscript). 1mL of streptavidin-agarose resin was mixed with 3mg of the respective peptides (P1 and P2 dissolved in water while P3 and P4 dissolved in DMSO) diluted in phosphate buffer saline (PBS, pH 7.4) and left on a rotating mixer overnight. On the following day the resin was washed with PBS to remove non-bound peptide. The P200,000g fraction obtained from the traditional ultracentrifugation method was incubated with P1-agarose overnight. The next day, after washing with 10mM phosphate buffer (20 column volumes) the bound vesicles were eluted with 100mM glycine buffer (pH 2.4). For affinity chromatography on whole urine, 50mL urine (SN2000g) was concentrated to 2mL using a Vivaspin (MWCO 5 or 300kDa, Sartorius, Goettingen). This reduction in volume was performed to reduce urine to manageable volumes without reaching high ionic strength or changing pH which would happen if dialysis and vacuum concentration was used. This 2mL volume was diluted to 4mL with 20mM phosphate buffer without salt and incubated with peptide-agarose on rotation overnight at 4°C. The non-bound fraction from peptide-agarose was collected the following day and resin was washed with phosphate buffer (20 column volumes). For elution 100mM glycine buffer (pH 2.4) was incubated with resin for 2 hours at $+4^{o}$C while rotating. The eluate was dialysed in a membrane with 3.5 kDa 'cut-off' against water and concentrated in a speed vac concentrator. This concentrated eluate was used for SDS-PAGE and WB, as described in Chapter 2. It was found that reproducible results were obtained for these peptide columns up to 15-20 intermittent runs.

Reduction, alkylation, trypsin digestion and mass spectrometric analysis were performed as described in Chapter 3.

4.3 Results

4.3.1 Selection of peptides

A number of studies have reported the typical lipid composition of exosomes derived from various sources. Interestingly, the phospholipid composition of the reticulocyte-derived exosomes is quite similar to the plasma memebrane of the cell (Vidal *et al.*, 1989). Phosphatidylcholine (PC) is the major phospholipid followed by phosphatidylethanolamine (PE) and then phosphatidylserine (PS) and phosphatidylinositol (PI) while sphingomyelin (SM) is similar to combined PS+PI. In case of mast cell and dendritic cell-derived exosomes, SM is enriched in exosomes compared to parent cell (Laulagnier *et al.*, 2004). Cholesterol (CL) is much more enriched in MHC II-bearing exosomes from B lymphocytes compared to the parent cell (Wubbolts *et al.*, 2003) while in case of exosomes from erythrocytes and mast cells CL amount is highly similar to parent cell plasma membrane (Vidal *et al.*, 1989; Laulagnier *et al.*, 2004).

Here, we sought to develop a lipid affinity-based method for isolation of urinary membrane vesicles from whole urine. Notably, our approach appears theoretically important as short peptides displaying binding to any of the above-mentioned lipids would be suitable for the task to isolate vesicles. For this purpose, we did a thorough literature search for short peptides and their binding properties to various lipids and five peptides with high affinity to various lipids were selected. The peptides selected are presented in Table 4.1.

4.3.2 Peptide affinity chromatography

Five peptides which bind various lipids (Table 4.1) were selected and 4-5 glycines were added to their N-terminal domain as a spacer to reduce steric hindrance resulting from proximity to the bead attached. These peptides with N-terminal glycine biotinylated allowing immobilisation to streptavidin-agarose, were obtained by peptide synthesis as supplied

commercially. To confirm if this approach using a lipid-binding peptide will work to isolate membrane vesicles, we first incubated P200,000g with P1-agarose and eluted the bound fraction.

4.3.3 PS binding peptide binding to exosomal pellet from urine and MS analysis

A small fraction of exosomes and the majority of ectosomes expose phosphatidylserine on their surface. Therefore, a crude pellet of exosomes (P200) was applied to P1-agarose to isolate only this fraction. This fraction was reduced, alkylated and trypsin digested and all proteins identified using LC-MS/MS. This served as positive control that this peptide (P1) actually is able to bind to membrane vesicles.

A complete list of the proteins identified in P1-peptide enriched fraction of P200,000g is given in Appendix 4.A. A total of 295 proteins were identified unambiguously. These proteins were compared to previously published datasets of urinary exosomes and 'exosome-like' vesicles. A comparison of P1-peptide-enriched proteins with two previous urine exosomal studies combined together (Pistikun *et al.,*. and Gonzalez *et al.,* combined as Knepper) (Pisitkun, Shen & Knepper, 2004, Gonzales *et al.,* 2009) and one study on urinary 'exosome-like' vesicles (Hogan *et al.,*) (Hogan *et al.,* 2009) is shown in Figure 4.1.

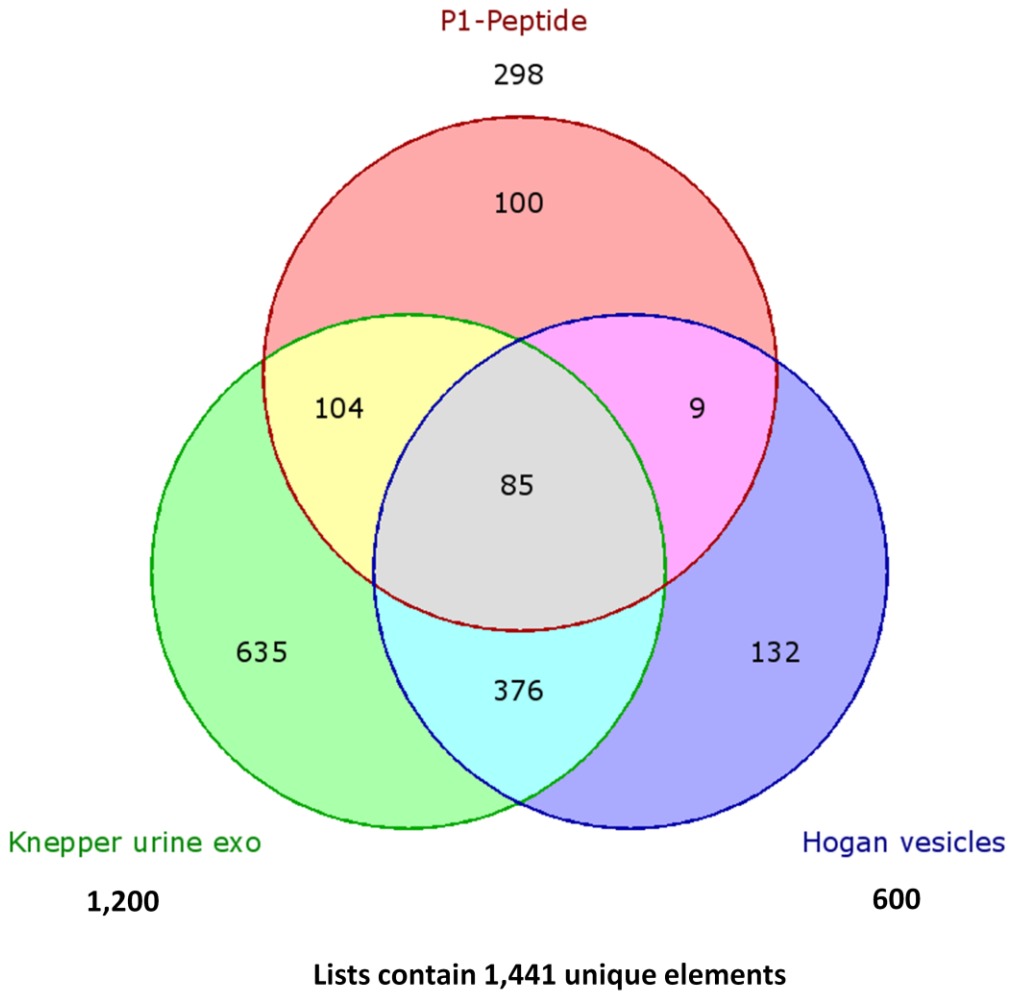

Figure 4.1: Comparison of P1-peptide enriched vesicle proteins with urinary exosomes (two studies combined as Knepper) (Pisitkun, Shen & Knepper, 2004) (Gonzales *et al.*, 2009) and exosome-like vesicles (Hogan) (Hogan *et al.*, 2009). For comparison, all the identifiers were converted to Unigene identifiers.

Eighty-five proteins (28% of our identifications) were found to be shared among all three studies while other than these, P1-peptide-enriched proteins are more in common with urinary exosomes (Knepper, 63% of P1 identifications common with this set) than with 'exosome-like' vesicles (Hogan, 31% of P1 proteins common with Hogan). However, very recently a study was published which has identified the highest number of proteins in urinary exosomes (3,280 proteins) to date (Wang *et al.*, 2011). This study has used the same isolation

protocol for urinary exosomes as that of previous studies (Pisitkun, Shen & Knepper, 2004, Gonzales *et al.*, 2009). Figure 4.2 shows the comparison of our protein list to the proteins they identified.

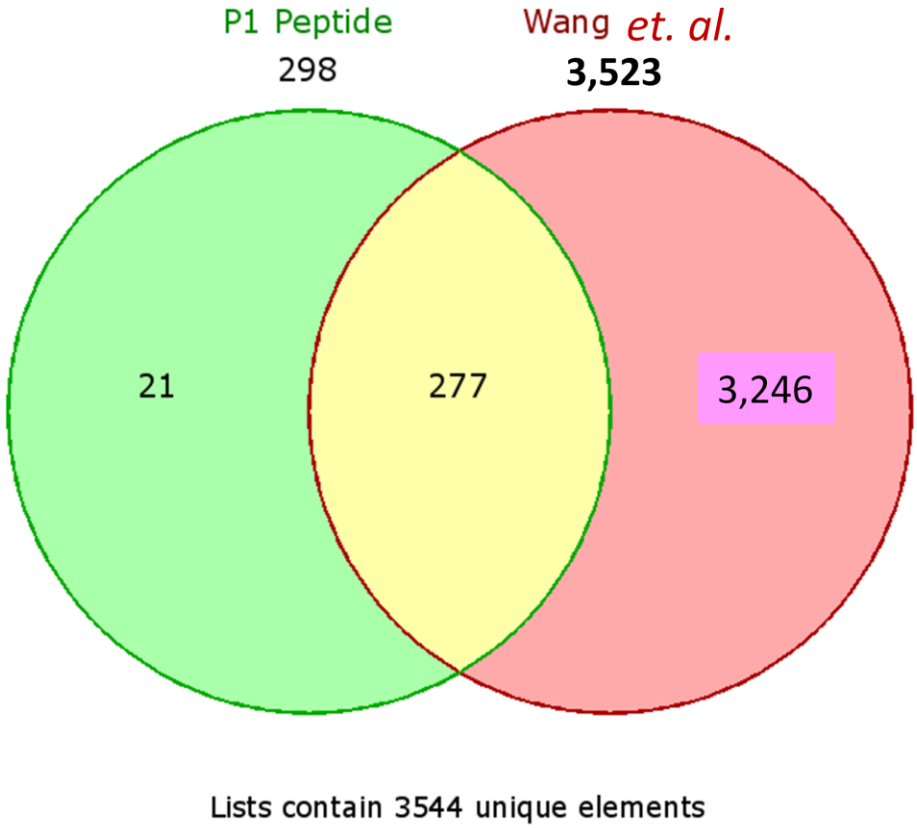

Figure 4.2: Comparison of P1-peptide-enriched vesicles with that of the largest protein identification study of urinary exosomes (*Wang et al.,*) (Wang *et al.*, 2011). Gene identifiers were converted to Unigene for comparison.

As can be seen in Figure 4.2, most of the proteins (93%) identified in our study are common to that of the largest previous study of urinary exosomes. This suggests that what we have enriched as vesicles is in common to the traditional protocol (ultracentrifugation-based isolation) used for urinary exosomes. It is to be noted that the differential centrifugation method will isolate a heterogenous mixture of different types of vesicles present in urine. Another study has previously been published which used complement receptor-1 (CR-1)

antibodies to enrich membrane vesicles from the urine of various patients having different types of kidney diseases (Lescuyer *et al.*, 2008). CR-1 is an ectosomal marker which is known to be expressed by the glomerular visceral epithelial cells, podocytes, in the kidney (Pascual *et al.*, 1994). When anti-CR-1 antibody-mediated enrichment of vesicles was used, 76 proteins from different patient samples were identified. To establish whether this is a different population of vesicles than our P1-peptide-enriched vesicles (other urinary exosome studies), we compared these proteins with that of largest urinary exosome proteome study and our protein list (Figure 4.3).

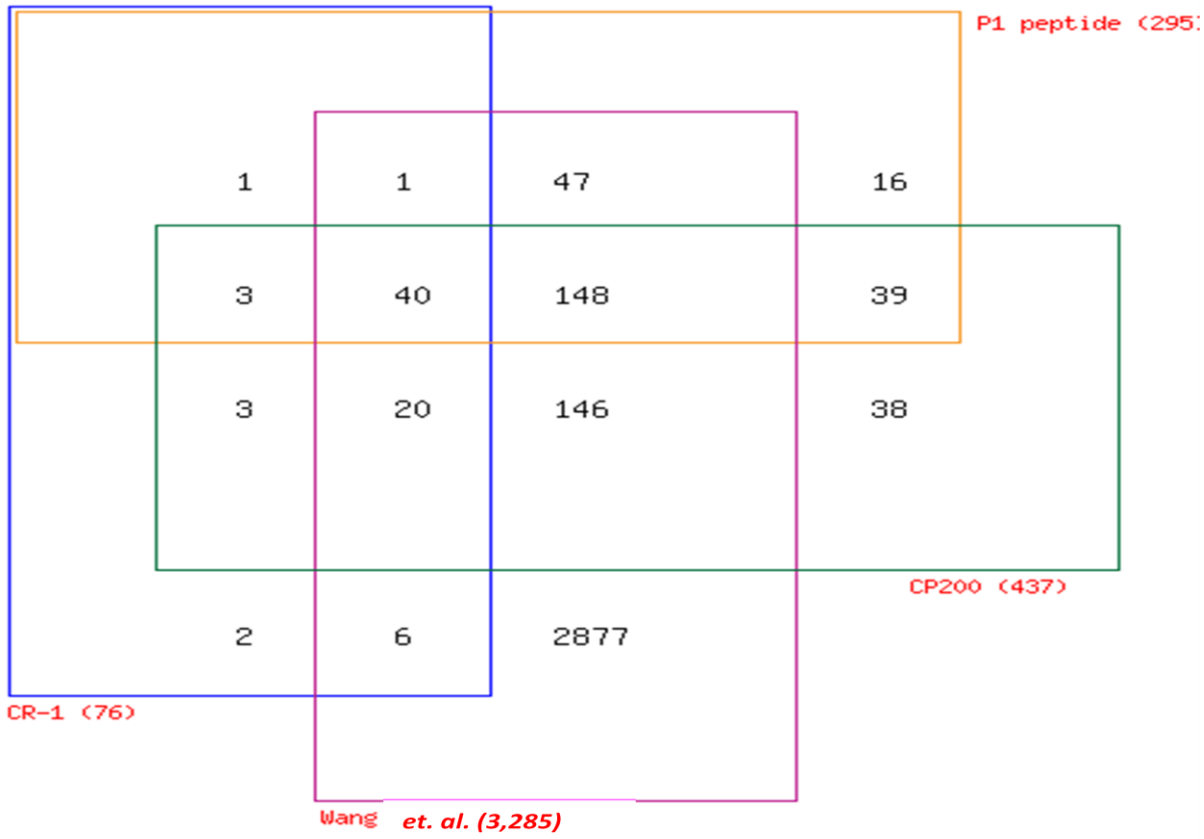

Figure 4.3: Comparison of anti-CR-1 immunopurified vesicles (Lescuyer *et al.*, 2008) proteome with that of largest urinary exosome proteome (Wang *et al.*,) (Wang *et al.*, 2011), our protein list (P1 peptide) and whole proteome identifications from Chapter 3 (CP200).

It can be seen in Figure 4.3 that all the proteins from CR-1 study, bar 9, have been identified in urinary exosome previously and a large proportion (60%) were found in our study (P1-peptide) as well. Another 30% were identified in the whole proteome study reported in Chapter 3.

Our protein list was annotated using IPA software (Redwood city, CA) and Blast2Go software (Conesa *et al.*, 2005) and graphics were generated for cellular component-based, molecular function and biological process-based categories. Figure 4.4 shows the annotation of our protein list according to the cellular component group.

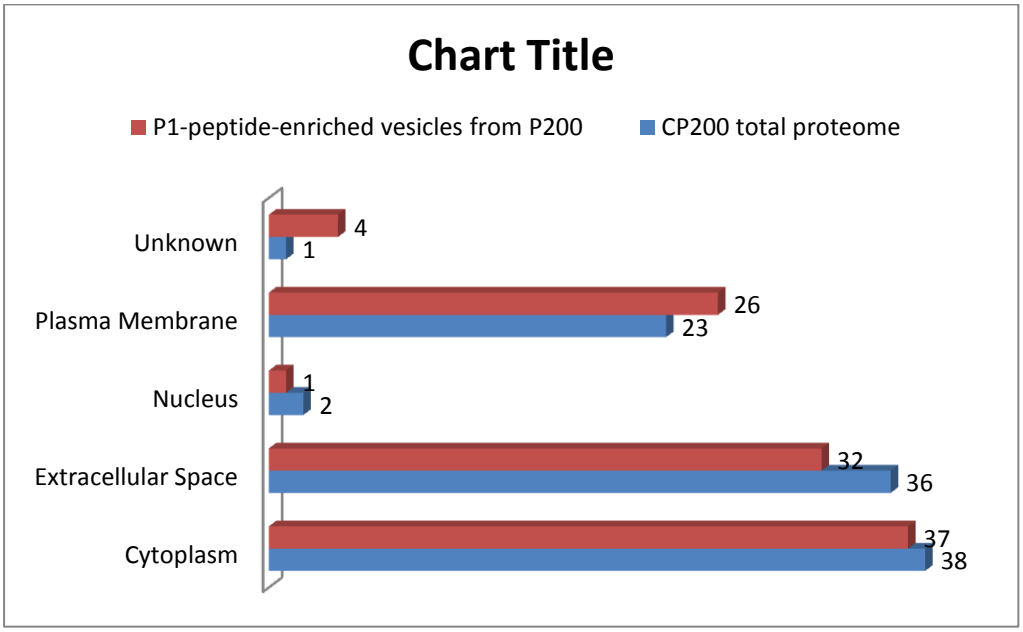

Figure 4.4: Annotation of our protein list according to the cellular component to which they belong, using IPA software. For comparison, the total proteome list from Chapter 3 is also shown (CP200).

Thirty-two% proteins belong to extracellular space while 37% were identified as cytosolic proteins. There are a large proportion of membrane proteins with 26% proteins belonging to membrane fraction. Only 1% proteins belong to nucleus. It can be seen in the Figure that by comparison all categories are similarly represented in total proteome of P200,000g from

Chapter 3. This suggests that what we have enriched using P1-peptides are similar to vesicles present in P200,000g isolated using differential centrifugation. The annotation of the protein list according to molecular function is shown in Figure 4.5.

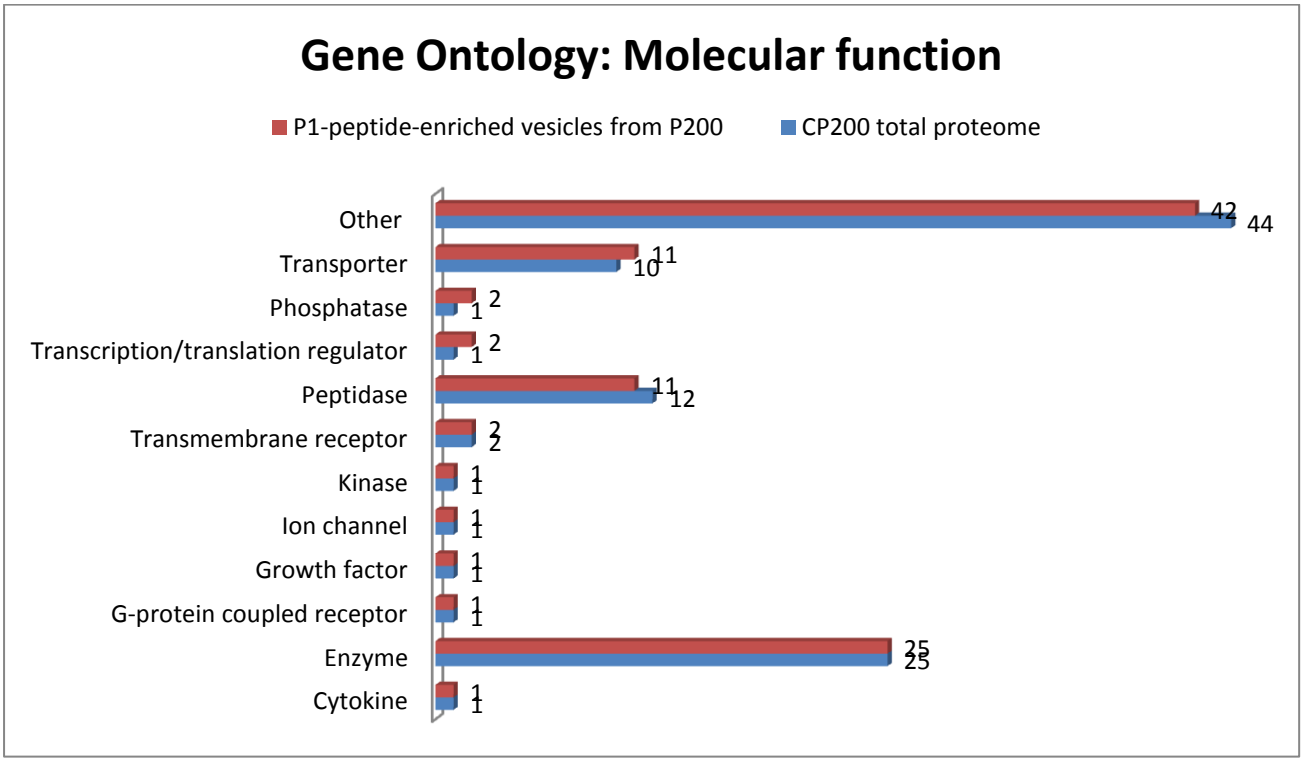

Figure 4.5: Annotation of our protein list according to their molecular functions using Blast2Go.

Forty-two percent proteins were annotated as 'other', which is comparable to those found in total proteome from chapter 3. Enzymes comprise 25 % of the proteins while peptidases and transporters make 11 % each. One % each of growth factors, G-protein coupled receptors, kinases and ion channels were found. Two % each of transmembrane receptor, transcription/translation regulators and phosphatases were found as well. Again, all the categories are similarly represented. The G-protein coupled receptor (GPCRC5C) was found in both the CP200 from Chapter 3 and P1-peptide-enriched vesicles here. However, smoothened, frizzled family receptor and frizzled family receptor 4 were unique to CP200

and P1-peptide, respectively. Similarly, in kinases, phosphoglycerate kinase 1 was common to both datasets. However, creatine kinase (brain) and phosphofrucokinase (liver) were unique to CP200 and P1-peptide, respectively.

Figure 4.6 shows the annotation of our protein list using Blast2Go software (Conesa *et al.*, 2005) according to biological processes in which they are involved. Seventy transport proteins were identified in our list followed by 64 proteins involved in regulation of biological quality and 62 involved in cellular component organization. Forty-two proteins are involved in the defense response while 61 proteins take part in immune response.

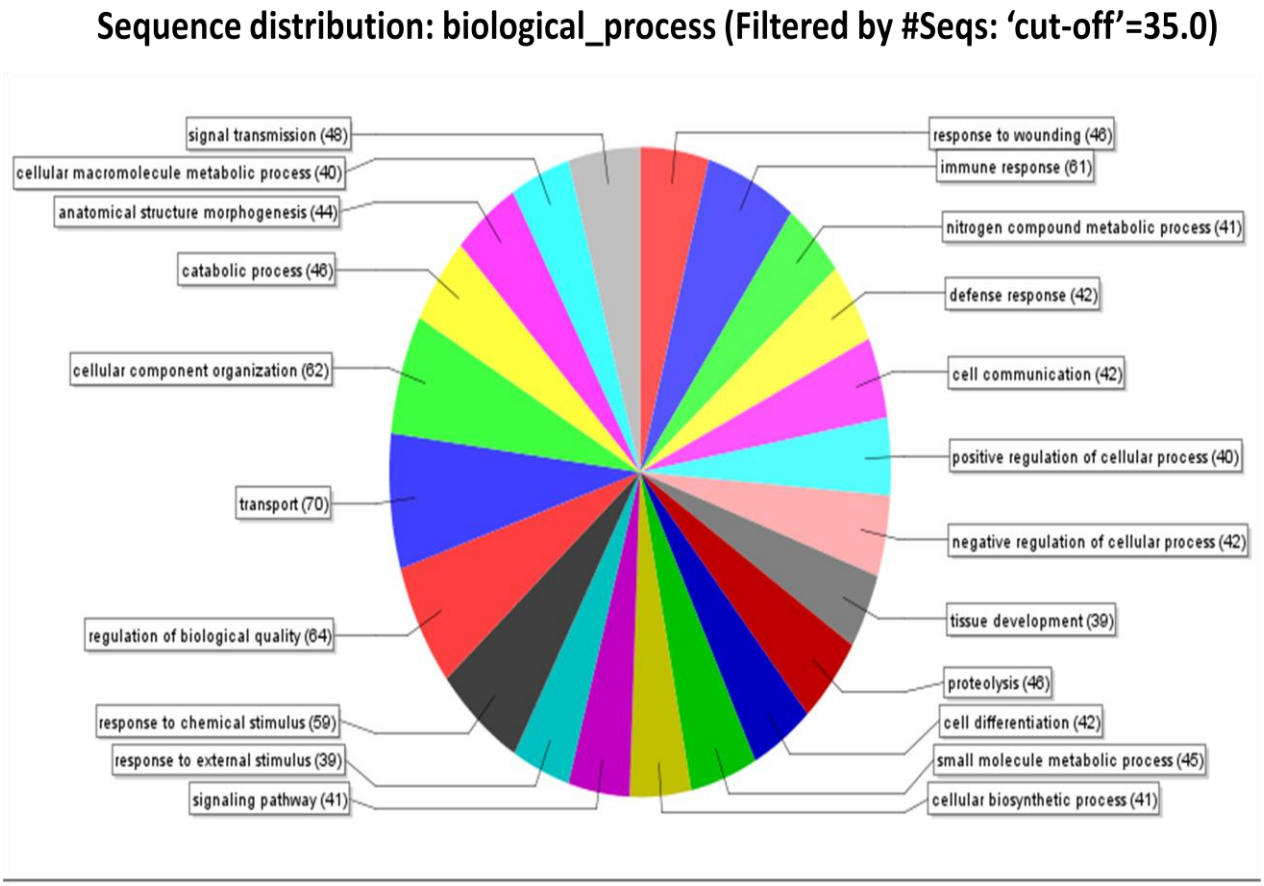

Figure 4.6: Annotation of our protein list according to biological processes in which these proteins are involved using Blast2Go software (Conesa *et al.*, 2005).

Forty-six proteins involved in proteolysis while 40 and 42 proteins involved in positive and negative regulation of cellular processes respectively were found in our list. Notably, 41 proteins involved in signalling pathways and 48 proteins involved in signal transmission were also identified.

4.3.4 Isolation of membrane vesicles from minimally processed urine

After confirming that our peptide-agarose matrix is able to bind membrane vesicles as evidenced by proteins identified in the elution of P1-agarose from P200,000g, all four peptide-agarose columns were applied to purify membrane vesicles from whole urine cleared of cells (Supernatant 2000xg). Phosphate buffer (10mM, pH 7.2) was chosen as binding buffer and concentrated urine (50mL was reduced to 2mL by Vivaspin MWCO 300kDa; 5kDa filter also works fine) was incubated with peptide-agarose overnight at rotation at +4^{o}C. After washing with binding buffer, elution was performed with reduced pH (100mM glycine, pH 2.4). The flow chart of the whole method is shown in Figure 4.7.

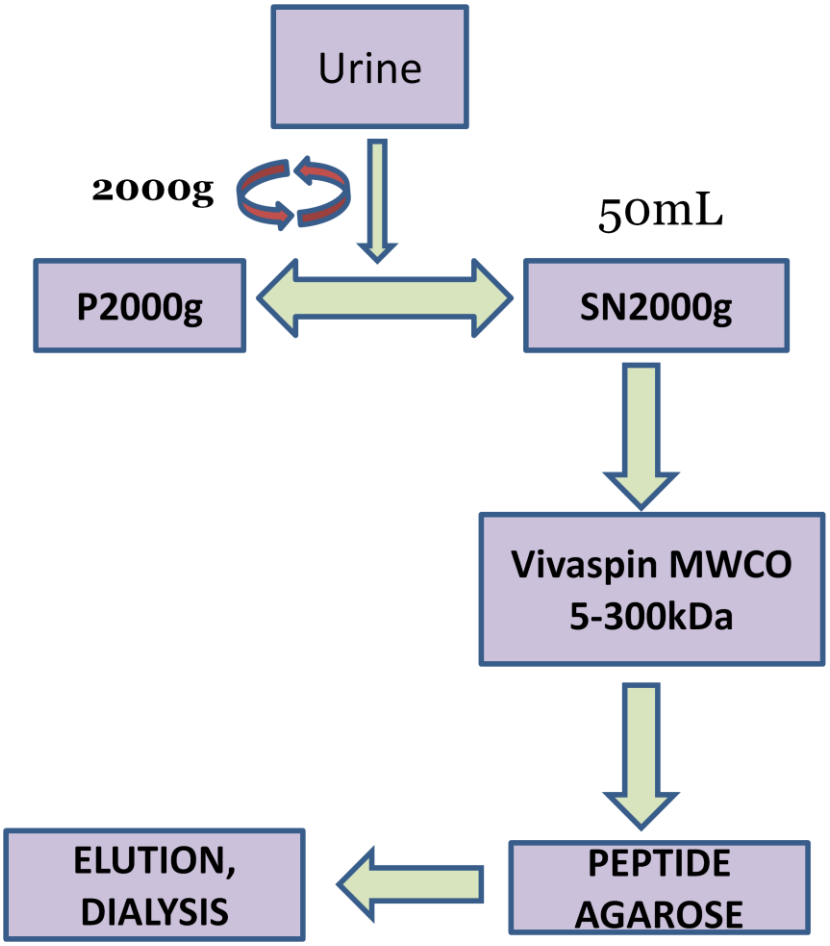

Figure 4.7: The workflow for peptide-affinity chromatography.

The same chromatography was performed on all four peptides separately. In another experiment, the non-bound fraction of P1-agarose was applied to P2-agarose to recover vesicles which are left in the urine. The obtained elution was dialysed to remove glycine and SDS-PAGE was performed after reducing the volume by speed vacuum centrifugation.

Figure 4.8 shows the SDS-PAGE of elution of all four peptides and P2 elution of P1 non-bound.

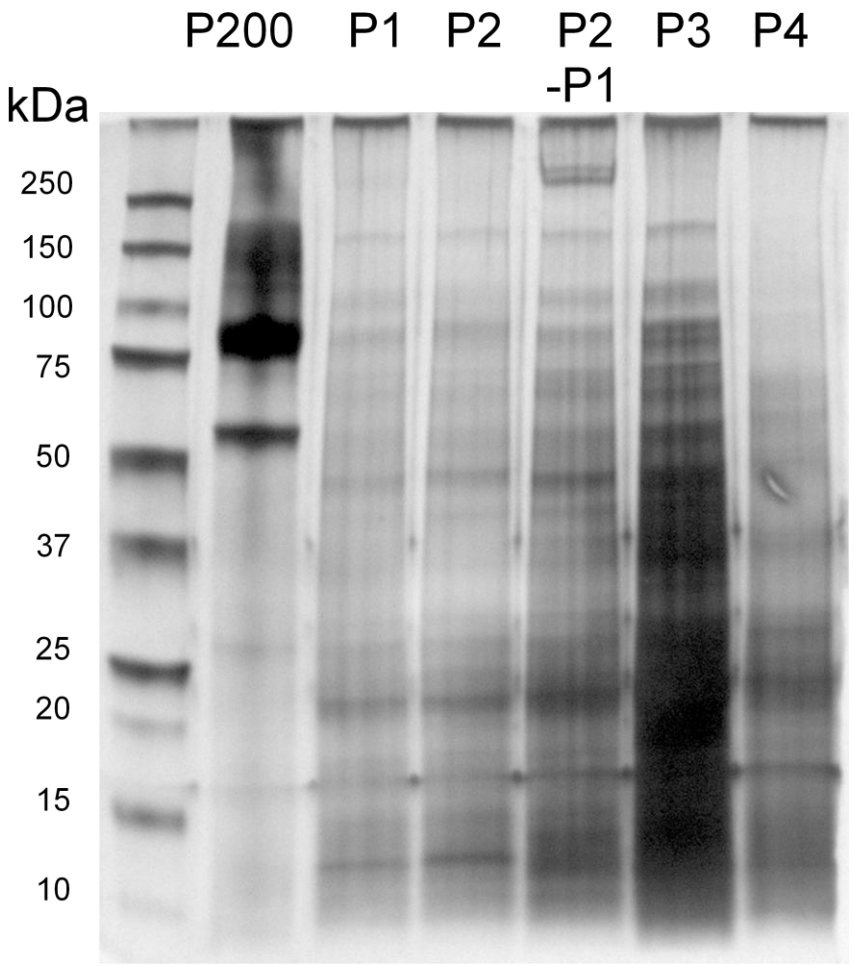

Figure 4.8: Shows the silver stained SDS-PAGE of elutions of all four peptide-agarose columns (P1-P4) in non-reducing conditions. P2-P1 means that non-bound of P1 was applied to P2-agarose. P200 is the crude exosomal fraction obtained using the traditional differential centrifugation method (Pisitkun, Shen & Knepper, 2004). P1-P4 her means the vesicles eluted from these peptides were loaded onto the gel.

It can be seen in Figure 4.8 that all four peptides enrich proteins which resemble a complex pattern from low to high molecular weight suggesting presence of membrane vesicles. Whole urine also has proteins in the same range as these peptide elution fraction but THP and albumin are more abundant in the urine and break the pattern of multiple proteins present in similar intensity. This is not the case here suggesting peptide-eluted fraction is a sub-fraction of urine. Approximately, the same amount of protein (3µg by the Bradford assay) was loaded

onto every lane. It can be clearly seen that in the crude exosomal fraction, obtained by traditional differential centrifugation method, the major proteins are Tamm-Horsfall glycoprotein (THP) at ~80kDa and albumin at ~65kDa. All other proteins are not clearly resolved and are present in low amounts. This is not the case with all four peptide-agarose elutions which show a complex pattern with least interference of THP and albumin. P2-P1 shows two bands above 250kDa which are not present in other peptide elutions even in P2-agarose elutions when applied to whole urine. Apart from this the, major pattern of proteins is similar in all four peptide elutions although difference can be found such as presence of several proteins which can be seen to be less or more enriched among four peptides.

4.3.4.1 Western blotting with exosomal markers and abundant proteins

All four peptide-agarose elutions were immunoblotted for exosomal marker CD63 (Figures 4.9). CD63 (Figure 4.9) is present as 48kDa band in P1, P2 and P2-P1 elutions in high intensity band while less intense in P3 and P4. This 48kDa band appears to be enriched in peptide-agarose elutions compared to P200 (crude exosomal fraction).

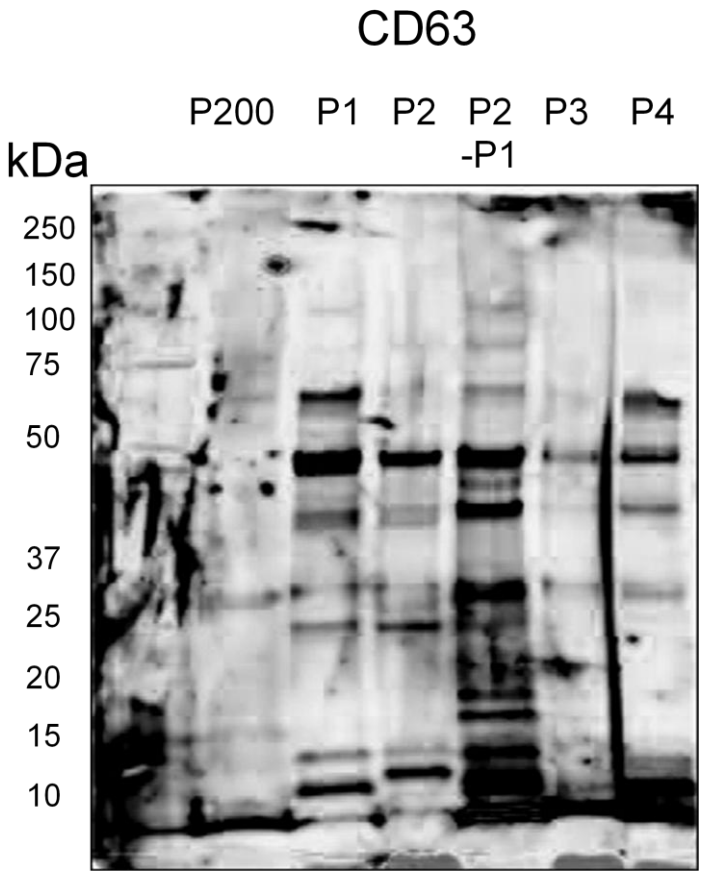

Figure 4.9: Western blot of all four peptide-agarose elutions using anti-CD63 antibodies in reducing conditions. P200 is the crude vesicular pellet obtained from urine using differential centrifugation method. P1 to P4 are elution fractions of peptide-agaroses when urine to applied to these peptide columns. P2-P1 means that non-bound fraction of P1-peptide-agarose was applied to P2 peptide-agarose and bound fraction of P2 was eluted.

There is another band at 65kDa which is most intense in P1 compared to P200 while less so in other peptide elutions. P1 and P2-P1 elutions have a band at 100kDa while another band at 75kDa is present in all elutions of P1 and P2 but not in P3 and P4. Looking at the overall pattern, P1 is similar to P2 and P3 is similar to P4. P2-P1 although a bit similar to P1 and P2, seems to be different from all other peptides. CD63 is a tetraspanin protein which is considered as a marker of exosomal membrane vesicles. The pattern found here for CD63 has some similarity with DTT, and CHAPS-treated P200,000g and P18,000g as reported in

Chapter 2 (Figure 2.4, Panel A). However, multiple low molecular weight bands are present here which were not visible in those fractions. In comparison with crude pellet P18,000g (Figure 2.4, Chapter 2) the band at 50 and 25 kDa is common however the band at ~65kDa found here in P1 and P2 peptide is not present in crude P18,000g.

Using the traditional method of isolation of membrane vesicles, a high interference of THP and albumin is observed in the exosomal fraction. We immuno-blotted all peptide eluates and non-bound fractions for THP and all eluates for albumin and observed the efficiency of peptide-affinity chromatography in enriching vesicles without interference by these proteins.

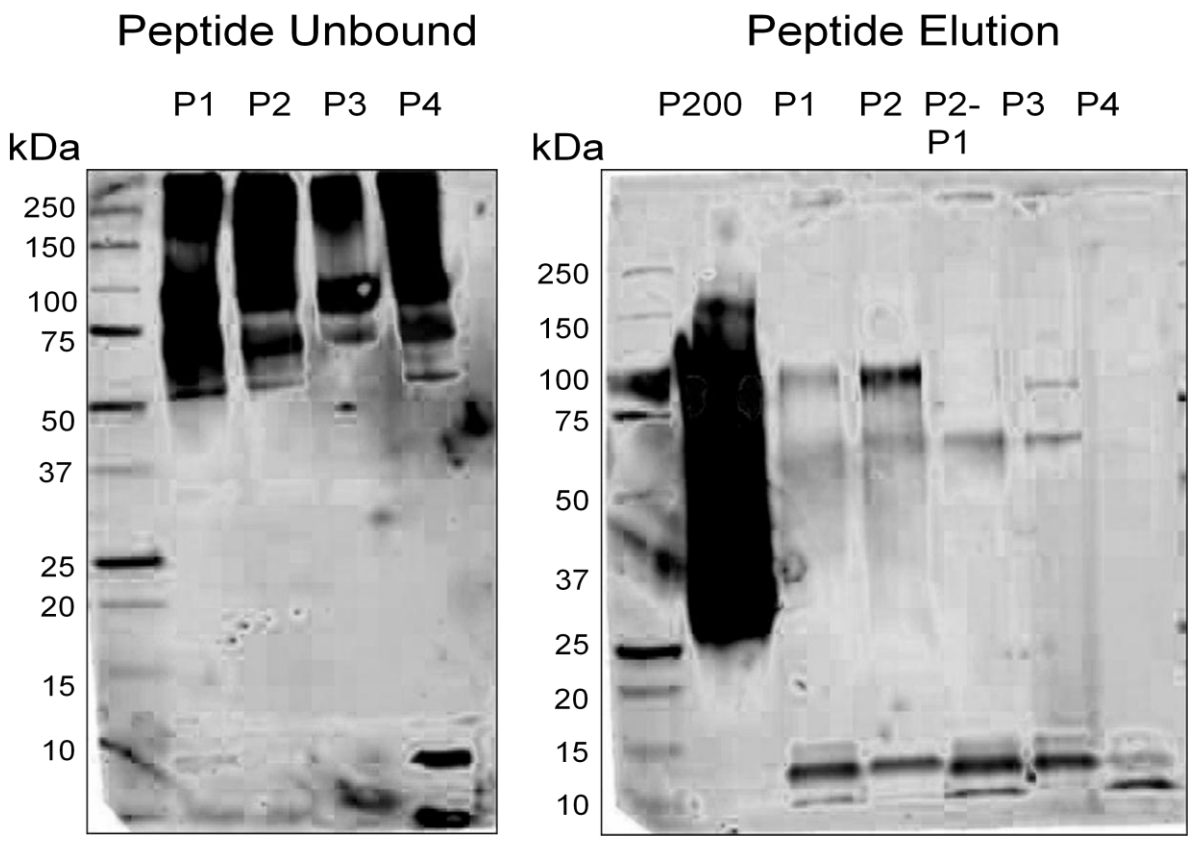

Figure 4.10: Western blot of peptide-agarose non-bound and elution using anti-THP antibodies. P200 is the crude vesicular pellet obtained from urine using differential centrifugation method. P1 to P4 are elution fractions of peptide-agaroses when urine to applied to these peptide columns. P2-P1 means that non-bound fraction of P1-peptide-agarose was applied to P2 peptide-agarose and bound fraction of P2 was eluted.

The non-bound fraction of peptide-agarose chromatography has significant amount of THP, as can be seen in Figure 4.10. However, in eluates, THP is present in negligible amounts when compared to crude P200 (exosomal fraction) obtained by traditional centrifugation although very small fragments (~12kDa) can be seen in all peptide elutions. P2 elution has more THP compared to P1 and P3 elutions but when P1 non-bound is applied to P2 THP is absent in the elution fraction. Peptide 4 elution also has almost undetectable levels of THP.

Figure 4.11 shows that full length albumin (~66kDa) is present in large amounts in crude exosomal pellets obtained by the traditional centrifugation method. Higher order aggregates of albumin and smaller fragments are also present in significant amounts in this pellet. However, all the peptide elutions are essentially free of fragments and aggregates of albumin while full-length albumin (whole protein ~66kDa) is present in a trace amount. Once again a similar pattern to THP is observed here with slightly more albumin (although non-significant amounts) being present in the P2 elution and when non-bound of P1 is applied to P2-agarose the eluate of P2 becomes free of it.

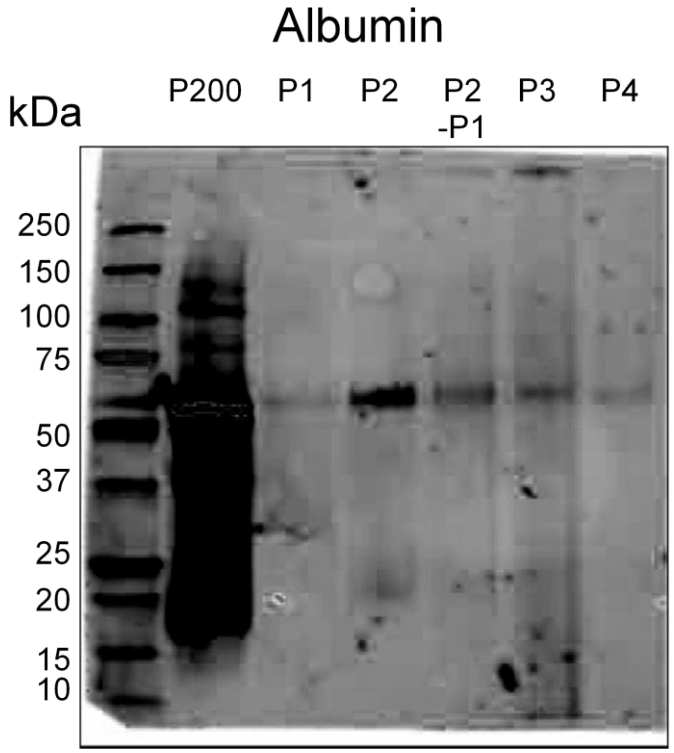

Figure 4.11: All peptide-agarose elutions were immunoblotted for albumin using a monoclonal anti-albumin antibody. P200 is the crude vesicular pellet obtained from urine using differential centrifugation method. P1 to P4 are elution fractions of peptide-agaroses when urine to applied to these peptide columns. P2-P1 means that non-bound fraction of P1-peptide-agarose was applied to P2 peptide-agarose and bound fraction of P2 was eluted.

Finally, to validate that the isolation of membrane vesicles is due to the lipid binding ability of the peptide and not due to the non-specific adsorption to stationary phase matrix, the whole procedure was repeated on streptavidin-agarose alone without the immobilised peptides. Figure 4.12 shows the SDS-PAGE of our negative control and the P1 peptide-agarose eluate.

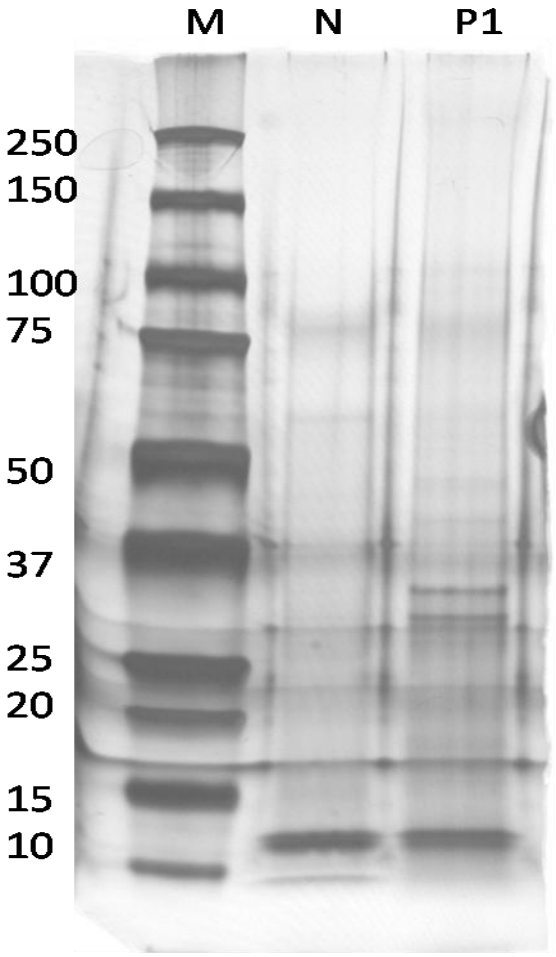

Figure 4.12: Silver stained SDS-PAGE of negative control (Streptavidin agarose) and P1-peptide eluate. N: streptavidin agarose (without any peptide immobilised) and P1: P1 peptide agarose.

There are three proteins which are common to the negative control and to P1-peptide agarose, as can be seen in Figure 4.12,. One is approximately 80kDa which could be THP and the other two being 60kDa and 12kDa proteins. These three proteins might be becoming enriched with elution non-specifically owing to their interaction with the stationary phase matrix or

streptavidin. The overall pattern of P1-peptide agarose is similar to one shown in Figure 4.8. To exclude binding of these peptides to any exosomal or membrane vesicle proteins, we probed the P200,000g, P18,000g and SN200,000g with these peptides by western blotting using a previously published protocol (Melrose, Ghosh & Patel, 1995). No bands were observed suggesting these peptides are enriching vesicles through lipid affinity only.

4.3.4.2 Transmission electron microscopy (TEM) of vesicles isolated from minimally processed urine using P1 and P2

Peptide P3 and P4 had lower levels of CD63. Therefore, they were excluded from further analysis. Elutions from P1 and P2 were further analysed by TEM to assess the morphology and size of the vesicles enriched. Figure 4.13 shows the representative vesicles enriched by peptide P1.

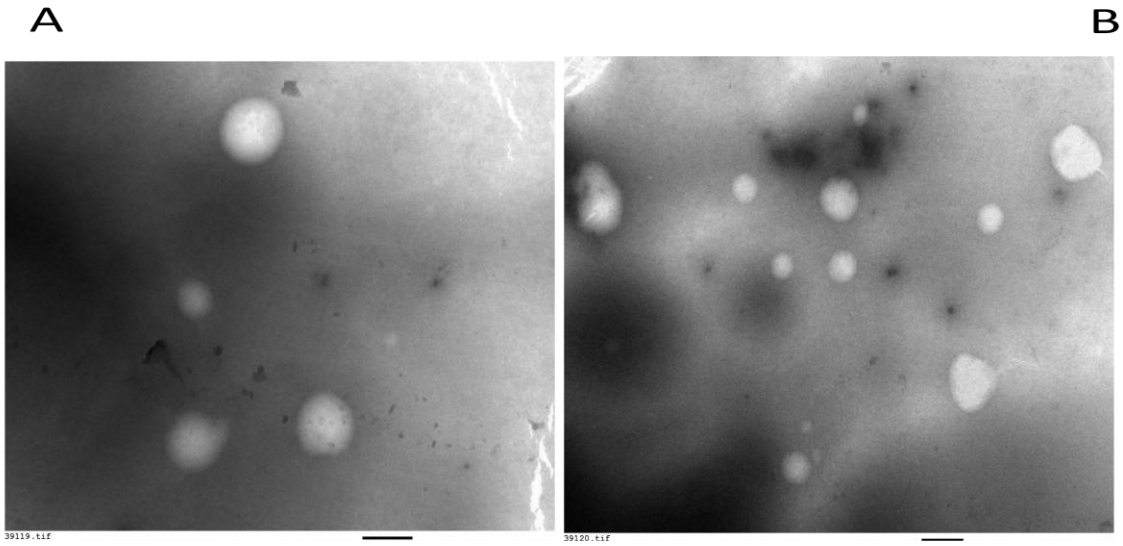

Figure 4.13: TEM analyses of vesicles enriched by PS binding peptide (P1). The scale bar in both the panels of figures is 500nm and direct magnification in panel A is 6000x while in panel B, it is 5000x. The bars shown below the picture are 500nm.

The figure 4.14 shows the representative TEM picture of vesicles enriched by PC binding peptide (P2).

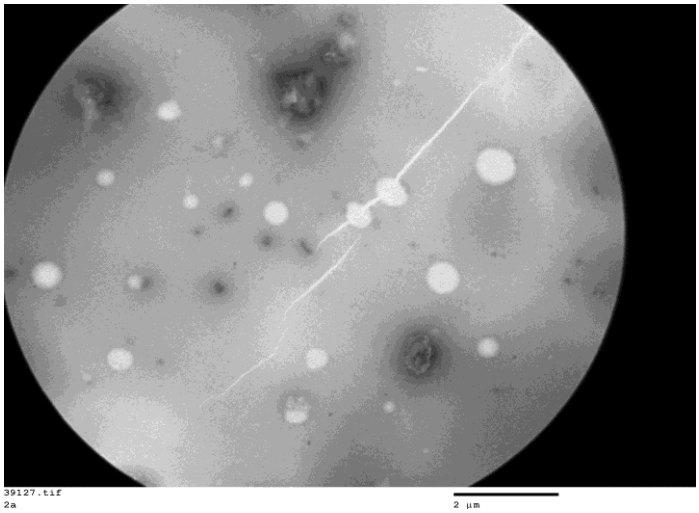

Figure 4.14: TEM analyses of vesicles enriched by PC binding peptide (P2). The scale bar shown below the picture in is 2000nm and direct magnification is 2500x.

Multiple pictures were obtained for both the peptides and size of all the vesicles calculated using ImageJ software. A graph was created for size distribution of vesicles enriched by both the peptides (Figure 4.15).

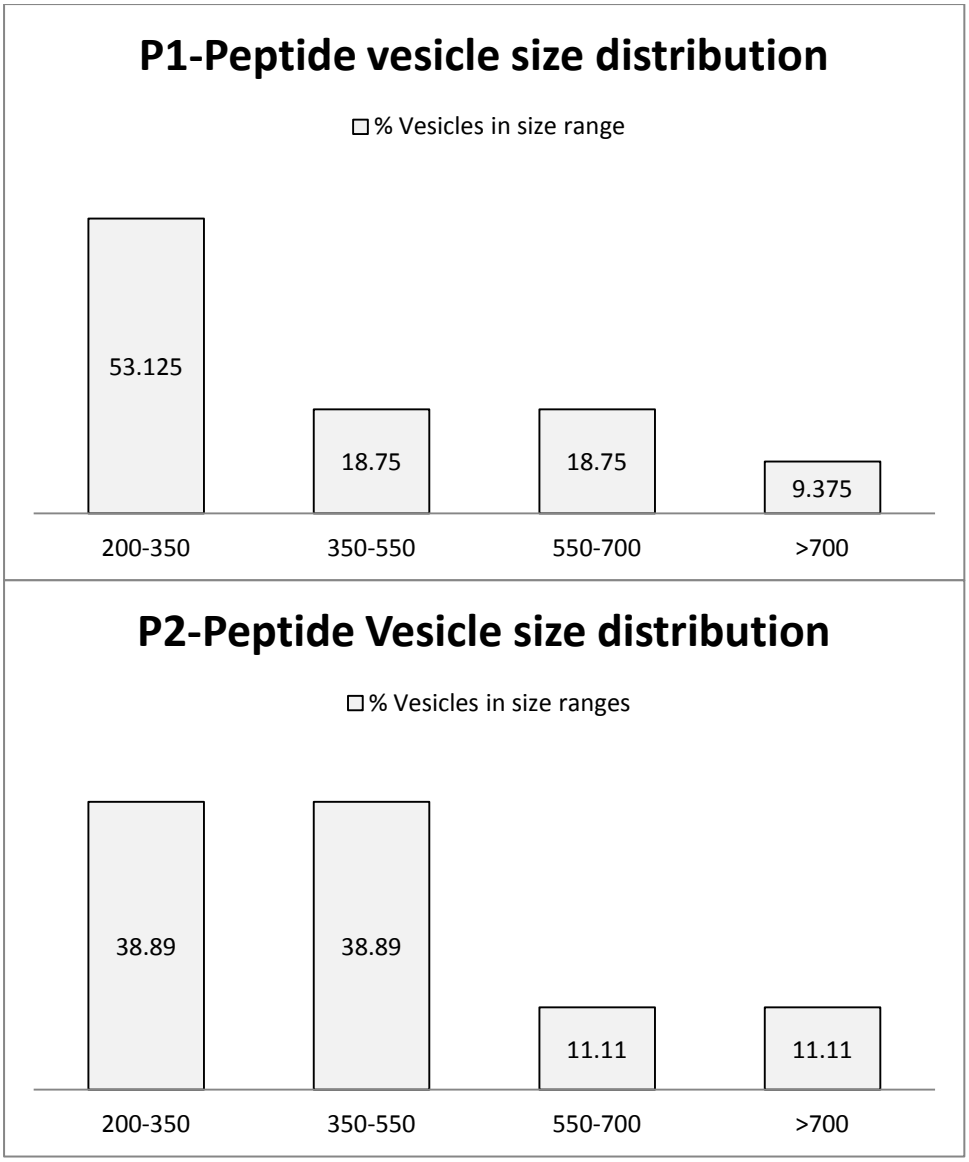

Figure 4.15: Size distribution of vesicles enriched by both the peptides (P1 and P2). X-axis is the size of vesicles in nanometers (nm) and Y-axis is the percent vesicles in any given size range.

The majority of the vesicles in the both the peptides were from 200-550nm (72% for P1 and 77% for P2) while lesser number of vesicles were found in 550-700nm range (18% for P1 and 11% for P2). There were approximately 10% of vesicles which were bigger than 700 and lesser than 1µm for both the peptides. This clearly shows that, although P1 can bind to vesicles in the exosomal fraction (section 4.3.3), both the peptides (P1 and P2) when applied

to minimally processed whole urine, enrich vesicles (200-1000nm) which have been given the name microvesicles or microparticles in the literature. The published procedure for vesicle counting from TEM pictures employs 20 image fields (Fernandez-Llama *et al.*, 2010). We have captured 7 fields for P1 peptide and 5 fields for P2 peptides. Therefore the size distribution presented in figure 4.15 has a limitation for extrapolation to the whole sample. This size distribution is crude and only provides a rough idea about the percentage of vesicles having a given size.

4.4 Discussion

A novel method to enrich membrane vesicles from urine was developed. Five peptides were designed based on the published literature, which show binding to PS, PC, CL and SM. Three of these peptides are naturally present in various proteins which bind the cognate lipids listed in Table1, while two peptides (KRES and FREL) are synthetic which have been shown to bind to PC (Navab *et al.,* 2005). These 2 PC binding peptides were joined together in a single sequence separated by 3 glycine residues while 5 glycines were used to separate them from the agarose bead. All of the other three peptides were similarly inserted in a sequence. phosphate buffer saline (PBS pH 7.4) was initially chosen as binding buffer because lipid-binding proteins (parent proteins of these peptides) bind the respective lipids intracellularly (pH 7.2) and the synthetic peptides has been shown to bind to PC at this pH (Navab *et al.,* 2005). However, urine has THP which precipitates at 150mM salt concentration (Kobayashi & Fukuoka, 2001), and therefore, NaCl was removed from the buffer. Finally only 10mM phosphate buffer (pH 7.2) was used as the binding buffer. Urine was centrifuged at 2000xg and the pellet containing cells was discarded because shed or dead cells present in urine would bind to these lipid-binding columns.

The traditional method of urinary exosome isolation was differential centrifugation (Pisitkun, Shen & Knepper, 2004; Gonzales *et al.,* 2009; Wang *et al.,* 2011). However, this method enriches for soluble proteins like THP and albumin in large quantities which are known to interfere with the subsequent analysis of vesicles and mask low abundance proteins. THP also causes a lot of membrane vesicles to pellet down at low speed by entrapping them in its fibers (Fernandez-Llama *et al.,* 2010). DTT treatment has been proposed to resolve this problem and it does increase the yield but causes a lot of THP to precipitate with the high speed exosomal fraction (P200,000g) instead of low speed (P18,000g) (Fernandez-Llama *et al.,* 2010). Other methods like density-gradient ultracentrifugation have less interference from

soluble protein but they are labour-intensive and not suitable for clinical applications. Our method, as shown in the results section, does not enrich for soluble proteins like THP and albumin while enriching CD63, which is a membrane vesicle marker. CD63, a tetraspanin, is traditionally an exosomal marker (exosomes are 40-100nm in size) but, as can be seen in the TEM picture of P1 and P2 peptide elution, only vesicles bigger than 200nm are present. These results question the specificity of CD63 as a stringent exosomal marker. CD63 has been shown to float in high density regions (1.26-1.29 g/ml) of a sucrose gradient. Although density floatation has not been determined for larger microvesicles or microparticles, a density of 1.26-1.29g/mL would be expected for vesicles much larger than exosomes. In another study, CD63 was detected in bigger microvesicles (100-1000nm) secreted by platelets although it was more enriched in exosomes. This evidence from other studies, as well as our study, present a picture in which CD63, although enriched in exosomal fractions from various sources, is also present on bigger microvesicles.

Only a fraction of exosomes, and most of the ectosomes and microvesicles (100-1000nm), are expected to expose PS while all vesicles would be expected to expose PC. We have fractionated the traditional high speed pellet (200,000g) known to isolate exosomes using a PS binding peptide. In the PS-exposing fraction of exosomal pellet enriched by P1-peptide agarose, CD63, CD82 and CD9 were identified but TSG101 and alix, which reflect the endosomal origin of the exosomes, were absent. CD133 or prominin-1, which is a marker of membrane particles (Marzesco *et al.,* 2005), was also identified in our PS-binding peptide-enriched fraction. Recently it was shown that hematopoietic stem cells release CD133-positive vesicles upon differentiation and CD63 inside the cell co-localises with CD133 (Bauer *et al.,* 2011). Therefore, distinct populations of membrane vesicles are all present in the high speed pellet from the traditional centrifugation method. This leads us to suggest that the high speed pellet generated using the traditional differential centrifugation method is a

heterogeneous mixture of different types of vesicles derived from different pathways of trafficking in cells. This view is also supported by our finding that, PS exposing vesicles, from the high speed pellet, have the majority of the proteins common to the largest urinary exosomes proteomic study (Wang *et al.,* 2011). The relevance of PS exposure only on a small fraction of exosomes is not clear; therefore a method to fractionate and purify only the PS-exposing exosomes from other populations of exosomes will help future functional studies on these vesicles. Moreover, MFG-E8 or lactadherin was identified to be present in crude as well as CHAPS and DTT-treated P18,000g and P200,000g pellets (Figure 2.4 Panel C, Chapter 2). This protein can bind integrins present on exosomes (Taylor *et al.,* 1997) as well as PS which might be present on surface of different types of vesicles. This protein could potentially cross-link various populations of vesicles and cause them to pellet down at a given centrifugal force (e.g. 18,000g or 200,000g).

The vivaspin filtration column was used in our protocol because it can reduce the volume of urine to a manageable volume and maintain the same ionic strength at the same time. Therefore vivaspin-concentrated samples will be free of high ionic strength (typically seen in other concentration methods such as vacuum concentration) which might interfere with peptide vesicle binding. When P1 and P2 peptides are employed to isolate membrane vesicles from minimally processed whole urine, these peptides enrich mostly bigger vesicles (200-1000nm). However, P1 peptide can bind to exosomal fractions enriching a sub-population of exosomes (Section 4.3.3) or 'exosome-like' vesicles. There could be two reasons for this discrepancy:

1. Microvesicles (100-1000nm) expose a greater amount of PS on their surface and, therefore, smaller vesicles are not enriched which expose only small amounts of PS. This was shown to be true in case for platelet microvesicles and exosomes where only 18% of exosomes expose PS on their surface while most microvesicles expose PS.

2. There are a much higher number of microvesicles present in urine and small sub-populations of exosomes exposing PS lose out in competition with microvesicles for binding to P1 peptide resin. Although crude information about numbers of smaller exosome and 'exosome-like' vesicles has been reported (Fernandez-Llama *et al.,* 2010), there is no data about numbers of bigger microvesicles in urine. Therefore, no comparisons can be made and this remains only a hypothesis.

There is no data for PC exposure on membrane vesicles in the literature and we propose the above same two arguments for explaining the lack of presence of smaller exosomes and 'exosome-like' vesicles in P2 elution fractions. Bigger microvesicles (100-1000nm) having procoagulant activity in the urine of rabbits has been described (Wiggins *et al.,* 1987). These vesicles are thought to bud from glomerular epithelial cells and are also visible in the proximal tubular lumen. The procoagulant activity was 'tissue factor/Factor VII-like' which was associated with microvesicles (100-1000nm). No report of such vesicles exists in human urine samples therefore we are the first to show the presence of microvesicles in urine which expose PS on their surface and might contain procoagulant activity.

This method is very simple and can be performed in any laboratory without the need for expensive instrumentation or trained professionals. In the future, an automated method for membrane vesicles isolation from urine using these peptides bound to stationary phase can be envisaged. Moreover, most of the proteins identified in urinary vesicles from patients of various types of nephropathies (CR-1) are common with our study (60% overlap between both the studies) and this suggests the clinical utility of our method of isolation of membrane vesicles.

4.5 Reference

Alvarez, M.L., Khosroheidari, M., Ravi, R.K. & DiStefano, J.K. (2012), "Comparison of protein, microRNA, and mRNA yields using different methods of urinary exosome isolation for the discovery of kidney disease biomarkers", *KIDNEY INTERNATIONAL*,vol. 82, no. 9, pp. 124-132 .

Atmeh, R.F. & Shabsoug, B. (1997), "Detection and semiquantitation of albumin forms in fresh human plasma separated on gradient polyacrylamide gel by means of electroblotting on agarose gel matrix", *ELECTROPHORESIS*, vol. 18, no. 11, pp. 2055-2058.

Bauer, N., Wilsch-Braeuninger, M., Karbanova, J., Fonseca, A., Strauss, D., Freund, D., Thiele, C., Huttner, W.B., Bornhaeuser, M. & Corbeil, D. (2011), "Haematopoietic stem cell differentiation promotes the release of prominin-1/CD133-containing membrane vesicles-a role of the endocytic-exocytic pathway", *EMBO MOLECULAR MEDICINE*, vol. 3, no. 7, pp. 398-409.

Cheruvanky, A., Zhou, H., Pisitkun, T., Kopp, J.B., Knepper, M.A., Yuen, P.S.T. & Star, R.A. (2007), "Rapid isolation of urinary exosomal biomarkers using a nanomembrane ultrafiltration concentrator", *AMERICAN JOURNAL OF PHYSIOLOGY-RENAL PHYSIOLOGY*, vol. 292, no. 5, pp. F1657-F1661.

Conesa, A., Gotz, S., Garcia-Gomez, J., Terol, J., Talon, M. & Robles, M. (2005), "Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research", *BIOINFORMATICS*, vol. 21, no. 18, pp. 3674-3676.

Fernandez-Llama, P., Khositseth, S., Gonzales, P.A., Star, R.A., Pisitkun, T. & Knepper, M.A. (2010), "Tamm-Horsfall protein and urinary exosome isolation", *JOURNAL OF HYPERTENSION*, vol. 28, no. A, pp. E164.

Gonzales, P.A., Pisitkun, T., Hoffert, J.D., Tchapyjnikov, D., Star, R.A., Kleta, R., Wang, N.S. & Knepper, M.A. (2009), "Large-Scale Proteomics and Phosphoproteomics of Urinary Exosomes", *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY*, vol. 20, no. 2, pp. 363-379.

Hara, M., Yanagihara, T., Hirayama, Y., Ogasawara, S., Kurosawa, H., Sekine, S. & Kihara, I. (2010), "Podocyte membrane vesicles in urine originate from tip vesiculation of podocyte microvilli", *HUMAN PATHOLOGY*, vol. 41, no. 9, pp. 1265-1275.

Hogan, M.C., Manganelli, L., Woollard, J.R., Masyuk, A.I., Masyuk, T.V., Tammachote, R., Huang, B.Q., Leontovich, A.A., Beito, T.G., Madden, B.J., Charlesworth, M.C., Torres, V.E., LaRusso, N.F., Harris, P.C. & Ward, C.J. (2009), "Characterization of PKD Protein-Positive Exosome-Like Vesicles", *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY*, vol. 20, no. 2, pp. 278-288.

Igarashi, K., Kaneda, M., Yamaji, A., Saido, T., Kikkawa, U., Ono, Y., Inoue, K. & Umeda, M. (1995), "A novel phosphatidylserine-binding peptide motif defined by an antiidiotypic monoclonal-antibody - localization of phosphatidylserine-specific binding-sites on protein-kinase-c and phosphatidylserine decarboxylase", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 270, no. 49, pp. 29075-29078.

Keller, S., Ridinger, J., Rupp, A., Janssen, J. & Altevogt, P. (2011), "Body fluid derived exosomes as a novel template for clinical diagnostics", *JOURNAL OF TRANSLATIONAL MEDICINE*, vol. 9, no. 1, pp. 86.

Kobayashi, K. & Fukuoka, S. (2001), "Conditions for solubilization of Tamm-Horsfall protein/uromodulin in human urine and establishment of a sensitive and accurate enzyme-linked immunosorbent assay (ELISA) method", *ARCHIVES OF BIOCHEMISTRY AND BIOPHYSICS*, vol. 388, no. 1, pp. 113-120.

Laulagnier, K., Motta, C., Hamdi, S., Roy, S., Fauvelle, F., Pageaux, J., Kobayashi, T., Salles, J., Perret, B., Bonnerot, C. & Record, M. (2004), "Mast cell- and dendritic cell-derived exosomes display a specific lipid composition and an unusual membrane organization", *BIOCHEMICAL JOURNAL*, vol. 380, no. Part 1, pp. 161-171.

Lescuyer, P., Pernin, A., Hainard, A., Bigeire, C., Burgess, J.A., Zimmermann-Ivol, C., Sanchez, J., Schifferli, J.A., Hochstrasser, D.F. & Moll, S. (2008), "Proteomic analysis of a podocyte vesicle-enriched fraction from human normal and pathological urine samples", *PROTEOMICS CLINICAL APPLICATIONS*, vol. 2, no. 7-8, pp. 1008-1018.

Li, H., Yao, Z., Degenhardt, B., Teper, G. & Papadopoulos, V. (2001), "Cholesterol binding at the cholesterol recognition/ interaction amino acid consensus (CRAC) of the peripheral-type benzodiazepine receptor and inhibition of steroidogenesis by an HIV TAT-CRAC peptide", *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, vol. 98, no. 3, pp. 1267-1272.

Mahfoud, R., Garmy, N., Maresca, M., Yahi, N., Puigserver, A. & Fantini, J. (2002), "Identification of a common sphingolipid-binding domain in Alzheimer, prion, and HIV-1 proteins", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 277, no. 13, pp. 11292-11296.

Marzesco, A., Janich, P., Wilsch-Brauninger, M., Dubreuil, V., Langenfeld, K., Corbeil, D. & Huttner, W. (2005), "Release of extracellular membrane particles carrying the stem cell marker prominin-1 (CD133) from neural progenitors and other epithelial cells", *JOURNAL OF CELL SCIENCE*, vol. 118, no. 13, pp. 2849-2858.

Melrose, J., Ghosh, P. & Patel, M. (1995), "biotinylated aprotinin - a versatile probe for the detection of serine proteinases on western blots", *INTERNATIONAL JOURNAL OF BIOCHEMISTRY \& CELL BIOLOGY*, vol. 27, no. 9, pp. 891-904.

Navab, M., Anantharamaiah, G., Reddy, S., Hama, S., Hough, G., Frank, J., Grijalva, V., Ganesh, V., Mishra, V., Palgunachari, M. & Fogelman, A. (2005), "Oral small peptides render HDL antiinflammatory in mice and monkeys and reduce atherosclerosis in ApoE null mice", *CIRCULATION RESEARCH*, vol. 97, no. 6, pp. 524-532.

Pascual, M., Steiger, G., Sadallah, S., Paccaud, J., Carpentier, J., James, R. & Schifferli, J. (1994), "Identification of membrane-bound CR-1 (CD35) in human urine - evidence for its release by glomerular podocytes", *JOURNAL OF EXPERIMENTAL MEDICINE*, vol. 179, no. 3, pp. 889-899.

Pisitkun, T., Shen, R. & Knepper, M. (2004), "Identification and proteomic profiling of exosomes in human urine", *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, vol. 101, no. 36, pp. 13368-13373.

Rood, I.M., Deegens, J.K.J., Merchant, M.L., Tamboer, W.P.M., Wilkey, D.W., Wetzels, J.F.M. & Klein, J.B. (2010), "Comparison of three methods for isolation of urinary microvesicles to identify biomarkers of nephrotic syndrome", *KIDNEY INTERNATIONAL*, vol. 78, no. 8, pp. 810-816.

Taylor, M., Couto, J., Scallan, C., Ceriani, R. & Peterson, J. (1997), "Lactadherin (Formerly BA46), a membrane-associated glycoprotein expressed in human milk and breast carcinomas, promotes Arg-Gly-Asp (RGD)-dependent cell adhesion", *DNA AND CELL BIOLOGY*, vol. 16, no. 7, pp. 861-869.

Vidal, M., Sainte-Marie, J., Philippot, J.R. & Bienvenue, A. (1989), "Asymmetric distribution of phospholipids in the membrane of vesicles released during in vitro maturation of guinea pig reticulocytes: Evidence precluding a role for an aminophospholipid translocase", *JOURNAL OF CELLULAR PHYSIOLOGY*, vol. 140, no. 3, pp. 455-462.

Wang, Z., Hill, S., Luther, J.M., Hachey, D.L. & Schey, K.L. (2011), "Proteomic analysis of urine exosomes by multidimensional protein identification technology (MudPIT)", *PROTEOMICS*, vol. 12, no. 2, pp. 329-338.

Welton, J.L., Khanna, S., Giles, P.J., Brennan, P., Brewis, I.A., Staffurth, J., Mason, M.D. & Clayton, A. (2010), "Proteomic analysis of bladder cancer exosomes", *MOLECULAR & CELLULAR PROTEOMICS*, vol. 9, no. 6, pp. 1324-1338.

Wiggins, R., Glatfelter, A., Kshirsagar, B. & Beals, T. (1987), "Lipid microvesicles and their association with procoagulant activity in urine and glomeruli of rabbits with nephrotoxic nephritis", *LABORATORY INVESTIGATION*, vol. 56, no. 3, pp. 264-272.

Wubbolts, R., Leckie, R., Veenhuizen, P., Schwarzmann, G., Mobius, W., Hoernschemeyer, J., Slot, J., Geuze, H. & Stoorvogel, W. (2003), "Proteomic and biochemical analyses of human B cell-derived exosomes - Potential implications for their function and multivesicular body formation", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 278, no. 13, pp. 10963-10972.

Supplementary table S4.1: Total proteins identified in the membrane vesicles enriched from P200,000g using P1-Peptide agarose. Uniprot accessions, sequence coverage, PSM, peptides found, number of amino acid in the proteins, molecular weight, calculated PI, mascot score and description of the proteins is given.

Accession	Coverage %	# PSMs	# Peptides	# AAs	MW [kDa]	calc. pI	Score	Description
P62258	12.16	2	2	255	29.2	4.74	135.78	14-3-3 protein epsilon OS=Homo sapiens GN=YWHAE PE=1 SV=1 - [1433E_HUMAN]
P63104	4.90	1	1	245	27.7	4.79	64.56	14-3-3 protein zeta/delta OS=Homo sapiens GN=YWHAZ PE=1 SV=1 - [1433Z_HUMAN]
Q9BUT1	4.90	1	1	245	26.7	7.65	82.20	3-hydroxybutyrate dehydrogenase type 2 OS=Homo sapiens GN=BDH2 PE=1 SV=2 - [BDH2_HUMAN]
Q8TCD5	8.96	1	1	201	23.4	6.64	80.01	5'(3')-deoxyribonucleotidase, cytosolic type OS=Homo sapiens GN=NT5C PE=1 SV=2 - [NT5C_HUMAN]
P17858	1.92	1	1	780	85.0	7.50	65.71	6-phosphofructokinase, liver type OS=Homo sapiens GN=PFKL PE=1 SV=6 - [K6PL_HUMAN]
Q96IU4	16.19	3	3	210	22.3	6.40	108.53	Abhydrolase domain-containing protein 14B OS=Homo sapiens GN=ABHD14B PE=1 SV=1 - [ABHEB_HUMAN]
Q13510	25.57	12	7	395	44.6	7.62	474.76	Acid ceramidase OS=Homo sapiens GN=ASAH1 PE=1 SV=5 - [ASAH1_HUMAN]
P60709	17.87	6	5	375	41.7	5.48	325.20	Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1 - [ACTB_HUMAN]
P23526	6.71	3	3	432	47.7	6.34	122.91	Adenosylhomocysteinase OS=Homo sapiens GN=AHCY PE=1 SV=4 - [SAHH_HUMAN]
Q06278	1.20	1	1	1338	147.8	7.17	66.61	Aldehyde oxidase OS=Homo sapiens GN=AOX1 PE=2 SV=2 - [ADO_HUMAN]
P01011	19.86	6	6	423	47.6	5.52	343.23	Alpha-1-antichymotrypsin OS=Homo sapiens GN=SERPINA3 PE=1 SV=2 - [AACT_HUMAN]
P01009	33.25	15	11	418	46.7	5.59	702.93	Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3 - [A1AT_HUMAN]
P08697	15.27	5	5	491	54.5	6.29	209.54	Alpha-2-antiplasmin OS=Homo sapiens GN=SERPINF2 PE=1 SV=3 - [A2AP_HUMAN]
P02765	3.81	1	1	367	39.3	5.72	60.41	Alpha-2-HS-glycoprotein OS=Homo sapiens GN=AHSG PE=1 SV=1 - [FETUA_HUMAN]
P04745	20.16	6	6	511	57.7	6.93	301.99	Alpha-amylase 1 OS=Homo sapiens GN=AMY1A PE=1 SV=2 - [AMY1_HUMAN]
P19961	16.05	6	6	511	57.7	7.09	321.43	Alpha-amylase 2B OS=Homo sapiens GN=AMY2B PE=1 SV=1 - [AMY2B_HUMAN]
P06733	8.53	2	2	434	47.1	7.39	131.52	Alpha-enolase OS=Homo sapiens GN=ENO1 PE=1 SV=2 - [ENOA_HUMAN]
P06280	6.76	2	2	429	48.7	5.60	138.47	Alpha-galactosidase A OS=Homo sapiens GN=GLA PE=1 SV=1 - [AGAL_HUMAN]

P17050	6.57	2	2	411	46.5	5.19	105.96	Alpha-N-acetylgalactosaminidase OS=Homo sapiens GN=NAGA PE=1 SV=2 - [NAGAB_HUMAN]
P54802	23.69	14	12	743	82.1	6.54	686.14	Alpha-N-acetylglucosaminidase OS=Homo sapiens GN=NAGLU PE=1 SV=1 - [ANAG_HUMAN]
P19801	3.73	2	2	751	85.3	7.09	92.58	Amiloride-sensitive amine oxidase [copper-containing] OS=Homo sapiens GN=ABP1 PE=1 SV=4 - [ABP1_HUMAN]
P15144	28.23	24	18	967	109.5	5.48	1093.36	Aminopeptidase N OS=Homo sapiens GN=ANPEP PE=1 SV=4 - [AMPN_HUMAN]
Q9UBD6	2.71	1	1	479	53.1	6.39	76.92	Ammonium transporter Rh type C OS=Homo sapiens GN=RHCG PE=1 SV=1 - [RHCG_HUMAN]
Q9UKU9	5.88	3	3	493	57.1	7.53	118.68	Angiopoietin-related protein 2 OS=Homo sapiens GN=ANGPTL2 PE=2 SV=1 - [ANGL2_HUMAN]
P12821	1.07	1	1	1306	149.6	6.39	71.25	Angiotensin-converting enzyme OS=Homo sapiens GN=ACE PE=1 SV=1 - [ACE_HUMAN]
P50995	5.35	2	2	505	54.4	7.65	120.56	Annexin A11 OS=Homo sapiens GN=ANXA11 PE=1 SV=1 - [ANX11_HUMAN]
P01008	13.58	4	4	464	52.6	6.71	200.24	Antithrombin-III OS=Homo sapiens GN=SERPINC1 PE=1 SV=1 - [ANT3_HUMAN]
P05090	38.10	16	10	189	21.3	5.15	548.73	Apolipoprotein D OS=Homo sapiens GN=APOD PE=1 SV=1 - [APOD_HUMAN]
P02649	21.14	6	5	317	36.1	5.73	319.65	Apolipoprotein E OS=Homo sapiens GN=APOE PE=1 SV=1 - [APOE_HUMAN]
P29972	5.58	1	1	269	28.5	7.42	110.95	Aquaporin-1 OS=Homo sapiens GN=AQP1 PE=1 SV=3 - [AQP1_HUMAN]
P00966	10.19	3	3	412	46.5	8.02	165.18	Argininosuccinate synthase OS=Homo sapiens GN=ASS1 PE=1 SV=2 - [ASSY_HUMAN]
P15289	3.35	1	1	507	53.6	6.07	88.12	Arylsulfatase A OS=Homo sapiens GN=ARSA PE=1 SV=3 - [ARSA_HUMAN]
P25705	2.35	1	1	553	59.7	9.13	60.23	ATP synthase subunit alpha, mitochondrial OS=Homo sapiens GN=ATP5A1 PE=1 SV=1 - [ATPA_HUMAN]
P06576	2.46	1	1	529	56.5	5.40	88.64	ATP synthase subunit beta, mitochondrial OS=Homo sapiens GN=ATP5B PE=1 SV=3 - [ATPB_HUMAN]
O75882	7.35	9	8	1429	158.4	7.31	385.18	Attractin OS=Homo sapiens GN=ATRN PE=1 SV=2 - [ATRN_HUMAN]
P98160	1.62	5	5	4391	468.5	6.51	227.69	Basement membrane-specific heparan sulfate proteoglycan core protein OS=Homo sapiens GN=HSPG2 PE=1 SV=3 - [PGBM_HUMAN]
P15291	5.03	1	1	398	43.9	8.65	96.50	Beta-1,4-galactosyltransferase 1 OS=Homo sapiens GN=B4GALT1 PE=1 SV=5 - [B4GT1_HUMAN]
P16278	24.08	15	14	677	76.0	6.57	595.22	Beta-galactosidase OS=Homo sapiens GN=GLB1 PE=1 SV=2 - [BGAL_HUMAN]
P08236	4.30	2	2	651	74.7	7.02	104.80	Beta-glucuronidase OS=Homo sapiens GN=GUSB PE=1 SV=2 - [BGLR_HUMAN]
P06865	5.86	2	2	529	60.7	5.16	115.60	Beta-hexosaminidase subunit alpha OS=Homo sapiens GN=HEXA PE=1 SV=1 - [HEXA_HUMAN]

P07686	1.80	1	1	556	63.1	6.76	62.70	Beta-hexosaminidase subunit beta OS=Homo sapiens GN=HEXB PE=1 SV=3 - [HEXB_HUMAN]
Q93088	19.95	6	6	406	45.0	7.03	332.85	Betaine--homocysteine S-methyltransferase 1 OS=Homo sapiens GN=BHMT PE=1 SV=2 - [BHMT1_HUMAN]
O00462	1.48	1	1	879	100.8	5.52	79.44	Beta-mannosidase OS=Homo sapiens GN=MANBA PE=1 SV=3 - [MANBA_HUMAN]
P52848	0.79	1	1	882	100.8	7.97	62.44	Bifunctional heparan sulfate N-deacetylase/N-sulfotransferase 1 OS=Homo sapiens GN=NDST1 PE=1 SV=1 - [NDST1_HUMAN]
P21810	2.99	1	1	368	41.6	7.52	68.04	Biglycan OS=Homo sapiens GN=BGN PE=1 SV=2 - [PGS1_HUMAN]
Q5VW32	6.33	2	2	411	46.4	7.65	98.65	BRO1 domain-containing protein BROX OS=Homo sapiens GN=BROX PE=1 SV=1 - [BROX_HUMAN]
P12830	2.61	3	2	882	97.4	4.73	133.91	Cadherin-1 OS=Homo sapiens GN=CDH1 PE=1 SV=3 - [CADH1_HUMAN]
Q9BYE9	6.49	6	6	1310	141.5	4.50	251.84	Cadherin-related family member 2 OS=Homo sapiens GN=CDHR2 PE=1 SV=2 - [CDHR2_HUMAN]
Q9HBB8	4.50	2	2	845	88.2	4.93	86.73	Cadherin-related family member 5 OS=Homo sapiens GN=CDHR5 PE=1 SV=3 - [CDHR5_HUMAN]
P00918	6.15	1	1	260	29.2	7.40	69.24	Carbonic anhydrase 2 OS=Homo sapiens GN=CA2 PE=1 SV=2 - [CAH2_HUMAN]
P22792	8.81	4	3	545	60.6	5.99	201.13	Carboxypeptidase N subunit 2 OS=Homo sapiens GN=CPN2 PE=1 SV=2 - [CPN2_HUMAN]
P07858	12.68	4	3	339	37.8	6.30	222.70	Cathepsin B OS=Homo sapiens GN=CTSB PE=1 SV=3 - [CATB_HUMAN]
P07339	5.10	2	2	412	44.5	6.54	115.82	Cathepsin D OS=Homo sapiens GN=CTSD PE=1 SV=1 - [CATD_HUMAN]
P09668	12.54	3	2	335	37.4	8.07	195.04	Cathepsin H OS=Homo sapiens GN=CTSH PE=1 SV=4 - [CATH_HUMAN]
P11717	3.17	5	5	2491	274.1	5.91	167.49	Cation-independent mannose-6-phosphate receptor OS=Homo sapiens GN=IGF2R PE=1 SV=2 - [MPRI_HUMAN]
P13987	15.63	3	2	128	14.2	6.48	154.26	CD59 glycoprotein OS=Homo sapiens GN=CD59 PE=1 SV=1 - [CD59_HUMAN]
P08962	7.56	4	3	238	25.6	7.81	153.86	CD63 antigen OS=Homo sapiens GN=CD63 PE=1 SV=2 - [CD63_HUMAN]
P27701	3.75	1	1	267	29.6	5.24	61.76	CD82 antigen OS=Homo sapiens GN=CD82 PE=1 SV=1 - [CD82_HUMAN]
P21926	15.35	2	2	228	25.4	7.15	138.54	CD9 antigen OS=Homo sapiens GN=CD9 PE=1 SV=4 - [CD9_HUMAN]
P00450	27.98	20	18	1065	122.1	5.72	809.80	Ceruloplasmin OS=Homo sapiens GN=CP PE=1 SV=1 - [CERU_HUMAN]
P11597	25.15	11	11	493	54.7	6.09	406.75	Cholesteryl ester transfer protein OS=Homo sapiens GN=CETP PE=1 SV=2 - [CETP_HUMAN]
Q8WWI5	6.09	3	3	657	73.3	8.60	98.54	Choline transporter-like protein 1 OS=Homo sapiens GN=SLC44A1 PE=1 SV=1 - [CTL1_HUMAN]
Q8IWA5	14.02	7	7	706	80.1	8.62	270.76	Choline transporter-like protein 2 OS=Homo sapiens GN=SLC44A2 PE=1 SV=2 - [CTL2_HUMAN]

Q53GD3	7.61	10	6	710	79.2	8.59	361.93	Choline transporter-like protein 4 OS=Homo sapiens GN=SLC44A4 PE=2 SV=1 - [CTL4_HUMAN]
P10909	26.95	13	10	449	52.5	6.27	565.78	Clusterin OS=Homo sapiens GN=CLU PE=1 SV=1 - [CLUS_HUMAN]
P12109	5.06	3	3	1028	108.5	5.43	141.93	Collagen alpha-1(VI) chain OS=Homo sapiens GN=COL6A1 PE=1 SV=3 - [CO6A1_HUMAN]
P39059	1.95	2	2	1388	141.6	5.00	99.58	Collagen alpha-1(XV) chain OS=Homo sapiens GN=COL15A1 PE=1 SV=2 - [COFA1_HUMAN]
Q9HBJ8	5.86	1	1	222	25.2	5.63	58.38	Collectrin OS=Homo sapiens GN=TMEM27 PE=1 SV=1 - [TMM27_HUMAN]
Q9NZP8	6.37	4	2	487	53.5	7.20	191.48	Complement C1r subcomponent-like protein OS=Homo sapiens GN=C1RL PE=1 SV=2 - [C1RL_HUMAN]
P01024	12.69	17	16	1663	187.0	6.40	765.93	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2 - [CO3_HUMAN]
P0C0L4	5.85	5	5	1744	192.7	7.08	248.32	Complement C4-A OS=Homo sapiens GN=C4A PE=1 SV=1 - [CO4A_HUMAN]
P02748	3.94	2	2	559	63.1	5.59	81.66	Complement component C9 OS=Homo sapiens GN=C9 PE=1 SV=2 - [CO9_HUMAN]
Q2VPA4	2.11	1	1	569	62.7	7.23	59.58	Complement component receptor 1-like protein OS=Homo sapiens GN=CR1L PE=1 SV=3 - [CR1L_HUMAN]
P08603	1.06	1	1	1231	139.0	6.61	75.96	Complement factor H OS=Homo sapiens GN=CFH PE=1 SV=4 - [CFAH_HUMAN]
Q12860	1.87	2	2	1018	113.2	5.90	74.32	Contactin-1 OS=Homo sapiens GN=CNTN1 PE=1 SV=1 - [CNTN1_HUMAN]
O60494	22.77	77	55	3623	398.4	5.35	3521.98	Cubilin OS=Homo sapiens GN=CUBN PE=1 SV=4 - [CUBN_HUMAN]
P01034	10.96	1	1	146	15.8	8.75	91.29	Cystatin-C OS=Homo sapiens GN=CST3 PE=1 SV=1 - [CYTC_HUMAN]
Q96KP4	7.37	2	2	475	52.8	5.97	131.04	Cytosolic non-specific dipeptidase OS=Homo sapiens GN=CNDP2 PE=1 SV=2 - [CNDP2_HUMAN]
P24855	6.74	2	2	282	31.4	4.91	105.13	Deoxyribonuclease-1 OS=Homo sapiens GN=DNASE1 PE=1 SV=1 - [DNAS1_HUMAN]
P81605	22.73	2	2	110	11.3	6.54	118.08	Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2 - [DCD_HUMAN]
P15924	0.91	2	2	2871	331.6	6.81	121.81	Desmoplakin OS=Homo sapiens GN=DSP PE=1 SV=3 - [DESP_HUMAN]
P16444	19.95	6	5	411	45.6	6.15	272.09	Dipeptidase 1 OS=Homo sapiens GN=DPEP1 PE=1 SV=3 - [DPEP1_HUMAN]
P53634	24.19	10	8	463	51.8	6.99	360.09	Dipeptidyl peptidase 1 OS=Homo sapiens GN=CTSC PE=1 SV=1 - [CATC_HUMAN]
Q9UHL4	17.48	9	6	492	54.3	6.32	387.29	Dipeptidyl peptidase 2 OS=Homo sapiens GN=DPP7 PE=1 SV=3 - [DPP2_HUMAN]
P27487	12.01	8	8	766	88.2	6.04	392.73	Dipeptidyl peptidase 4 OS=Homo sapiens GN=DPP4 PE=1 SV=2 - [DPP4_HUMAN]
Q12805	27.99	12	9	493	54.6	5.07	549.25	EGF-containing fibulin-like extracellular matrix protein 1 OS=Homo sapiens GN=EFEMP1 PE=1 SV=2 - [FBLN3_HUMAN]
P68104	8.87	2	2	462	50.1	9.01	89.82	Elongation factor 1-alpha 1 OS=Homo sapiens GN=EEF1A1 PE=1

								SV=1 - [EF1A1_HUMAN]
Q9Y2E5	3.77	3	3	1009	113.9	7.24	150.15	Epididymis-specific alpha-mannosidase OS=Homo sapiens GN=MAN2B2 PE=1 SV=4 - [MA2B2_HUMAN]
P33947	10.38	2	2	212	24.4	8.72	112.87	ER lumen protein retaining receptor 2 OS=Homo sapiens GN=KDELR2 PE=1 SV=1 - [ERD22_HUMAN]
P56537	7.35	1	1	245	26.6	4.68	70.15	Eukaryotic translation initiation factor 6 OS=Homo sapiens GN=EIF6 PE=1 SV=1 - [IF6_HUMAN]
P43005	3.44	1	1	524	57.1	5.71	77.44	Excitatory amino acid transporter 3 OS=Homo sapiens GN=SLC1A1 PE=1 SV=2 - [EAA3_HUMAN]
P15311	2.90	1	1	586	69.4	6.27	88.81	Ezrin OS=Homo sapiens GN=EZR PE=1 SV=4 - [EZRI_HUMAN]
P35555	0.45	1	1	2871	312.1	4.93	64.99	Fibrillin-1 OS=Homo sapiens GN=FBN1 PE=1 SV=2 - [FBN1_HUMAN]
P02671	12.70	12	6	866	94.9	6.01	558.20	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2 - [FIBA_HUMAN]
Q14314	12.30	5	4	439	50.2	7.39	264.28	Fibroleukin OS=Homo sapiens GN=FGL2 PE=1 SV=1 - [FGL2_HUMAN]
P02751	12.57	19	17	2386	262.5	5.71	843.15	Fibronectin OS=Homo sapiens GN=FN1 PE=1 SV=4 - [FINC_HUMAN]
P23142	7.54	5	5	703	77.2	5.22	207.01	Fibulin-1 OS=Homo sapiens GN=FBLN1 PE=1 SV=4 - [FBLN1_HUMAN]
Q5D862	2.55	2	2	2391	247.9	8.31	75.61	Filaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1 - [FILA2_HUMAN]
P15328	7.39	2	2	257	29.8	7.97	89.55	Folate receptor alpha OS=Homo sapiens GN=FOLR1 PE=1 SV=3 - [FOLR1_HUMAN]
Q5SZK8	2.02	5	5	3169	350.9	5.03	167.96	FRAS1-related extracellular matrix protein 2 OS=Homo sapiens GN=FREM2 PE=1 SV=2 - [FREM2_HUMAN]
Q9ULV1	2.98	1	1	537	59.8	7.27	71.30	Frizzled-4 OS=Homo sapiens GN=FZD4 PE=1 SV=2 - [FZD4_HUMAN]
P05062	5.49	2	2	364	39.4	7.87	76.62	Fructose-bisphosphate aldolase B OS=Homo sapiens GN=ALDOB PE=1 SV=2 - [ALDOB_HUMAN]
P54803	2.19	1	1	685	77.0	6.64	85.96	Galactocerebrosidase OS=Homo sapiens GN=GALC PE=1 SV=2 - [GALC_HUMAN]
Q08380	30.94	20	13	585	65.3	5.27	1039.17	Galectin-3-binding protein OS=Homo sapiens GN=LGALS3BP PE=1 SV=1 - [LG3BP_HUMAN]
O00182	4.51	1	1	355	39.5	9.17	88.46	Galectin-9 OS=Homo sapiens GN=LGALS9 PE=1 SV=2 - [LEG9_HUMAN]
Q92820	20.13	6	5	318	35.9	7.11	242.07	Gamma-glutamyl hydrolase OS=Homo sapiens GN=GGH PE=1 SV=2 - [GGH_HUMAN]
P19440	4.75	2	2	569	61.4	7.12	120.62	Gamma-glutamyltranspeptidase 1 OS=Homo sapiens GN=GGT1 PE=1 SV=2 - [GGT1_HUMAN]
P13284	4.21	1	1	261	29.1	4.98	83.20	Gamma-interferon-inducible lysosomal thiol reductase OS=Homo sapiens GN=IFI30 PE=1 SV=2 - [GILT_HUMAN]
P17900	15.54	2	2	193	20.8	5.31	86.43	Ganglioside GM2 activator OS=Homo sapiens GN=GM2A PE=1 SV=4 - [SAP3_HUMAN]

P07093	7.54	2	2	398	44.0	9.29	79.58	Glia-derived nexin OS=Homo sapiens GN=SERPINE2 PE=1 SV=1 - [GDN_HUMAN]
P06744	3.05	1	1	558	63.1	8.32	63.37	Glucose-6-phosphate isomerase OS=Homo sapiens GN=GPI PE=1 SV=4 - [G6PI_HUMAN]
Q16769	11.08	3	3	361	40.9	6.61	135.96	Glutaminyl-peptide cyclotransferase OS=Homo sapiens GN=QPCT PE=1 SV=1 - [QPCT_HUMAN]
Q07075	5.75	4	3	957	109.2	5.47	206.36	Glutamyl aminopeptidase OS=Homo sapiens GN=ENPEP PE=1 SV=3 - [AMPE_HUMAN]
P21266	25.78	5	5	225	26.5	5.54	276.42	Glutathione S-transferase Mu 3 OS=Homo sapiens GN=GSTM3 PE=1 SV=3 - [GSTM3_HUMAN]
P04406	33.43	8	7	335	36.0	8.46	336.18	Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens GN=GAPDH PE=1 SV=3 - [G3P_HUMAN]
P51654	5.17	2	2	580	65.5	6.37	128.15	Glypican-3 OS=Homo sapiens GN=GPC3 PE=1 SV=1 - [GPC3_HUMAN]
Q92896	1.44	2	2	1179	134.5	6.90	93.01	Golgi apparatus protein 1 OS=Homo sapiens GN=GLG1 PE=1 SV=2 - [GSLG1_HUMAN]
Q8NBJ4	5.74	2	2	401	45.3	4.97	108.52	Golgi membrane protein 1 OS=Homo sapiens GN=GOLM1 PE=1 SV=1 - [GOLM1_HUMAN]
Q9NQ84	18.82	6	5	441	48.2	8.43	303.35	G-protein coupled receptor family C group 5 member C OS=Homo sapiens GN=GPRC5C PE=1 SV=2 - [GPC5C_HUMAN]
P63096	6.21	2	2	354	40.3	5.97	78.40	Guanine nucleotide-binding protein G(i) subunit alpha-1 OS=Homo sapiens GN=GNAI1 PE=1 SV=2 - [GNAI1_HUMAN]
Q9UBI6	15.28	1	1	72	8.0	8.97	73.78	Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-12 OS=Homo sapiens GN=GNG12 PE=1 SV=3 - [GBG12_HUMAN]
P62873	11.47	3	3	340	37.4	6.00	140.70	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1 OS=Homo sapiens GN=GNB1 PE=1 SV=3 - [GBB1_HUMAN]
P62879	11.47	3	3	340	37.3	6.00	159.26	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-2 OS=Homo sapiens GN=GNB2 PE=1 SV=3 - [GBB2_HUMAN]
P54652	2.50	1	1	639	70.0	5.74	75.89	Heat shock-related 70 kDa protein 2 OS=Homo sapiens GN=HSPA2 PE=1 SV=1 - [HSP72_HUMAN]
P69905	10.56	1	1	142	15.2	8.68	70.76	Hemoglobin subunit alpha OS=Homo sapiens GN=HBA1 PE=1 SV=2 - [HBA_HUMAN]
P68871	53.74	7	6	147	16.0	7.28	276.73	Hemoglobin subunit beta OS=Homo sapiens GN=HBB PE=1 SV=2 - [HBB_HUMAN]
Q93099	2.25	1	1	445	49.9	6.96	62.23	Homogentisate 1,2-dioxygenase OS=Homo sapiens GN=HGD PE=1 SV=2 - [HGD_HUMAN]
Q86YZ3	9.23	6	6	2850	282.2	10.04	284.89	Hornerin OS=Homo sapiens GN=HRNR PE=1 SV=2 - [HORN_HUMAN]
P01876	37.68	17	12	353	37.6	6.51	687.77	Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=2 - [IGHA1_HUMAN]
P01877	29.12	11	8	340	36.5	6.10	431.51	Ig alpha-2 chain C region OS=Homo sapiens GN=IGHA2 PE=1 SV=3 - [IGHA2_HUMAN]
P01857	28.79	6	6	330	36.1	8.19	249.68	Ig gamma-1 chain C region OS=Homo sapiens GN=IGHG1 PE=1

								SV=1 - [IGHG1_HUMAN]
P01859	7.98	2	2	326	35.9	7.59	176.29	Ig gamma-2 chain C region OS=Homo sapiens GN=IGHG2 PE=1 SV=2 - [IGHG2_HUMAN]
P01742	10.26	4	1	117	12.5	6.57	113.67	Ig heavy chain V-I region EU OS=Homo sapiens PE=1 SV=1 - [HV101_HUMAN]
P01743	22.22	2	2	117	12.9	8.92	90.03	Ig heavy chain V-I region HG3 OS=Homo sapiens PE=4 SV=1 - [HV102_HUMAN]
P06331	10.96	1	1	146	16.2	8.28	80.27	Ig heavy chain V-II region ARH-77 OS=Homo sapiens PE=4 SV=1 - [HV209_HUMAN]
P01766	25.00	4	2	120	13.2	6.57	213.66	Ig heavy chain V-III region BRO OS=Homo sapiens PE=1 SV=1 - [HV305_HUMAN]
P01781	18.97	2	2	116	12.7	8.48	110.43	Ig heavy chain V-III region GAL OS=Homo sapiens PE=1 SV=1 - [HV320_HUMAN]
P01771	9.09	1	1	121	13.6	9.36	67.49	Ig heavy chain V-III region HIL OS=Homo sapiens PE=1 SV=1 - [HV310_HUMAN]
P01765	26.09	3	2	115	12.3	9.13	171.68	Ig heavy chain V-III region TIL OS=Homo sapiens PE=1 SV=1 - [HV304_HUMAN]
P01762	15.57	3	1	122	13.5	9.72	142.18	Ig heavy chain V-III region TRO OS=Homo sapiens PE=1 SV=1 - [HV301_HUMAN]
P01779	26.72	2	2	116	12.4	9.73	121.71	Ig heavy chain V-III region TUR OS=Homo sapiens PE=1 SV=1 - [HV318_HUMAN]
P01764	18.80	3	2	117	12.6	8.28	164.88	Ig heavy chain V-III region VH26 OS=Homo sapiens PE=1 SV=1 - [HV303_HUMAN]
P01834	80.19	13	6	106	11.6	5.87	685.01	Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 SV=1 - [IGKC_HUMAN]
P01593	31.48	4	2	108	12.0	5.99	218.14	Ig kappa chain V-I region AG OS=Homo sapiens PE=1 SV=1 - [KV101_HUMAN]
P01596	16.82	1	1	107	11.7	9.41	102.63	Ig kappa chain V-I region CAR OS=Homo sapiens PE=1 SV=1 - [KV104_HUMAN]
P01597	16.67	12	1	108	11.7	9.36	186.48	Ig kappa chain V-I region DEE OS=Homo sapiens PE=1 SV=1 - [KV105_HUMAN]
P01613	30.36	2	2	112	12.2	5.36	108.29	Ig kappa chain V-I region Ni OS=Homo sapiens PE=1 SV=1 - [KV121_HUMAN]
P01611	16.67	1	1	108	11.6	7.28	77.26	Ig kappa chain V-I region Wes OS=Homo sapiens PE=1 SV=1 - [KV119_HUMAN]
P01616	11.61	1	1	112	12.0	9.29	86.34	Ig kappa chain V-II region MIL OS=Homo sapiens PE=1 SV=1 - [KV203_HUMAN]
P01617	21.24	1	1	113	12.3	6.00	117.46	Ig kappa chain V-II region TEW OS=Homo sapiens PE=1 SV=1 - [KV204_HUMAN]
P01620	31.19	2	2	109	11.8	8.48	119.81	Ig kappa chain V-III region SIE OS=Homo sapiens PE=1 SV=1 - [KV302_HUMAN]
P06312	29.75	3	3	121	13.4	5.25	96.97	Ig kappa chain V-IV region (Fragment) OS=Homo sapiens GN=IGKV4-1 PE=4 SV=1 - [KV401_HUMAN]
P01625	36.84	4	3	114	12.6	7.93	148.67	Ig kappa chain V-IV region Len OS=Homo sapiens PE=1 SV=2 - [KV402_HUMAN]
P04208	11.93	1	1	109	11.7	6.54	109.57	Ig lambda chain V-I region WAH OS=Homo sapiens PE=1 SV=1 -

								[LV106_HUMAN]
P80748	21.62	2	2	111	11.9	5.08	114.11	Ig lambda chain V-III region LOI OS=Homo sapiens PE=1 SV=1 - [LV302_HUMAN]
P01714	25.00	2	2	108	11.4	6.52	116.26	Ig lambda chain V-III region SH OS=Homo sapiens PE=1 SV=1 - [LV301_HUMAN]
P0CG04	77.36	10	6	106	11.3	7.87	493.61	Ig lambda-1 chain C regions OS=Homo sapiens GN=IGLC1 PE=1 SV=1 - [LAC1_HUMAN]
P0CG05	75.47	10	6	106	11.3	7.24	514.52	Ig lambda-2 chain C regions OS=Homo sapiens GN=IGLC2 PE=1 SV=1 - [LAC2_HUMAN]
P01871	9.29	3	3	452	49.3	6.77	162.54	Ig mu chain C region OS=Homo sapiens GN=IGHM PE=1 SV=3 - [IGHM_HUMAN]
Q9Y6R7	7.25	14	11	5405	571.6	5.34	607.42	IgGFc-binding protein OS=Homo sapiens GN=FCGBP PE=1 SV=3 - [FCGBP_HUMAN]
P01591	16.98	2	2	159	18.1	5.24	103.02	Immunoglobulin J chain OS=Homo sapiens GN=IGJ PE=1 SV=4 - [IGJ_HUMAN]
Q969P0	1.63	1	1	613	65.0	8.00	65.85	Immunoglobulin superfamily member 8 OS=Homo sapiens GN=IGSF8 PE=1 SV=1 - [IGSF8_HUMAN]
P18065	3.69	1	1	325	34.8	7.50	68.44	Insulin-like growth factor-binding protein 2 OS=Homo sapiens GN=IGFBP2 PE=1 SV=2 - [IBP2_HUMAN]
Q16270	15.25	3	3	282	29.1	7.90	129.74	Insulin-like growth factor-binding protein 7 OS=Homo sapiens GN=IGFBP7 PE=1 SV=1 - [IBP7_HUMAN]
P35858	2.64	1	1	605	66.0	6.79	79.36	Insulin-like growth factor-binding protein complex acid labile subunit OS=Homo sapiens GN=IGFALS PE=1 SV=1 - [ALS_HUMAN]
Q9Y287	7.89	2	2	266	30.3	5.14	99.50	Integral membrane protein 2B OS=Homo sapiens GN=ITM2B PE=1 SV=1 - [ITM2B_HUMAN]
Q14624	13.76	8	8	930	103.3	6.98	332.69	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4 - [ITIH4_HUMAN]
P53990	11.54	5	5	364	39.7	5.35	184.03	IST1 homolog OS=Homo sapiens GN=KIAA0174 PE=1 SV=1 - [IST1_HUMAN]
P06870	18.32	4	3	262	28.9	4.83	192.89	Kallikrein-1 OS=Homo sapiens GN=KLK1 PE=1 SV=2 - [KLK1_HUMAN]
P29622	21.08	7	7	427	48.5	7.75	265.06	Kallistatin OS=Homo sapiens GN=SERPINA4 PE=1 SV=3 - [KAIN_HUMAN]
P13645	50.68	45	23	584	58.8	5.21	2028.78	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6 - [K1C10_HUMAN]
P13646	10.92	6	5	458	49.6	4.96	261.83	Keratin, type I cytoskeletal 13 OS=Homo sapiens GN=KRT13 PE=1 SV=4 - [K1C13_HUMAN]
P02533	48.73	18	18	472	51.5	5.16	808.54	Keratin, type I cytoskeletal 14 OS=Homo sapiens GN=KRT14 PE=1 SV=4 - [K1C14_HUMAN]
P19012	10.53	5	5	456	49.2	4.77	266.13	Keratin, type I cytoskeletal 15 OS=Homo sapiens GN=KRT15 PE=1 SV=2 - [K1C15_HUMAN]
P08779	39.96	17	15	473	51.2	5.05	722.80	Keratin, type I cytoskeletal 16 OS=Homo sapiens GN=KRT16 PE=1 SV=4 - [K1C16_HUMAN]
P35527	63.24	40	25	623	62.0	5.24	1805.64	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1

								SV=3 - [K1C9_HUMAN]
P04264	51.09	48	30	644	66.0	8.12	1995.10	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6 - [K2C1_HUMAN]
Q7Z794	3.82	3	2	576	61.7	5.85	113.44	Keratin, type II cytoskeletal 1b OS=Homo sapiens GN=KRT77 PE=1 SV=2 - [K2C1B_HUMAN]
P35908	61.97	33	29	639	65.4	8.00	1621.97	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2 - [K22E_HUMAN]
P12035	6.36	4	4	629	64.5	6.48	203.80	Keratin, type II cytoskeletal 3 OS=Homo sapiens GN=KRT3 PE=1 SV=2 - [K2C3_HUMAN]
P19013	3.75	3	2	534	57.2	6.61	114.19	Keratin, type II cytoskeletal 4 OS=Homo sapiens GN=KRT4 PE=1 SV=4 - [K2C4_HUMAN]
P13647	28.47	19	16	590	62.3	7.74	735.76	Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 PE=1 SV=3 - [K2C5_HUMAN]
P02538	32.27	20	17	564	60.0	8.00	731.15	Keratin, type II cytoskeletal 6A OS=Homo sapiens GN=KRT6A PE=1 SV=3 - [K2C6A_HUMAN]
P04259	27.48	19	15	564	60.0	8.00	677.45	Keratin, type II cytoskeletal 6B OS=Homo sapiens GN=KRT6B PE=1 SV=5 - [K2C6B_HUMAN]
Q14CN4	4.11	3	2	511	55.8	6.89	139.87	Keratin, type II cytoskeletal 72 OS=Homo sapiens GN=KRT72 PE=1 SV=2 - [K2C72_HUMAN]
Q7RTS7	3.97	3	3	529	57.8	7.71	127.60	Keratin, type II cytoskeletal 74 OS=Homo sapiens GN=KRT74 PE=1 SV=2 - [K2C74_HUMAN]
P01042	18.79	12	9	644	71.9	6.81	537.37	Kininogen-1 OS=Homo sapiens GN=KNG1 PE=1 SV=2 - [KNG1_HUMAN]
Q9Y2S2	13.79	3	3	319	35.4	6.18	144.67	Lambda-crystallin homolog OS=Homo sapiens GN=CRYL1 PE=1 SV=3 - [CRYL1_HUMAN]
P18428	19.75	9	7	481	53.3	6.70	378.83	Lipopolysaccharide-binding protein OS=Homo sapiens GN=LBP PE=1 SV=3 - [LBP_HUMAN]
P00338	18.37	6	5	332	36.7	8.27	275.52	L-lactate dehydrogenase A chain OS=Homo sapiens GN=LDHA PE=1 SV=2 - [LDHA_HUMAN]
P07195	26.65	8	6	334	36.6	6.05	335.02	L-lactate dehydrogenase B chain OS=Homo sapiens GN=LDHB PE=1 SV=2 - [LDHB_HUMAN]
P98164	27.78	132	86	4655	521.6	5.08	5205.16	Low-density lipoprotein receptor-related protein 2 OS=Homo sapiens GN=LRP2 PE=1 SV=3 - [LRP2_HUMAN]
P11117	7.09	3	3	423	48.3	6.74	130.41	Lysosomal acid phosphatase OS=Homo sapiens GN=ACP2 PE=1 SV=3 - [PPAL_HUMAN]
P10253	7.46	6	6	952	105.3	5.99	264.77	Lysosomal alpha-glucosidase OS=Homo sapiens GN=GAA PE=1 SV=3 - [LYAG_HUMAN]
P10619	5.63	2	2	480	54.4	6.61	138.23	Lysosomal protective protein OS=Homo sapiens GN=CTSA PE=1 SV=2 - [PPGB_HUMAN]
P42785	1.81	1	1	496	55.8	7.21	78.21	Lysosomal Pro-X carboxypeptidase OS=Homo sapiens GN=PRCP PE=1 SV=1 - [PCP_HUMAN]
P11279	6.47	2	2	417	44.9	8.75	128.58	Lysosome-associated membrane glycoprotein 1 OS=Homo sapiens GN=LAMP1 PE=1 SV=3 - [LAMP1_HUMAN]
P13473	6.34	2	2	410	44.9	5.63	125.07	Lysosome-associated membrane glycoprotein 2 OS=Homo sapiens GN=LAMP2 PE=1 SV=2 -

								[LAMP2_HUMAN]
P61626	8.11	1	1	148	16.5	9.16	65.16	Lysozyme C OS=Homo sapiens GN=LYZ PE=1 SV=1 - [LYSC_HUMAN]
P14174	7.83	1	1	115	12.5	7.88	58.96	Macrophage migration inhibitory factor OS=Homo sapiens GN=MIF PE=1 SV=4 - [MIF_HUMAN]
O43451	3.88	4	4	1857	209.7	5.50	226.10	Maltase-glucoamylase, intestinal OS=Homo sapiens GN=MGAM PE=1 SV=5 - [MGA_HUMAN]
O00187	7.73	6	4	686	75.7	5.77	319.05	Mannan-binding lectin serine protease 2 OS=Homo sapiens GN=MASP2 PE=1 SV=3 - [MASP2_HUMAN]
P33908	19.45	14	9	653	72.9	6.47	636.82	Mannosyl-oligosaccharide 1,2-alpha-mannosidase IA OS=Homo sapiens GN=MAN1A1 PE=1 SV=3 - [MA1A1_HUMAN]
Q9NR34	2.86	1	1	630	70.9	7.46	120.17	Mannosyl-oligosaccharide 1,2-alpha-mannosidase IC OS=Homo sapiens GN=MAN1C1 PE=1 SV=1 - [MA1C1_HUMAN]
P02795	19.67	1	1	61	6.0	7.83	71.21	Metallothionein-2 OS=Homo sapiens GN=MT2A PE=1 SV=1 - [MT2_HUMAN]
P08571	4.80	1	1	375	40.1	6.23	58.33	Monocyte differentiation antigen CD14 OS=Homo sapiens GN=CD14 PE=1 SV=2 - [CD14_HUMAN]
P15941	2.23	2	2	1255	122.0	7.47	84.01	Mucin-1 OS=Homo sapiens GN=MUC1 PE=1 SV=3 - [MUC1_HUMAN]
Q9HC84	3.58	5	5	5703	590.1	6.67	193.26	Mucin-5B OS=Homo sapiens GN=MUC5B PE=1 SV=2 - [MUC5B_HUMAN]
Q9H8L6	5.27	3	3	949	104.3	5.86	235.00	Multimerin-2 OS=Homo sapiens GN=MMRN2 PE=1 SV=2 - [MMRN2_HUMAN]
Q7Z7M0	5.69	11	10	2845	302.9	6.87	443.30	Multiple epidermal growth factor-like domains protein 8 OS=Homo sapiens GN=MEGF8 PE=1 SV=2 - [MEGF8_HUMAN]
O95865	7.37	1	1	285	29.6	6.01	74.60	N(G),N(G)-dimethylarginine dimethylaminohydrolase 2 OS=Homo sapiens GN=DDAH2 PE=1 SV=1 - [DDAH2_HUMAN]
O96009	16.90	4	4	420	45.4	6.61	209.00	Napsin-A OS=Homo sapiens GN=NAPSA PE=1 SV=1 - [NAPSA_HUMAN]
P08473	20.40	11	11	750	85.5	5.73	628.52	Neprilysin OS=Homo sapiens GN=MME PE=1 SV=2 - [NEP_HUMAN]
P59665	19.15	2	2	94	10.2	6.99	88.97	Neutrophil defensin 1 OS=Homo sapiens GN=DEFA1 PE=1 SV=1 - [DEF1_HUMAN]
Q5JS37	3.46	1	1	347	38.3	6.43	66.13	NHL repeat-containing protein 3 OS=Homo sapiens GN=NHLRC3 PE=2 SV=1 - [NHLC3_HUMAN]
P14543	1.68	1	1	1247	136.3	5.29	95.14	Nidogen-1 OS=Homo sapiens GN=NID1 PE=1 SV=3 - [NID1_HUMAN]
P61970	29.13	2	2	127	14.5	5.38	155.79	Nuclear transport factor 2 OS=Homo sapiens GN=NUTF2 PE=1 SV=1 - [NTF2_HUMAN]
Q6UX06	16.27	6	6	510	57.2	5.69	281.54	Olfactomedin-4 OS=Homo sapiens GN=OLFM4 PE=1 SV=1 - [OLFM4_HUMAN]
P78380	12.82	2	2	273	30.9	7.28	119.08	Oxidized low-density lipoprotein receptor 1 OS=Homo sapiens GN=OLR1 PE=1 SV=1 - [OLR1_HUMAN]

P00558	4.08	1	1	417	44.6	8.10	76.65	Phosphoglycerate kinase 1 OS=Homo sapiens GN=PGK1 PE=1 SV=3 - [PGK1_HUMAN]
O15162	6.29	1	1	318	35.0	4.94	68.00	Phospholipid scramblase 1 OS=Homo sapiens GN=PLSCR1 PE=1 SV=1 - [PLS1_HUMAN]
Q9BTY2	3.21	1	1	467	54.0	6.25	65.49	Plasma alpha-L-fucosidase OS=Homo sapiens GN=FUCA2 PE=1 SV=2 - [FUCO2_HUMAN]
P05155	17.20	7	7	500	55.1	6.55	431.29	Plasma protease C1 inhibitor OS=Homo sapiens GN=SERPING1 PE=1 SV=2 - [IC1_HUMAN]
P05154	42.12	28	15	406	45.7	9.26	1070.79	Plasma serine protease inhibitor OS=Homo sapiens GN=SERPINA5 PE=1 SV=2 - [IPSP_HUMAN]
P00747	3.83	1	1	810	90.5	7.24	92.68	Plasminogen OS=Homo sapiens GN=PLG PE=1 SV=2 - [PLMN_HUMAN]
P01833	29.97	23	17	764	83.2	5.74	1007.64	Polymeric immunoglobulin receptor OS=Homo sapiens GN=PIGR PE=1 SV=4 - [PIGR_HUMAN]
P0CG48	23.65	2	2	685	77.0	7.66	117.31	Polyubiquitin-C OS=Homo sapiens GN=UBC PE=1 SV=1 - [UBC_HUMAN]
Q9UHG3	3.96	2	2	505	56.6	6.18	104.25	Prenylcysteine oxidase 1 OS=Homo sapiens GN=PCYOX1 PE=1 SV=3 - [PCYOX_HUMAN]
P07602	10.31	7	4	524	58.1	5.17	219.10	Proactivator polypeptide OS=Homo sapiens GN=PSAP PE=1 SV=2 - [SAP_HUMAN]
Q9H3G5	24.16	11	9	476	54.1	5.62	421.42	Probable serine carboxypeptidase CPVL OS=Homo sapiens GN=CPVL PE=1 SV=2 - [CPVL_HUMAN]
P01133	41.92	66	35	1207	133.9	5.85	2644.30	Pro-epidermal growth factor OS=Homo sapiens GN=EGF PE=1 SV=2 - [EGF_HUMAN]
Q8WUM4	5.65	4	4	868	96.0	6.52	159.24	Programmed cell death 6-interacting protein OS=Homo sapiens GN=PDCD6IP PE=1 SV=1 - [PDC6I_HUMAN]
P12273	18.49	2	2	146	16.6	8.05	102.59	Prolactin-inducible protein OS=Homo sapiens GN=PIP PE=1 SV=1 - [PIP_HUMAN]
O43490	11.45	8	7	865	97.1	7.27	345.73	Prominin-1 OS=Homo sapiens GN=PROM1 PE=1 SV=1 - [PROM1_HUMAN]
Q14914	3.04	1	1	329	35.8	8.29	59.44	Prostaglandin reductase 1 OS=Homo sapiens GN=PTGR1 PE=1 SV=2 - [PTGR1_HUMAN]
P41222	36.32	5	4	190	21.0	7.80	265.73	Prostaglandin-H2 D-isomerase OS=Homo sapiens GN=PTGDS PE=1 SV=1 - [PTGDS_HUMAN]
O43653	23.58	3	3	123	12.9	5.29	170.58	Prostate stem cell antigen OS=Homo sapiens GN=PSCA PE=1 SV=1 - [PSCA_HUMAN]
P15309	2.59	1	1	386	44.5	6.24	73.09	Prostatic acid phosphatase OS=Homo sapiens GN=ACPP PE=1 SV=3 - [PPAP_HUMAN]
P02760	40.06	17	10	352	39.0	6.25	736.61	Protein AMBP OS=Homo sapiens GN=AMBP PE=1 SV=1 - [AMBP_HUMAN]
P31151	21.78	2	2	101	11.5	6.77	71.62	Protein S100-A7 OS=Homo sapiens GN=S100A7 PE=1 SV=4 - [S10A7_HUMAN]
P06702	11.40	2	1	114	13.2	6.13	159.64	Protein S100-A9 OS=Homo sapiens GN=S100A9 PE=1 SV=1 - [S10A9_HUMAN]
Q9C0H2	9.94	6	3	523	57.5	5.39	291.27	Protein tweety homolog 3 OS=Homo sapiens GN=TTYH3 PE=1 SV=3 - [TTYH3_HUMAN]

P00734	4.34	2	2	622	70.0	5.90	119.25	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2 - [THRB_HUMAN]
Q6V0I7	2.21	9	8	4981	542.4	4.94	407.50	Protocadherin Fat 4 OS=Homo sapiens GN=FAT4 PE=2 SV=2 - [FAT4_HUMAN]
Q9UN70	1.61	1	1	934	101.0	5.21	79.15	Protocadherin gamma-C3 OS=Homo sapiens GN=PCDHGC3 PE=1 SV=1 - [PCDGK_HUMAN]
Q495M3	5.80	2	2	483	53.2	8.12	101.50	Proton-coupled amino acid transporter 2 OS=Homo sapiens GN=SLC36A2 PE=2 SV=1 - [S36A2_HUMAN]
P00491	3.46	1	1	289	32.1	6.95	63.37	Purine nucleoside phosphorylase OS=Homo sapiens GN=PNP PE=1 SV=2 - [PNPH_HUMAN]
Q8NHP8	5.09	2	2	589	65.4	6.80	73.61	Putative phospholipase B-like 2 OS=Homo sapiens GN=PLBD2 PE=1 SV=2 - [PLBL2_HUMAN]
Q9H853	5.81	2	1	241	27.5	7.83	91.42	Putative tubulin-like protein alpha-4B OS=Homo sapiens GN=TUBA4B PE=5 SV=2 - [TBA4B_HUMAN]
P61026	16.50	3	3	200	22.5	8.38	94.69	Ras-related protein Rab-10 OS=Homo sapiens GN=RAB10 PE=1 SV=1 - [RAB10_HUMAN]
A6NIZ1	8.15	1	1	184	20.9	5.48	82.14	Ras-related protein Rap-1b-like protein OS=Homo sapiens PE=2 SV=1 - [RP1BL_HUMAN]
Q12913	2.32	2	2	1337	145.9	5.58	127.14	Receptor-type tyrosine-protein phosphatase eta OS=Homo sapiens GN=PTPRJ PE=1 SV=3 - [PTPRJ_HUMAN]
Q9HD89	51.85	8	4	108	11.4	6.86	376.55	Resistin OS=Homo sapiens GN=RETN PE=2 SV=1 - [RETN_HUMAN]
Q9HB40	5.53	5	2	452	50.8	5.81	323.19	Retinoid-inducible serine carboxypeptidase OS=Homo sapiens GN=SCPEP1 PE=1 SV=1 - [RISC_HUMAN]
P02753	20.40	2	2	201	23.0	6.07	164.73	Retinol-binding protein 4 OS=Homo sapiens GN=RBP4 PE=1 SV=3 - [RET4_HUMAN]
Q8WVN6	14.52	4	2	248	27.0	7.43	325.97	Secreted and transmembrane protein 1 OS=Homo sapiens GN=SECTM1 PE=1 SV=2 - [SCTM1_HUMAN]
Q13591	1.02	1	1	1074	120.5	7.21	63.70	Semaphorin-5A OS=Homo sapiens GN=SEMA5A PE=1 SV=3 - [SEM5A_HUMAN]
P34896	2.28	1	1	483	53.0	7.71	60.55	Serine hydroxymethyltransferase, cytosolic OS=Homo sapiens GN=SHMT1 PE=1 SV=1 - [GLYC_HUMAN]
Q96SA4	8.55	2	2	456	50.8	6.19	93.68	Serine incorporator 2 OS=Homo sapiens GN=SERINC2 PE=2 SV=2 - [SERC2_HUMAN]
P02787	37.97	34	17	698	77.0	7.12	1585.18	Serotransferrin OS=Homo sapiens GN=TF PE=1 SV=2 - [TRFE_HUMAN]
Q96P63	4.20	1	1	405	46.2	5.53	77.37	Serpin B12 OS=Homo sapiens GN=SERPINB12 PE=1 SV=1 - [SPB12_HUMAN]
P29508	12.56	4	4	390	44.5	6.81	135.85	Serpin B3 OS=Homo sapiens GN=SERPINB3 PE=1 SV=2 - [SPB3_HUMAN]
P02768	50.08	48	26	609	69.3	6.28	1979.24	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2 - [ALBU_HUMAN]
Q9HAT2	2.29	1	1	523	58.3	7.33	85.37	Sialate O-acetylesterase OS=Homo sapiens GN=SIAE PE=1 SV=1 - [SIAE_HUMAN]
P31639	2.23	1	1	672	72.8	7.47	62.08	Sodium/glucose cotransporter 2 OS=Homo sapiens GN=SLC5A2 PE=1 SV=1 - [SC5A2_HUMAN]

Q13621	4.09	3	3	1099	121.4	7.39	130.61	Solute carrier family 12 member 1 OS=Homo sapiens GN=SLC12A1 PE=1 SV=2 - [S12A1_HUMAN]
P22732	5.39	3	3	501	54.9	6.04	140.45	Solute carrier family 2, facilitated glucose transporter member 5 OS=Homo sapiens GN=SLC2A5 PE=1 SV=1 - [GTR5_HUMAN]
Q4U2R8	2.13	1	1	563	61.8	8.76	74.41	Solute carrier family 22 member 6 OS=Homo sapiens GN=SLC22A6 PE=1 SV=1 - [S22A6_HUMAN]
Q14515	3.16	1	1	664	75.2	4.81	85.54	SPARC-like protein 1 OS=Homo sapiens GN=SPARCL1 PE=1 SV=2 - [SPRL1_HUMAN]
P17405	1.75	1	1	629	69.7	7.28	69.10	Sphingomyelin phosphodiesterase OS=Homo sapiens GN=SMPD1 PE=1 SV=4 - [ASM_HUMAN]
P52823	9.31	2	2	247	27.6	7.99	111.17	Stanniocalcin-1 OS=Homo sapiens GN=STC1 PE=1 SV=1 - [STC1_HUMAN]
Q9UGT4	3.04	2	2	822	90.1	6.28	131.08	Sushi domain-containing protein 2 OS=Homo sapiens GN=SUSD2 PE=1 SV=1 - [SUSD2_HUMAN]
O00560	17.79	4	3	298	32.4	7.53	219.36	Syntenin-1 OS=Homo sapiens GN=SDCBP PE=1 SV=1 - [SDCB1_HUMAN]
Q8TB96	2.29	1	1	612	68.1	5.39	64.06	T-cell immunomodulatory protein OS=Homo sapiens GN=ITFG1 PE=1 SV=1 - [TIP_HUMAN]
O60635	5.39	2	1	241	26.3	5.25	78.20	Tetraspanin-1 OS=Homo sapiens GN=TSPAN1 PE=1 SV=2 - [TSN1_HUMAN]
O43657	5.31	1	1	245	27.5	8.10	95.09	Tetraspanin-6 OS=Homo sapiens GN=TSPAN6 PE=1 SV=1 - [TSN6_HUMAN]
P07996	0.94	1	1	1170	129.3	4.94	101.45	Thrombospondin-1 OS=Homo sapiens GN=THBS1 PE=1 SV=2 - [TSP1_HUMAN]
Q9UKU6	14.06	12	12	1024	116.9	6.99	464.27	Thyrotropin-releasing hormone-degrading ectoenzyme OS=Homo sapiens GN=TRHDE PE=2 SV=1 - [TRHDE_HUMAN]
P05543	15.66	5	5	415	46.3	6.30	289.98	Thyroxine-binding globulin OS=Homo sapiens GN=SERPINA7 PE=1 SV=2 - [THBG_HUMAN]
P61586	13.47	3	2	193	21.8	6.10	204.58	Transforming protein RhoA OS=Homo sapiens GN=RHOA PE=1 SV=1 - [RHOA_HUMAN]
Q14956	3.15	1	1	572	63.9	6.64	71.63	Transmembrane glycoprotein NMB OS=Homo sapiens GN=GPNMB PE=1 SV=2 - [GPNMB_HUMAN]
O43280	1.89	1	1	583	66.5	5.68	60.93	Trehalase OS=Homo sapiens GN=TREH PE=1 SV=2 - [TREA_HUMAN]
O14773	14.21	4	4	563	61.2	6.48	246.57	Tripeptidyl-peptidase 1 OS=Homo sapiens GN=TPP1 PE=1 SV=2 - [TPP1_HUMAN]
Q9H1C7	10.31	1	1	97	10.6	4.32	67.31	UPF0467 protein C5orf32 OS=Homo sapiens GN=C5orf32 PE=2 SV=1 - [CE032_HUMAN]
P00749	16.94	7	7	431	48.5	8.41	315.08	Urokinase-type plasminogen activator OS=Homo sapiens GN=PLAU PE=1 SV=2 - [UROK_HUMAN]
P07911	30.47	21	14	640	69.7	5.24	964.49	Uromodulin OS=Homo sapiens GN=UMOD PE=1 SV=1 - [UROM_HUMAN]
Q6EMK4	12.33	6	5	673	71.7	7.39	257.62	Vasorin OS=Homo sapiens GN=VASN PE=1 SV=1 - [VASN_HUMAN]
Q12907	19.66	5	5	356	40.2	6.95	258.74	Vesicular integral-membrane protein VIP36 OS=Homo sapiens GN=LMAN2 PE=1 SV=1 - [LMAN2_HUMAN]

P22891	8.25	2	2	400	44.7	5.97	79.33	Vitamin K-dependent protein Z OS=Homo sapiens GN=PROZ PE=1 SV=2 - [PROZ_HUMAN]
P04004	10.67	3	3	478	54.3	5.80	118.49	Vitronectin OS=Homo sapiens GN=VTN PE=1 SV=1 - [VTNC_HUMAN]
P21796	23.67	5	5	283	30.8	8.54	214.46	Voltage-dependent anion-selective channel protein 1 OS=Homo sapiens GN=VDAC1 PE=1 SV=2 - [VDAC1_HUMAN]
P38606	2.92	1	1	617	68.3	5.52	96.78	V-type proton ATPase catalytic subunit A OS=Homo sapiens GN=ATP6V1A PE=1 SV=2 - [VATA_HUMAN]
Q9UI12	3.31	1	1	483	55.8	6.48	70.12	V-type proton ATPase subunit H OS=Homo sapiens GN=ATP6V1H PE=1 SV=1 - [VATH_HUMAN]
O43895	13.20	6	6	674	75.6	6.04	241.89	Xaa-Pro aminopeptidase 2 OS=Homo sapiens GN=XPNPEP2 PE=1 SV=3 - [XPP2_HUMAN]
Q96DA0	25.96	4	4	208	22.7	7.39	223.46	Zymogen granule protein 16 homolog B OS=Homo sapiens GN=ZG16B PE=1 SV=3 - [ZG16B_HUMAN]

CHAPTER 5

GLYCOPROTEOMICS OF URINARY MEMBRANE VESICLES AND NOVEL METHOD FOR ENRICHMENT OF MEMBRANE VESICLES FROM MINIMALLY PROCESSED URINE

5.1 Introduction

The urinary proteome may serve as a rich source of biomarkers for uro-genital and systemic diseases which have been reviewed previously (Pisitkun, Johnstone & Knepper, 2006). Notably, urine collection is a non-invasive procedure which makes it an ideal source for biomarker discovery. Due to their complex structure and functions, the analysis of glycans of cellular glycoproteins and glycolipids is rapidly emerging as an important aspect of novel biomarkers either directly from tissue/cellular samples or as analyzed from body fluids (Nilsson *et al.,* 2009; Halim *et al.,* 2012). However, up to now, there have been only a few large-scale studies on urine glycoproteome analysis (Adachi *et al.,* 2006; Wang *et al.,* 2006). For this purpose, the membrane vesicles, including exosomes as secreted by many cell types along the uro genital tract are attractive sources of glycoanalytics. Notably, the exosomes display functions like antigen-presentation, cell to cell communication and immuno-modulation (Nieuwland & Sturk, 2010). These are specialised compartment of cells and not only do they mirror the physiological state of cells secreting them but also provide information about the environment into which they are secreted; i.e., immunosuppressive and pro-angiogenic environment of cancer may be mediated in part by exosomes (Nieuwland & Sturk, 2010).

Exosomes and other types of membrane vesicles have been discovered in urine and known to be secreted by epithelial cells lining the genitourinary system (Pisitkun, Shen & Knepper, 2004). Glycosylation is an important modality of post-translational modification which plays many roles including but not limited to cell adhesion, cell to cell communication, immune response etc (Bhatia & Mukhopadhyay, 1999). Glycosylation has also been shown to be very important for targeting of proteins to various compartments inside the cells. Glycans (the complex carbohydrate chains linked to proteins or lipids) present on glycoproteins are

important in sorting these proteins to membrane microdomains and have been shown to influence their intracellular trafficking (Huet *et al.,* 2003). Interestingly, it has been recently found that membrane vesicles may have a glycan signature which is distinct from the parent cell itself, suggesting that they originate from specialised membrane microdomains. This implies a role of glycosylation in membrane vesicle protein sorting (Batista *et al.,* 2011; Escrevente *et al.,* 2011). For example, HIV-1 particles were found to have a glycome common to membrane vesicles which was cell-specific implying that virus hijacks the machinery of infected cells and uses it to infect other cells (Krishnamoorthy *et al.,* 2009).

Exosome uptake in various cell types was shown to occur through clathrin-mediated endocytosis, phagocytosis and macropinocytosis (Hao *et al.,* 2007; Barres *et al.,* 2010). The uptake of exosomes, particularly by dendritic cells and macrophages, was shown to be inhibited by mannose, N-acetylglucosamine and lactose, respectively. An important mediator of this uptake turned out to be a C-type lectin in dendritic cells and galectin-5 in macrophages (Hao *et al.,* 2007; Barres *et al.,* 2010). All this points towards a system, in which, the exosomal glycosylation pattern is maintained as a specific feature of the cells or sub-cellular organelles secreting them. Consequently, the tagged exosomes appear fitted for a specific target cell uptake pathway and distinct downstream functions. Taken together, these finding suggest that better understanding of surface glycosylation patterns as well as the whole glycoproteome of exosome may help define the biology of exosome uptake by target cells e.g. the pathways involved. Glycoproteome of urinary membrane vesicles, therefore, will provide information about both the functional state of constituent proteins and also highlight the special signatures of proteins which are specifically targeted to membrane vesicles. On the other hand it appears that lectins provide useful tools for glycan profiling from complex samples and have been applied in many assay and array formats for dynamic glycan profiling (Pilcibello, Slawek & Mahal, 2007; Hsu & Mahal, 2006).

Here we have devised a fluorophore-linked lectin assay (FLLA) for glycan profiling of purified membrane vesicles and other fractions from urine obtained using the traditional differential centrifugation method. Eighteen lectins, falling into 7 broad classes based on their nominal glycan binding specificities, were chosen for this study. All these lectins were further used for verification in lectin blots of membrane vesicles and other urinary fractions to reveal consitituent glycoproteins bearing the respective nominal glycans. Based on the analysis of these results we were able to use lectin affinity chromatography, to enrich urinary membrane vesicles directly from whole urine. This is a novel application area of lectins which is simple and avoids the use of specialised techniques like ultracentrifugation and setting up density gradients. Furthermore, to gain functional information about classes of proteins sorted to distinct membrane vesicle classes and to broaden the membrane vesicle glycoproteome coverage, we also perfomed multiple lectin-affinity chromatography (LAC) using 9 different lectins from 7 different nominal specificity classes of lectins on membrane vesicle extracts and proceeded to subsequent identification of these proteins. To complement the LAC method, hydrazide chemistry was also applied to P18,000g and P200,000g pellets to enrich and subsequently identify the glycoproteins. Two hundred and eighty-two proteins in total were identified and classified according to their sub-cellular localization and functions.

5.2 Material and methods

5.2.1 Preparation of nano vesicle and other fractions of urine

Vesicle preparation has been described in detail in methods section of Chapter 2. A schematic representation of the methodology used to isolated vesicles is shown in Figure 5.1. P200,000g is a heterogeneous mixture of different types of vesicles and the method developed in Chapter 2 using CHAPS treatment might select for specific populations of vesicles. DTT and CHAPS treatment methods are complementary which is also evident by the large overlap between DTT and CHAPS SN200,000g (76%, Chapter 2). However, activity of proteins is better preserved in the CHAPS treatment method (Figure 2.5, chapter 2). But here, the aim is not to perform activity based proteomics therefore any of these two methods can be used. We have used DTT-treated pellet (P18DTTP18 and P200DTTP200) for lectin affinity chromatography and crude pellet (P18,000g and P200,000g) for FLLA and hydrazide enrichment.

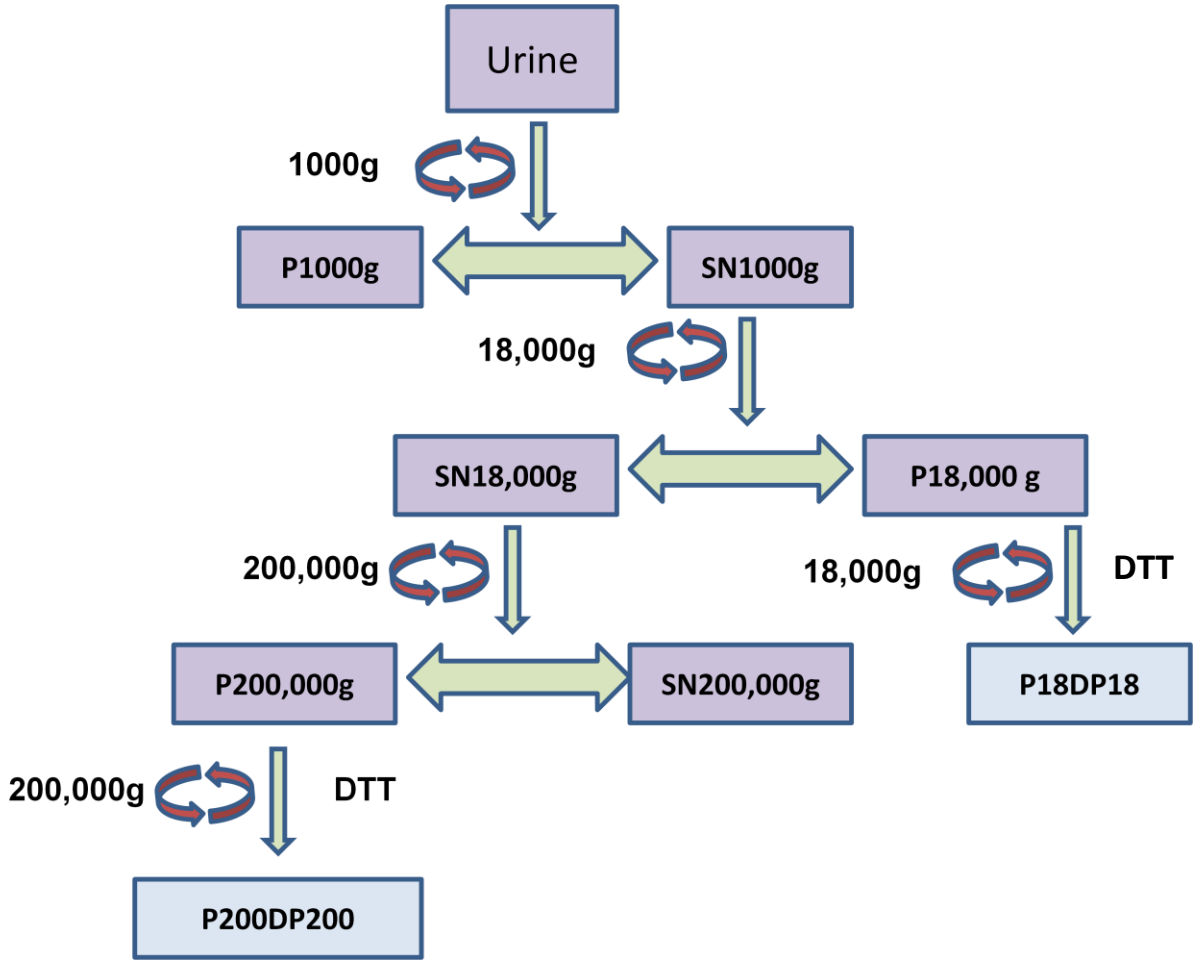

Figure 5.1: Schematic of vesicle purification from urine. Box highlighted in blue were used for lectin affinity chromatography. P18,000g, P200,000g crude and SN200,000g (Purple box) were used for FLLA and hydrazide enrichment.

5.2.2 THP Purification

THP was purified as described in methods section of Chapter 2.

5.2.3 Protein quantification, SDS-PAGE and Western Blotting

It was carried out as described in methods section of Chapter 2 (Musante *et al.*, 2012).

5.2.4 Fluorophore-linked lectin assay (FLLA)

Microtitre plate wells were coated overnight with 100µl/well of 100µg/mL different urine fractions at +37°C. All the subsequent steps were carried out at room temperature. Wells

were washed with PBS (pH 7.4) five times between each step. Wells were blocked with odyssey blocking buffer (LI-COR biosciences, Lincoln, NE) (diluted 1:1 with PBS) for 1 hour on agitation at room temperature. Biotinylated lectins (Vector laboratories, Burlingame, CA) were applied (100µl of 1:2000 dilutions from stocks of 2mg/mL) to the wells and incubated for 1 hour on agitation. Streptavidin conjugated to IRDye-800 (100µl of 1:5000 dilutions of company stock) was added to the wells and incubated on agitation for 1 hour. Finally wells were washed five times with PBS (pH 7.4) and image was acquired on Odyssey Infrared Laser Scanner (LI-COR Biosciences, Lincoln, NE) in 800nm channel at highest intensity. Intensities of fluorescence of wells were calculated using the Odyssey V3.0 software and imported into a text file. For the control (streptavidin fractions control) triplicate measurements were taken in which biotinylated lectins were substituted with PBS while rest of the procedure remained same. To be more stringent, additional controls were taken; (1) empty wells were blocked and lectin and streptavidin were added subsequently as per the protocol described (lectin blocking control) (2) only streptavidin was added to blocked wells (streptavidin blocking control). Streptavidin fraction control was subtracted from the final intensity of different fractions and streptavidin blocking control was subtracted from the lectin blocking control. This gave the values of lectin non-specific adsorption on blocking and lectin specific binding to different urine fractions. Finally lectin non-specific adsorption was subtracted from different fraction intensity to get specific lectin binding to different fractions.

5.2.5 Lectin blotting

SDS-PAGE was carried out as described in Chapter 2. Proteins were transferred to a nitrocellulose membrane using iBlot (Invitrogen, Life technologies, NY). Membranes were blocked for 1 hour at RT with the Odyssey blocking buffer solution (LI-COR Biosciences). All incubation steps with biotinylated lectins (1:2000, 1 hour) were performed in a 1:1 (v/v)

mixture of Odyssey blocking buffer and PBS with 0.1 % (v/v) Tween 20 (PBST). Blots were washed with PBST 6 times 10 minutes each. Finally streptavidin conjugated to IRDye-800 was prepared (1:5000) and incubated with blots for 1 hour and subsequently blots were washed 5 times with PBST and once with PBS. Images were acquired with Odyssey Infrared Laser Scanner (LI-COR Biosciences).

5.2.6 Hydrazide chemistry for enriching surface and other glycoproteins of P18,000g and P200,000g

A hydrazide chemistry-based approach was used to enrich glycoproteins from P18,000g and P200,000g (Crude pellets) as described previously (McDonald *et al.,* 2009). Briefly, P18,000g and P200,000g was suspended in 50mM Tris buffer, pH 7.4 and incubated with 10mM sodium metaperiodate for 1 hour at room temperature on rotation. Immediately after, the mixture was incubated with hydrazide resin (Sigma, St. Louis, MO) overnight at room temperature on rotation. Non-bound material was removed by washing with Tris buffer (pH 7.4). Here, the membrane vesicles in these two pellets were bound to the resin through surface glycoproteins. The resin was incubated with 1% (w/v) beta-octyl glucoside to lyse the vesicles. Non-bound material recovered was re-oxidised with 10mM sodium metaperiodate and incubated with hydrazide resin as described. These two resins thus generated will have surface glycoproteins bound to the resin in one and all other glycoproteins bound to the other. Then both resins were processed simultaneously. The resin was washed with 8M urea in Tris buffer (TU buffer, pH 7.4) to remove the non-bound material. The resins were then incubated with 50mM DTT in 50mM Tris buffer (pH7.4) for one hour at 37°C. DTT was removed by washing with TU buffer and resin was incubated with 65mM iodoacetamide in dark for 30 minutes. The resin was then washed with TU buffer and 1.5M NaCl in Tris buffer (pH 7.4) three alternate times to remove all the non-glycoproteins from the resin. The immobilised proteins were digested ‘on-column’ with 40ng/µL of sequencing grade trypsin (Promega,

Madison, WI) overnight at 37°C. The tryptic peptides were recovered the next day and enriched by Sep-Pak solid phase extraction columns (Waters, Milford, MA) as described in methods sections of Chapter 2. These enriched peptides were analysed by LC-MS/MS as described ahead.

5.2.7 Lectin affinity chromatography

Eight biotinylated lectins (Con-A, LCA, UEA, SNA, MAL-II, PHA-E, WGA, RCA120 and Jacalin, from Vector laboratories; refer to Table 5.1 for full names of these lectins) were immobilised on streptavidin agarose. Lectins (500µg each in PBS) were incubated with the agarose slurry overnight at +4°C on rotation. Next day the non-bound lectins were removed by washing with PBS. RCA120 already coupled to agarose was bought commercially (Vector laboratories). Six hundred microgram of P200,000g and P18,000g (DTT treated, see blue boxes in Figure 5.1) extract (prepared in 1% beta-octyl glucoside) was diluted in HEPES buffered-saline (100mM HEPES + .15M NaCl) to a 2mL final volume and incubated overnight with lectin-agarose at +4°C on rotation. For Con-A, 50mM tris buffer with .5M NaCl was used as a binding buffer. Some of these lectins (Con-A, LCA, PHA-E) require calcium and manganese ions for their binding ability to glycoproteins and 1mM of these divalent cations were added to the binding buffer of respective lectins. These cations were not used in FLLA and lectin blotting because phosphate in PBS might precipitate with calcium as calcium phosphate. Next day, the non-bound fraction was collected and lectin-agarose was incubated with respective elution buffers for 2 hours at +4°C on rotation and elution was recovered. Elution buffers were made by adding inhibiting/eluting sugars to binding buffers: 200mM galactose + 200mM lactose for RCA120, 200mM α-methyl mannoside + 200mM α-methyl glucoside for Con-A and LCA, 800mM galactose for Jacalin, 100mM L-fucose for UEA and 200mM lactose for MAL-II. SNA-agarose-bound proteins were eluted by 500mM lactose in binding buffer followed by 500mM lactose in 100mM acetic acid. PHA-E-bound

proteins were eluted by 100mM acetic acid. WGA-bound glycoproteins were eluted with 500mM N-acetylglucosamine. All the eluates were dialysed against water for further processing required for MS analysis.

5.2.8 Jacalin affinity chromatography for isoltion of membrane vesicles from minimally processed urine.

Five hundred microgram of biotinylated Jacalin was incubated with 1mL of streptavidin-agarose (SIGMA) overnight in PBS buffer on rotation. Non-bound material was removed the next day and newly generated Jacalin agarose was incubated with 4mL (50mL of SN2000g concentrated to 2mL with Vivaspin 'cut-off' 5 or 300 kDa (Sartorius) and diluted to 4mL with either 200mM HEPES buffer saline {HBS, 2X} or only 20mM phosphate buffer, pH 7.4) concentrated urine. After overnight incubation at +4°C on rotation, the non-bound fraction was removed and the Jacalin-agarose was washed with 20 column volumes (20mL) of 1X HEPES buffer saline or 10mM phosphate buffer. The bound fraction was eluted by incubating the resin with 800mM galactose in HBS or phosphate buffer for 2 hours.

5.2.9 MS analysis of proteins and database searching

Sample preparation for MS analysis was carried out as described in Chapter 3. Nano LC–MS/MS analysis was carried out using an Ultimate 3000 nanoLC system (Dionex, Sunnyvale, CA) coupled to a hybrid linear ion trap/Orbitrap mass spectrometer (LTQ Orbitrap XL; Thermo Fisher Scientific, Somerset, NJ). Five microlitres of digest were loaded onto a C18 trap column (C18 PepMap, 300μm ID × 5mm, 5μm particle size, 100Å pore size; Dionex) and desalted for 10 min using a flow rate of 25μL/min in 0.1% TFA. The trap column was then switched online with the analytical column (PepMap C18, 75μm ID × 250mm, 3μm particle and 100Å pore size; (Dionex)) and peptides were eluted with the following binary gradients of solvent A and B: 0–25% solvent B in 120min and 25–50% solvent B in a further 60min, where solvent A consisted of 2% acetonitrile (ACN) and 0.1% (v/v) formic acid in

water and solvent B consisted of 80% (v/v in water) ACN and 0.08% (v/v in water) formic acid. Column flow rate was set to 350nL/min.

Data were acquired with Xcalibur software, version 2.0.7 (Thermo Fisher Scientific). The mass spectrometer was operated in data-dependent mode and externally calibrated. Survey MS scans were acquired in the Orbitrap in the 400–1800 m/z range with the resolution set to a value of 60,000 at m/z 400. Lock mass was set at 445.120025u (protonated (Si (CH3)2O)6). Up to seven of the most intense ions (1+, 2+ and 3+) per scan were CID fragmented in the linear ion trap. A dynamic exclusion window was applied within 40s. All tandem mass spectra were collected using normalised collision energy of 35%, an isolation window of 3 m/z, and one microscan.

Proteins were identified using BioWorks 3.2 from Thermo Fisher Scientific using the HUPO criteria with XC scores of 1.8, 2.2, 3.75 for single, double and triple charged ions. A peptide probability score of 0.05 was also used. The database used was Human UniProt-SwissProt downloaded January 2012. Carboxymethylation of cysteine was set as fixed and oxidation of methionine as a variable modification. Two missed cleavages were allowed. The mass tolerance for precursor ions was 20ppm and the mass tolerance for fragment ions was 0.5Da. For some sample's spectra, that was searched using Mascot the protein score 'cut-off' was kept at 40.

5.2.10 Transmission electron microscopy

TEM was performed as described in the methods section of Chapter 2 (Musante *et al.*, 2012).

5.2.11 Bioinformatic analysis

Bioinformatic analysis was performed as described in Chapter 3. Additionally I2D (http://ophid.utoronto.ca/ophidv2.201/ppi.jsp) and iRefWeb (Turner *et al.*, 2010) databases were used to search for protein-protein interactions of proteins present in the dataset. SOSUI

(Hirokawa, Boon-Chieng & Mitaku, 1998) was used for classification of proteins according to the presence of transmembrane domains.

5.3 Results

5.3.1 Glycan profile of urinary membrane vesicles surface and other fraction of urine

FLLA was performed on isolated nano vesicle fraction (P200,000g), low speed pellet (P18,000g), 'nano vesicle-free' urine (SN 200,000g) and purified Tamm-Horsfall glycoprotein (THP). FLLA was performed with crude fractions because we wanted to develop a method with least number of steps and the ultimate aim is to develop a method which can be performed rapidly on minimally-processed urine. A list of all the lectins used in this study in different techniques can be found in Table 5.1. These lectins were specifically chosen because they cover the widest range of glycan binding specificities. Therefore, they are suited to profile the surface glycans of membrane vesicles.

Table 5.1: Lectins used in FLLA and their binding specificities along with their source organisms. Binding of lectin to various fractions of urine in FLLA is given (Please refer to Figure 5.2).

Abbreviations	**Origin**	**Nominal sugar binding specificity** (Gabius & Gabius, 1997; Liener, Sharon & Goldstein, 1986**;** Debray *et al.*, 1981)	**P18**	**P200**	**SN200**	**THP**
LCA/LcH	*Lens culinaris*	Fucα1-6GlcNAc ; α-Man ; α-Glc	-	-	-	-
PSA	*Pisum sativum*	Fucα1-6GlcNAc ; α-Man	-	+	-	-
UEA-I	*Ulex europeus*	Fucα1-2LacNAc ; α-Fuc	-	-	-	-
SNA	*Sambucus nigra*	Siaα2-6Gal ; GalNAc	+++	+++	-	+
MAL/ MAA I	*Maakia amurensis*	Galβ1-4GlcNAc	++	+	-	+
MAL II		Siaα2-3Gal	-	-	-	-
RCA[120]	*Ricinus communis*	Lac ; LacNAc	+++	++++++	+++	+
PHA (L)	*Phaseolus vulgaris*	Tetraantennary complex oligosaccharides	++	++	-	+
PHA (E)	*Phaseolus vulgaris*	NA2 ; bisecting GlcNAc	++	++++	-	++
ConA	*Canavalia ensiformis*	α-Man branched and terminal; terminal GlcNAc	+++	+++	+	+
WGA	*Triticum vulgaris*	(GlcNAc)n ; multivalent Sia ; β-GlcNAc, GalNAc	++	++	-	+
succ WGA		GlcNAc	-	-	-	-
PNA	*Arachis hypogeal*	Galβ1-3GalNAcα-Thr/Ser (T) ; terminal Galβ-OR	-	-	-	-

JAC/ Jacalin	*Artocarpus integrifolia*	Galβ1-3GalNAcα-Thr/Ser (T) ; GalNAcα-Thr/Ser (Tn)	+++	+++	+	-
GSL-I/ GSI	*Griffonia smplicifolia*	α-GalNAc ; GalNAcα-Thr/Ser (Tn) ; α –Gal	+	-	-	+
DBA	*Dolichos biflorus*	GalNAc α -Thr/Ser (Tn) ; GalNAc α 1-3GalNAc	++	+	-	+
SBA	*Glycine max*	Terminal GalNAc (especially GalNAc α 1-3Gal)	+	+	-	-
SJA	*Sophora japonica*	β-GalNac ; β-Gal	-	-	-	-

THP is the most abundant protein in normal human urine. As THP is present in all fractions of urine (Figure 2.3 from Chapter 2) and consists of a wide variety of glycans (Wu *et al.*, 2008) which make up 30% of total mass, we included THP in our FLLA study as a positive control. The same concentration of all the fractions was used in the study. The membrane vesicle fraction (P200,000g) consists of many different proteins (and potentially glycoproteins) in addition to THP (Panel A Figure 2.3 in Chapter 2) so FLLA signal in membrane vesicle fractions cannot be directly correlated and concluded to be because of THP. P18,000g as judged by SDS-PAGE (Panel A Figure 2.3 in Chapter 2) consisting more amount of THP, so any lectin binding to low speed pellet more than the purified THP fraction was considered to be due to other constituents. Multiple stringent process controls were also taken other than the normal background control as described in the methods section. Surface glycan profile of urinary membrane vesicles can be seen in Figure 5.2.

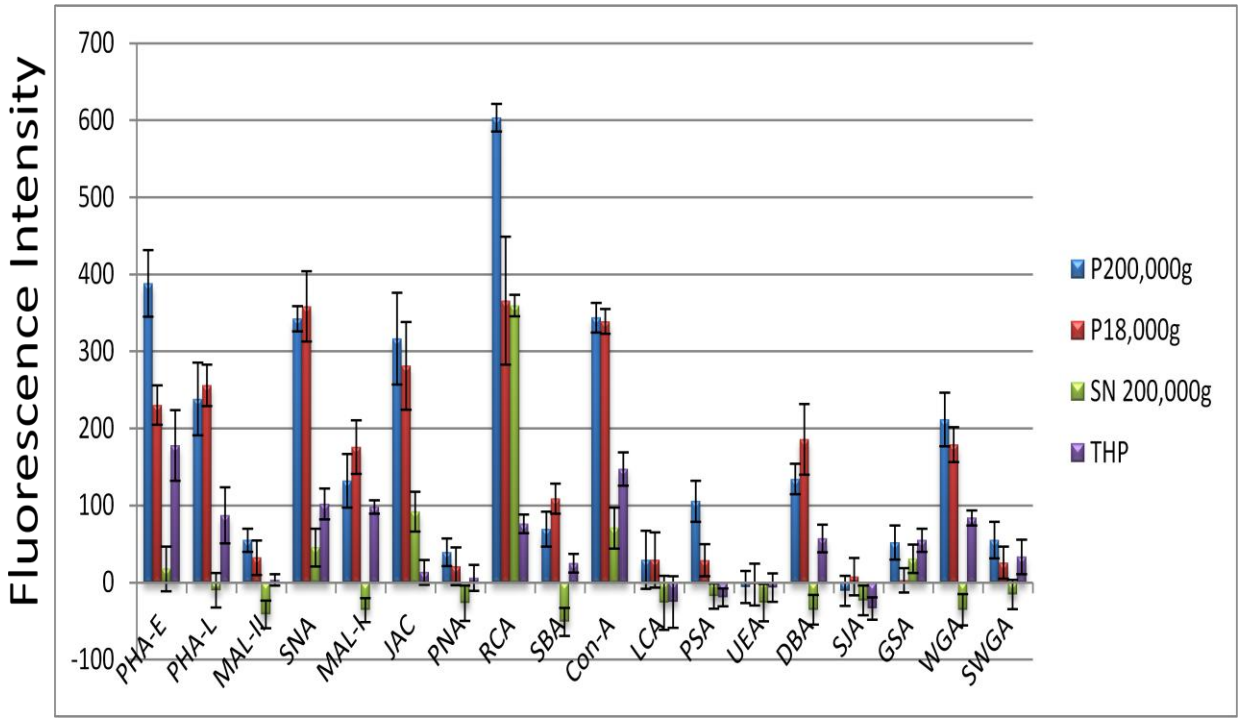

Figure 5.2: Fluorophore linked lectin assay (FLLA) was employed on membrane vesicles and other urine fractions. Fluorescence signal with standard deviation (based on triplicate measurements) is provided in the Figure. For every lectin, the bars from left to right are P200,000g, P18,000g, SN200,000g and purified THP which are also indicated by colour code in the legends on the right side of the figure.

Moreover, to confirm that signal generated from FLLA was due to the lectin-glycan interaction, four lectins were selected and the same procedure was repeated with the difference that inhibiting sugars of the respective lectins were added to the lectin solution. No significant signal above background was expected from this assay and results confirmed the specificity of the lectin binding (Figure 5.3).

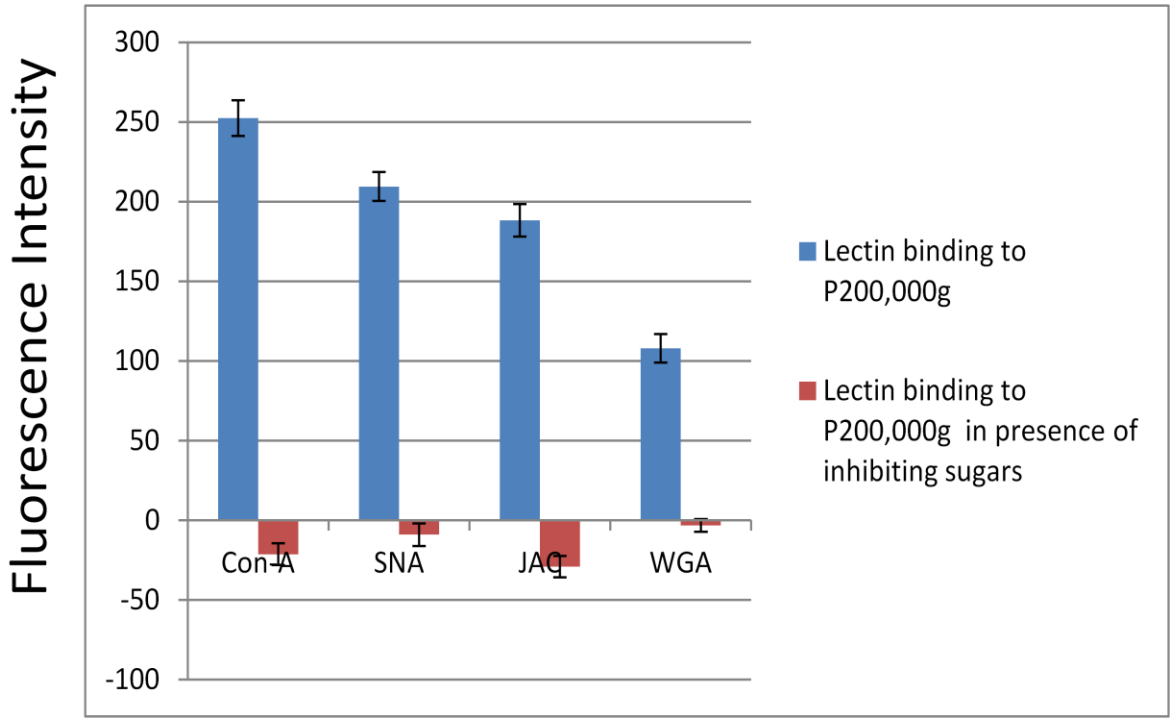

Figure 5.3: FLLA signal with P200,000g in presence and absence of inhibiting sugars (Con-A: 200 mM α-methyl mannoside, SNA: 500 mM lactose, JAC: 500 mM galactose, WGA: 500 mM N-acetylglucosamine) to demonstrate specificity of lectin binding to sugars in various fractions.

Some of the lectins having highest signal in the nano vesicle fraction representing surface glycan profile are indicated in Table 5.1. Briefly, tetrantennary complex oligosaccharides (PHA-L), bisecting GlcNAc complex structures (PHA-E), Sia α 2-6Gal ; GalNAc (SNA), Lac ; LacNAc (RCA120), α-Man branched and terminal; terminal GlcNAc (Con-A), (GlcNAc)n ; multivalent Sia ; β-GlcNAc, GalNAc (WGA), Gal β 1-3GalNAca-Thr/Ser (T) ; GalNAca-Thr/Ser (Tn) (JAC), GalNAcaThr/Ser (Tn) ; GalNAca1-3GalNAc (DBA) were enriched in the nano vesicles fraction. On the other hand, MAL-II, GSA, SWGA, SBA and PNA bound to much lesser degree to this fraction while SJA and UEA did not show any binding.

5.3.2 Lectin blotting

All the 18 lectins used in FLLA were also used in lectin blotting to reveal more information about constituent proteins of all these fractions. Only some of the lectins showing major binding to either membrane vesicle fraction (P200,000g) or P18,000g or SN 200,000g are shown in Figure 4. Lectin blotting results (Figure 5.4) are categorized according to the binding specificities of lectins.

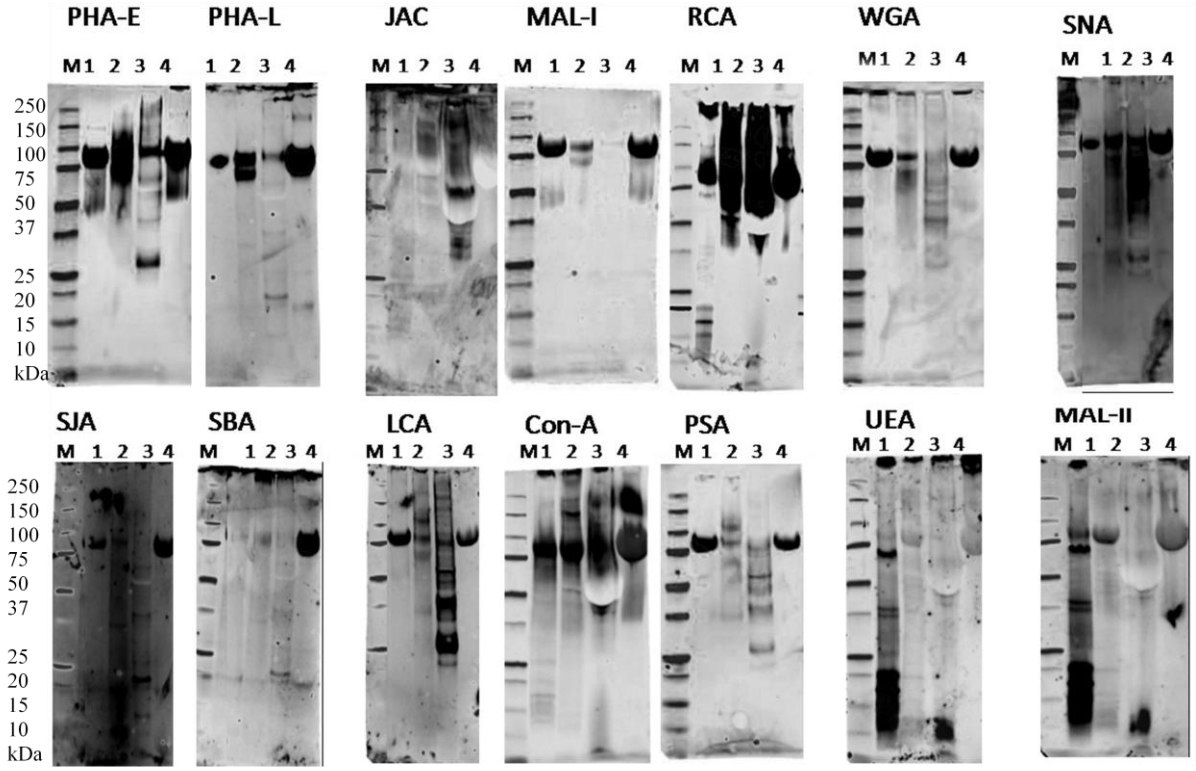

Figure 5.4: Lectin blloting using 14 lectins is shown. Lectins with shared binding specificity are grouped together horizontally except for MAL-II and SNA which are grouped together vertically. M: molecular weight markers; 1: P18,000g; 2: P200,000g; 3: SN 200,000g; 4: Purified THP.

As THP is present in all of the fractions and it is recognised by most lectins, any band of ~80kDa is considered to be due to THP. Notably, this is only an assumption and lectin blotting alone does not provide any proof for it and there may be more than one protein present in the same band. From top left, first panel comprising two lectins (PHA-E, PHA-L)

will display glycoproteins with tetrantennary complex oligosaccharides (PHA-L) and bisecting GlcNAc complex structures (PHA-E). PHA-E binding proteins in P200,000g are mainly present as a smear from 60kDa to 120kDa while in P18,000g the main band is that of THP (which can be compared to purified THP; Lane 4). In P18,000g, apart from THP, one weak band at 150kDa can be easily seen while in SN200,000g bands of 28, 40 and 60 kDa can be found apart from a usual band at molecular weight corresponding to THP. In blot of PHA-L there is a clear band of THP in P200,000g and equally intense band of 70kDa of unknown identity. In P18,000g there is no other band than that of THP while, in SN200,000g, bands of 20, 35 and 60kDa can be seen. This result can be compared to FLLA signal (Figure5.2) where intense signal in P200,000g can be seen for PHA-E and PHA-L while less so in P18,000g. No signal in FLLA is present for SN200,000 for PHA-L which is similar in lectin blotting. However, there are some bands present in lectin blot of SN200,000g of PHA-E while signal in FLLA for this fraction is absent (PHA-E in Fiigure 5.2). This can be explained if these proteins seen in lectin blot are present in the vesicle lumen and not on surface.

The second panel of three lectins (JAC, MAL-I and RCA120) show nominal specificity to galactose bearing structures although their binding specificities can be markedly different. All the precise binding specificity of lectins used in this study can be compared in Table 5.1. Jacalin showed binding to three proteins between 100kDa and 160kDa and three proteins between 50kDa and 75 kDa in P200,000g while in P18,000g showed a smear between 15kDa and 37kDa. In SN200,000g many bands were visible between 28 to 40kDa, one major band at 55kDa and a smear between 75 to 200kDa while there was no binding with THP (lane 4). Signal in FLLA for Jacalin is similar to lectin blotting for purified THP (No signal in both the techniques). P18,000g and P200,000, show good signal in FLLA for Jacalin while less so in lectin blotting. The converse is true for SN200,000g which shows some bands in lectin

blotting while FLLA signal is very low. MAL-I mainly showed two bands of 70kDa and 80kDa in P200,000g which can be directly compared to bands seen in PHA-L. In the SN200,000g, the usual THP band can be seen apart from the two very weak bands of 40kDa and 22 kDa. In the case of RCA^{120} with nominal glycan specificity to terminal galactose and N-acetylgalactosamine (similar to JAC), there is a huge smear of signal from 37kDa to 250kDa and above in both P200,000g and SN200,000g suggesting that majority of glycoproteins in urine in this range have this type of glycosylation. This is directly comparable with FLLA signal for RCA^{120}. In P18,000g many bands between 10kDa and 20kDa were visible apart from 2 bands at approximately 37kDa, one at 70kDa and one band at 150 kDa.

SJA and SBA binds to terminal N-acetylgalactosamine and galactose with the difference being that SJA prefers β-anomers. SJA showed a smear in P200,000g from very low molecular weight to ~80kDa where the band of THP was visible while some bands in the smear could be seen at 10, 20 and 35 kDa. In SN200,000g, one band at 20kDa, one at 37kDa and one at 65kDa and three bands below 15kDa could be seen. In SBA blots one band at 18 and one at 35 kDa could be seen in P200,000g while in SN200,000g three bands at 20, 37 and 50kDa could be seen. SBA and SJA did not have good signal in FLLA for any of the fractions which can be compared to lectin blots where signal is not intense, although present.

Con-A, LCA and PSA have been grouped together due to their shared nominal binding specificity to α-linked mannose although Con-A additionally recognises terminal and branched α-linked mannose and N-acetyl glucosamine residues while LCA and PSA also recognise fucose α1-6N-acetylglucosamine moieties and only LCA recognises α-linked glucose. In LCA P18,000g there is mainly the band of THP while one weak band at 50kDa can also be seen. In P200,000g bands at 160, 100, 60 and 37 kDa can also be seen while in SN200,000g many bands from between 25kDa to 250kDa are observed. In the case of Con-A

in all the fractions (P18,000g, P200,000g and SN200,000g) there was a smear from low to high molecular weight with many visible bands suggesting a number of proteins having α-mannose type of glycosylation which was expected.

In the PSA blot, bands of 37, 60, 80 and 100 could be seen in P200,000g while in P18000g THP was the major band recognised apart from a weak and at 37kDa. In SN200,000g bands of proteins with mol wt 25, 37, 45, 50 and 72 could be seen with considerable intensity. Con-A had good signal in FLLA for P18,000g and P200,000g fraction (which is comparable in lectin blots), less so for pure THP (similar in blotting comapred to P18,000g and P200,000g fractions) and poor signal for SN200,000g (same in lectin blot). In case of PSA, the best signal in FLLA was for P200,000g while other fractions did not have good signal. Multiple bands in P200,000g can be seen in the lectin blot for PSA. However signal intensity is better for P18,000g and pure THP which is in contrast with FLLA signal for these fractions.

LCA had no signal in FLLA in any of the fractions if we consider the standard deviation. However, in lectin blotting SN200,000g fraction has a lot of signal. This discrepancy in FLLA and lectin blotting signals for PSA and LCA can be partly explained by the fact that FLLA would show only surface glycan signal while lectin blotting can show all proteins regardless of their location inside or outside of the vesicles. WGA mainly binds glycan structures ending in single or multiple N-acetylglucosamine and even binds multivalent sialic acid residues. In P200,000g there was a smear from 35 to 75kDa and, apart from the distinct THP-specific band, there was a visible band at 120 kDa. There was not appreciable binding in P18,000g pellet apart from THP band while in SN200,000g there were multiple bands at 28, 40, 50, 55, 60 kDa and one band at 170kDa. In FLLA, WGA had similar signal in P18,000g and P200,000g while the half of that signal was for pure THP. However, in lectin blotting, WGA signal for P18,000g (THP band) and pure THP was visibly more than P200,000g which is in contrast with FLLA. For SN200,000g the signal of WGA in FLLA

was absent while multiple bands can be seen in this fraction in lectin blotting. Again these proteins might be present in the lumen of vesicles.

UEA1 binds to α-linked fucose residues and in P18,000g there was minute binding with THP while one intense band at 70kDa could be seen along with a doublet band at 37kDa, one band at 25kDa and an intense smear from 10kDa to 20kDa. MAL-II binds preferentially to α-2, 3 linked sialic acid while SNA binds to α-2, 6 linked sialic acid. In MAL-II blot pattern in P18,000g looked quite similar to UEA1 and accordingly glycan structures with fucose and α-2, 3 linked sialic acid have been reported (Wang *et al.*, 1990). The same was true of P200,000g but it could be concluded by the differential intensity of the THP bands that this similarity was not due to the non-specific binding of lectins to similar targets. Moreover in P200,000g there was a band at 40kDa in MAL-II which seemed to be absent in UEA1 blot. While, the pattern in SNA blot was quite different in P200,000g and apart from a more intense band of THP, a smear from 37 to 70kDa could be seen. In P18,000g, 22 and 27kDa bands were visible along with a smear from 40 to 80kDa. The signal in FLLA for UEA1 and MAL-II was comparable for all fractions other than P18,000g which has considerably more signal in lectin blotting. We have discussed previously that this could be due to the intra-vesicular proteins.

In the case of SNA lectin blotting had much more signal for SN200,000g compared to FLLA while P18,000g and P200,000g signal was much more in FLLA. In the case of THP the signal in the lectin blot was much more than P18,000g while the reverse was the case in FLLA. Speculatively, this could be due to the fact that denaturing treatment during SDS-PAGE could expose some glycans of THP which are otherwise hidden in native conditions of FLLA.

5.3.3 Lectin-affinity chromatography and hydrazide enrichment of glycoproteins from P18,000 and P200,000 extracts: glycoproteome of membrane vesicles

5.3.3.1 Hydrazide chemistry and surface glycome of urinary exosome and 'exosome-like' vesicles

To investigate which glycoproteins are present on the surface of the urinary exosome and 'exosome-like' vesicles, we employed hydrazide chemistry to enrich glycoproteins on surface of intact vesicles. The workflow for the technique is presented in Figure 5.12.

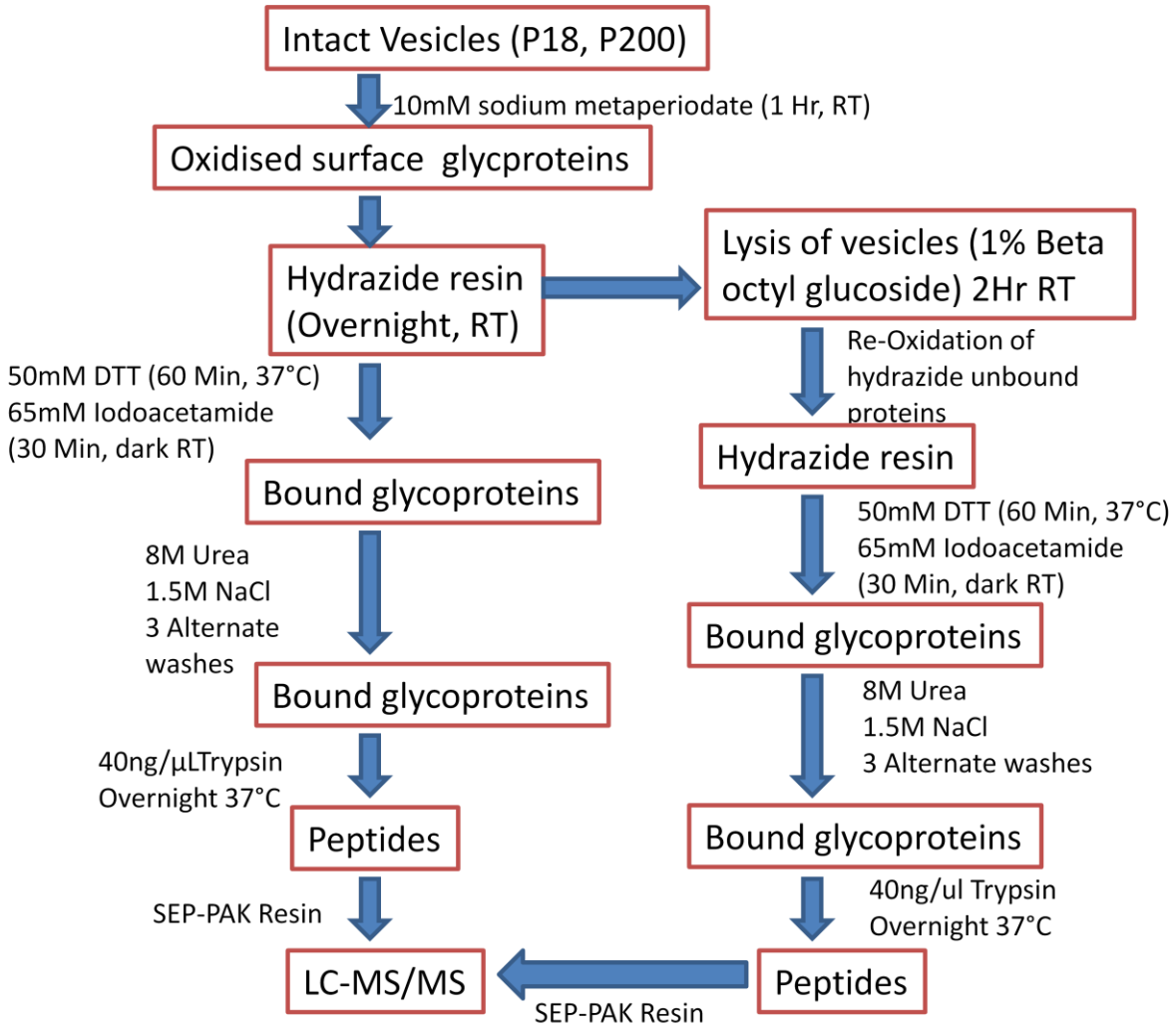

Figure 5.5: Workflow of the hydrazide chemistry methodology used in our anlysis. Crude pellets (P18,000g and P200,000g) were used as starting material.

This approach works by oxidising the diol moiety of the glycans to aldehyde which can bind covalently to the hydrazide resin. Non-bound and non-specifically reacting proteins were removed by three alternate washes with 8M urea and 1.5 M NaCl. This would result in almost complete removal of protein non-specifically bound to the resin matrix as well as proteins being retained due to their specific interactions with glycoproteins. This approach should result in identification of mostly glycosylated as well as non-enzymatically glycated proteins (Yaylayan, 2003; Mirzaei & Regnier, 2007; Zhang *et al.,* 2007). The surface proteins would be identified in the first part of the approach (Called P18S and P200S) while all other proteins would be identified in the second part after lysis of the vesicles and re-oxidation of the non-bound proteins (Called P18W and P200W). Although glycoproteins which would be binding to the surface of these vesicles by protein—protein, protein-glycan and protein-lipid interactions or by adsorption would also bind to the resin upon oxidation of the intact vesicles. Various numbers of proteins excluding keratins were identified in different samples (P200S: 92 and P200W: 41; P18S: 58 and P18W: 36). All these proteins are shown in supplementary table S5.1 in CD-ROM provided with the thesis. Comparison of all these proteins is presented in Figure 5.13.

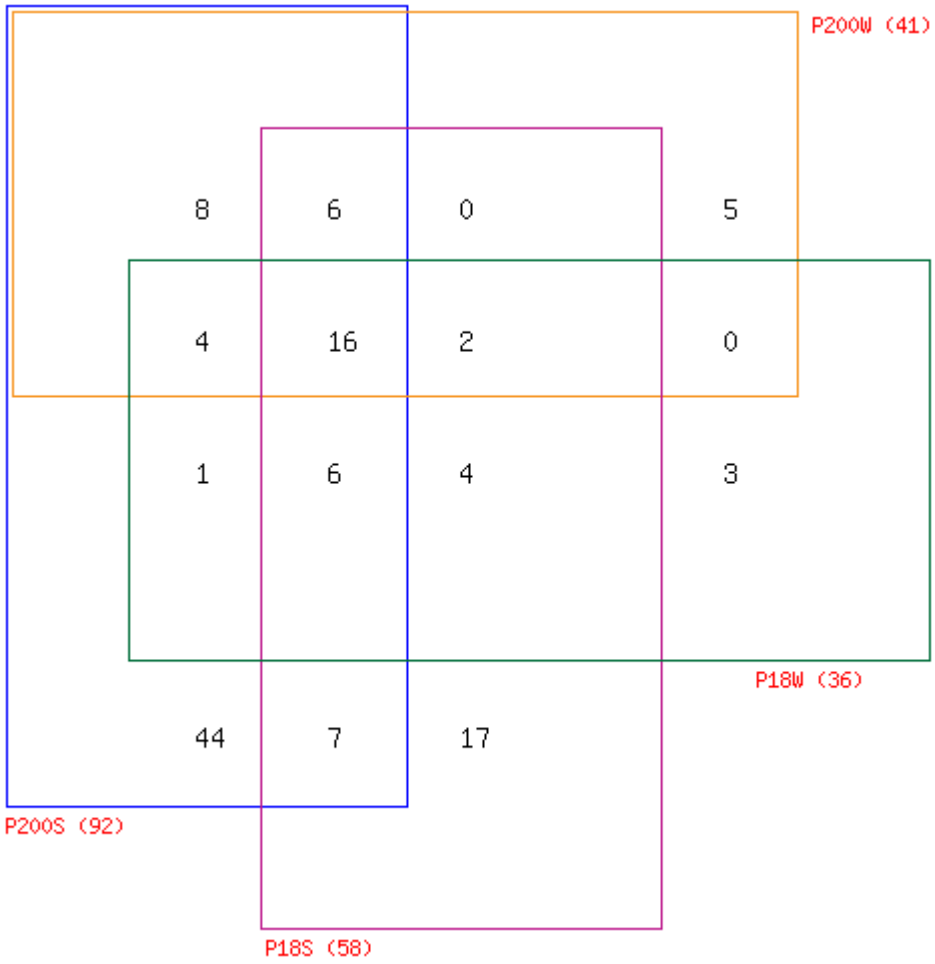

Figure 5.6: Comparison of proteins identified in various fractions of hydrazide chemistry analysis of P18 and P200. S means surface proteins (Intact vesicles used) W (Non-bound proteins from hydrazide resin after lysis of the vesicles).

Only 21 proteins were unique to P18S and 3 in P18W when compared to P200 identifications. These 24 proteins are indicated in the appendix at the end of this chapter. Out of these 24 proteins, 10 proteins are known to be glycosylated including 6 membrane proteins. These 6 membrane proteins are interesting and they are putative olfactory receptor 51H1, olfactomedin-4, solute carrier family 12 member 1, prostaglandin H2 D-isomerase, semenogelin-2, phosphoinositide-3-kinase interacting protein 1 and folate receptor alpha. The reason why these proteins were identified only in P18,000g remains to be proven but they are

clearly part of the surface of membrane vesicles. Ninety two proteins were identified in the P200S and 41 in P200W. However, only 7 proteins were unique to P200W. Out of these 7 proteins 3 are known glycoproteins namely angiotensinogen, serotransferrin and prostate specific antigen. All of these 3 are soluble secreted proteins and they probably end up in endosomes after endocytosis from where they can be secreted as part of membrane vesicles. Out of 92 proteins identified in P200S 61 proteins are annotated as known or potential glycoproteins (66% of the total). To establish the surface glycoproteins of the exosome and 'exosome-like' vesicles we classified these 61 proteins by SOSUI (Hirokawa, Boon-Chieng & Mitaku, 1998) (http://bp.nuap.nagoya-u.ac.jp/sosui/sosuiG/sosuigsubmit.html) and we found that 35 of these proteins are membrane proteins. Uniprot annotation was also used and it was found that 2 additional proteins are membrane-associated other than these 35. These are sulfhydryl oxidase 1 and CD14 molecule. These 37 proteins (60% of known glycoproteins identified in P200S) constitute the urinary membrane vesicles surface glycome.

Twenty four other proteins which are known to be glycosylated are present in P200S probably due to their interaction with vesicular surface or lysis of some vesicles upon freeze-thawing. However, there could be other reasons as well. For example, Napsin A, which is among these 24 proteins, is known to localise to MVB membranes and process pulmonary surfactant protein B into mature form (Ueno *et al.,* 2004). A number of proteins among these 24 proteins are protease inhibitors such as alpha-2-macroglobulins, alpha-1-antichymotrypsin and plasma serine protease inhibitor. Plasma serine protease inhibitor is associated with membrane of cells and microvesicles (Nishioka *et al.,* 1998). These reasons explain the presence of these proteins on surface of membrane vesicles. Other than these 24 proteins, thirty seven surface glycoproteins (known membrane proteins) are presented in Table 5.2.

Table 5.2: Presents the glycoproteins which are classified by either SOSUI and/or Uniprot as membrane associated or transmembrane proteins. These proteins were annotated by IPA (Ingenuity systems, Redwood city, CA). Uniprot accession, gene symbol Entrez gene name, location of the protein and type of molecule is indicated in the table.

Uniprot accession	Symbol	Entrez Gene Name	Location	Type (s)
P11117	ACP2	acid phosphatase 2, lysosomal	Cytoplasm	Phosphatise
P15144	ANPEP	alanyl (membrane) aminopeptidase	Plasma Membrane	Peptidase
P02649	APOE	apolipoprotein E	Extracellular Space	Transporter
P01024	C3	complement component 3	Extracellular Space	Peptidase
P08571	CD14	CD14 molecule	Plasma membrane	transmembrane receptor
P19835	CEL	carboxyl ester lipase (bile salt-stimulated lipase)	Extracellular Space	Enzyme
Q12860	CNTN1	contactin 1	Plasma Membrane	Enzyme
Q6UVK1	CSPG4	chondroitin sulfate proteoglycan 4	Plasma Membrane	Other
O60494	CUBN	cubilin (intrinsic factor-cobalamin receptor)	Plasma Membrane	transmembrane receptor
Q53TN4	CYBRD1	cytochrome b reductase 1	Cytoplasm	Enzyme
P27487	DPP4	dipeptidyl-peptidase 4	Plasma Membrane	Peptidase
P01133	EGF	epidermal growth factor	Extracellular Space	growth factor
O15197	EPHB6	EPH receptor B6	Plasma Membrane	Kinase
Q9NQ84	GPRC5C	G protein-coupled receptor, family C, group 5, member C	Plasma Membrane	G-protein coupled receptor
P98160	HSPG2	heparan sulfate proteoglycan 2	Plasma Membrane	Enzyme
P01591	IGJ	immunoglobulin J polypeptide, linker protein for immunoglobulin alpha and mu polypeptides	Extracellular Space	Other
P01042	KNG1	kininogen 1	Extracellular Space	Other
Q7L985	LINGO2	leucine rich repeat and Ig domain containing 2	Extracellular Space	Other
P98164	LRP2	low density lipoprotein receptor-related protein 2	Plasma Membrane	Transporter
P02788	LTF	Lactotransferrin	Extracellular Space	Peptidase

O43451	MGAM	maltase-glucoamylase (alpha-glucosidase)	Plasma Membrane	Enzyme
P08473	MME	membrane metallo-endopeptidase	Plasma Membrane	Peptidase
P15941	MUC1	mucin 1, cell surface associated	Plasma Membrane	transcription regulator
Q9BRK3	MXRA8	matrix-remodelling associated 8	unknown	Other
P54802	NAGLU	N-acetylglucosaminidase, alpha	Cytoplasm	Enzyme
P01833	PIGR	polymeric immunoglobulin receptor	Plasma Membrane	Transporter
O00592	PODXL	podocalyxin-like	Plasma Membrane	Kinase
O43490	PROM1	prominin 1	Plasma Membrane	Other
P00391	QSOX1	Quiescin Q6 sulfhydryl oxidase 1	Cytoplasm	Enzyme
Q86UN3	RTN4RL2	reticulon 4 receptor-like 2	Plasma Membrane	Other
P01009	SERPINA1	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1	Extracellular Space	Other
P05543	SERPINA7	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 7	Extracellular Space	Transporter
P55017	SLC12A3	solute carrier family 12 (sodium/chloride transporters), member 3	Plasma Membrane	Transporter
Q9UGT4	SUSD2	sushi domain containing 2	Extracellular Space	Other
P07911	UMOD	Uromodulin	Extracellular Space	Other
Q6EMK4	VASN	Vasorin	Plasma Membrane	Other
O43895	XPNPEP2	X-prolyl aminopeptidase (aminopeptidase P) 2, membrane-bound	Plasma Membrane	Peptidase

These 37 proteins were analysed by IPA software (Ingenuity systems, USA.) and some of the enriched bio-functions annotation are presented in the Table 5.3.

Table 5.3: Membrane glycoproteins of exosomal and exosome-like vesicles were enriched for some of the biological functions presented in the table. Category of functions, annotated function, p-Value of enrichement, molecules involved in the category and total number of molecules in the category is indicated in the table.

Category	Functions Annotation	p-Value	Molecules	# Molecules
Organ Morphology	morphology of kidney	1.68E-05	APOE,LRP2,NAGLU,PODXL, SLC12A3,UMOD	6
Renal and Urological System Development and Function	morphology of kidney	1.68E-05	APOE,LRP2,NAGLU,PODXL, SLC12A3,UMOD	6
Renal and Urological System Development and Function	proliferation of mesangial cells	9.76E-05	APOE,EGF,HSPG2	3
Cellular Development	proliferation of mesangial cells	9.76E-05	APOE,EGF,HSPG2	3
Cellular Growth and Proliferation	proliferation of mesangial cells	9.76E-05	APOE,EGF,HSPG2	3
Organ Morphology	morphology of kidney cells	8.19E-04	NAGLU,PODXL	2
Renal and Urological System Development and Function	morphology of kidney cells	8.19E-04	NAGLU,PODXL	2
Cell Morphology	morphology of kidney cells	8.19E-04	NAGLU,PODXL	2
Organ Morphology	abnormal morphology of renal tubule	2.04E-03	LRP2,SLC12A3,UMOD	3
Renal and Urological System Development and Function	abnormal morphology of renal tubule	2.04E-03	LRP2,SLC12A3,UMOD	3
Organismal Injury and Abnormalities	Fibrosis	3.48E-03	EGF,UMOD	2
Organ Morphology	abnormal morphology of nephrin	8.14E-03	LRP2,PODXL	2
Renal and Urological System Development and Function	abnormal morphology of nephrin	8.14E-03	LRP2,PODXL	2
Organ Morphology	mass of kidney	9.34E-03	APOE,EGF	2
Renal and Urological System Development and Function	mass of kidney	9.34E-03	APOE,EGF	2
Organ Development	mass of kidney	9.34E-03	APOE,EGF	2
Tissue Development	mass of kidney	9.34E-03	APOE,EGF	2

Embryonic Development	mass of kidney	9.34E-03	APOE,EGF	2
Organismal Development	mass of kidney	9.34E-03	APOE,EGF	2
Cell Death and Survival	apoptosis of kidney cells	1.16E-02	APOE,EGF	2
Renal and Urological System Development and Function	kidney development	1.38E-02	APOE,EGF,PODXL	3
Organ Development	kidney development	1.38E-02	APOE,EGF,PODXL	3
Tissue Development	kidney development	1.38E-02	APOE,EGF,PODXL	3
Embryonic Development	kidney development	1.38E-02	APOE,EGF,PODXL	3
Organismal Development	kidney development	1.38E-02	APOE,EGF,PODXL	3
Cell Death and Survival	cell viability	1.85E-02	EGF,UMOD	2
Cellular Development	proliferation of B-lymphocyte derived cell lines	2.38E-02	EGF,HSPG2	2
Cellular Growth and Proliferation	proliferation of B-lymphocyte derived cell lines	2.38E-02	EGF,HSPG2	2
Hematological System Development and Function	proliferation of B-lymphocyte derived cell lines	2.38E-02	EGF,HSPG2	2

Enrichment of one network was found by IPA among these 37 membrane glycoproteins. This network is regulation of some of these proteins by LIM homeobox 1 (LHX1) transcription factor. This network is shown in Figure 5.7.

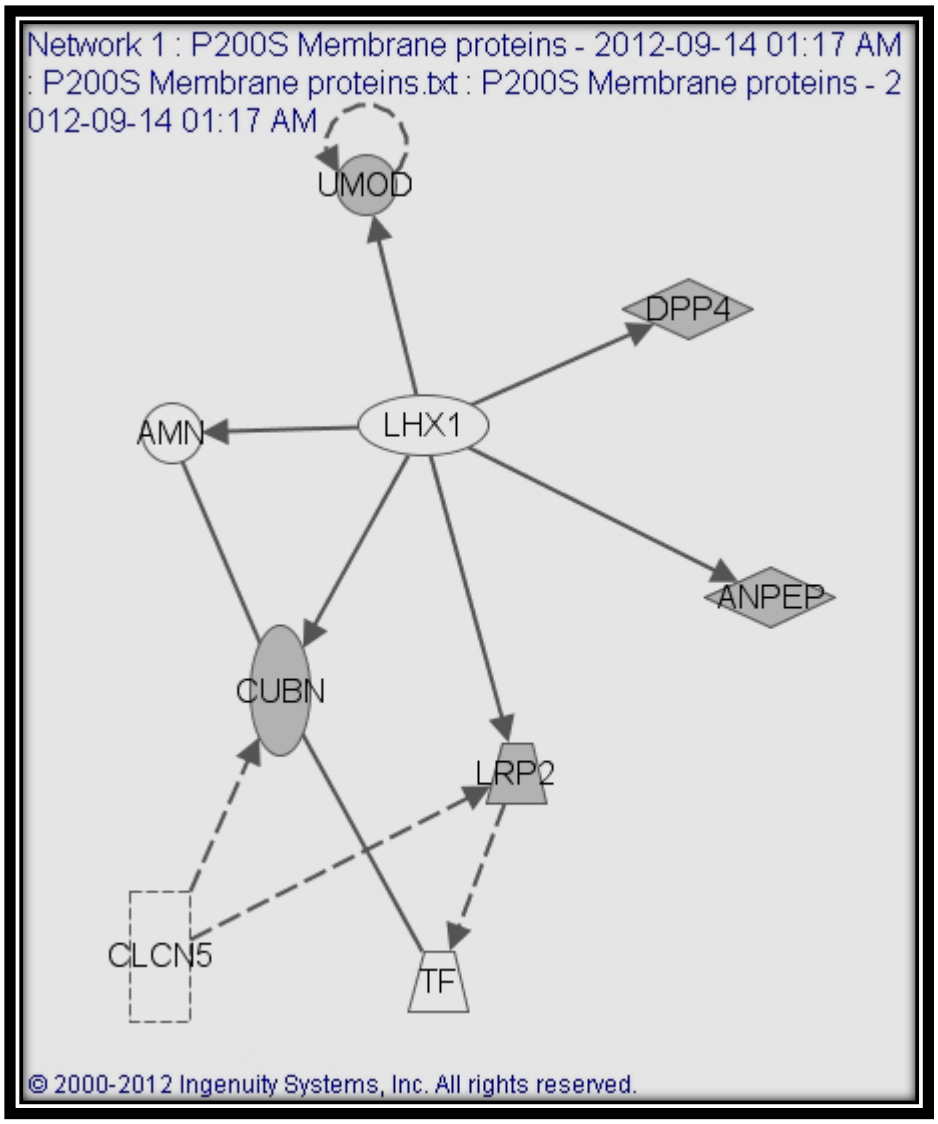

Figure 5.7: Presents the network enriched in membrane glycoproteins by IPA. LHX1: LIM homeobox 1 regulated proteins present in membrane glycoprotein identifications. Gene symbols have been used. The expression of all of these proteins is regulated by LHX1 transcription factor.

LHX1 has an important role in kidney field area specification during kidney development by mesoderm (Cirio *et al.*, 2011). These regulated proteins play important roles in kidney development as well as maintenance. These regulated proteins including cubulin and megalin maintain endocytic pathways in kidney to clear the lumen of nephrons. Dipeptidylpeptidase 4 (DPP4) and aminopeptidase N (ANPEP) control the processing of regulatory peptides in

tubular epithelial cells by regulating the cellular response to peptide hormones (Shipp & Look, 1993).

The protocol we have used for capturing glycoproteins from P200 is very stringent. After the oxidised proteins bound to hydrazide resin we washed the columns thrice with alternate 8M urea and 1.5M NaCl. This washing should get rid of all the non-specifically adsorbing proteins from the column and only the bound proteins should remain which would be expected to either be glycsylated or non-enzymatically glycated. These 31 candidate proteins are indicated in the appendix at the end of this chapter. Among these proteins, there are 4 membrane proteins; Sodium/potassium-transporting ATPase subunit alpha-1, gelsolin, nephrocystin-3 and stomatin. The presence of these 31 proteins in hydrazide-bound fraction could be explained in two ways with one being the interaction of these proteins with the glcosylated proteins found. We explored this option and using I2D (http://ophid.utoronto.ca/ophidv2.201/ppi.jsp) and iRefWeb (Turner *et al.,* 2010) (http://wodaklab.org/iRefWeb/search) databases we found interacting partners of these candidate proteins in the glycosylated protein list. We present the network that emerged in Figure 5.8.

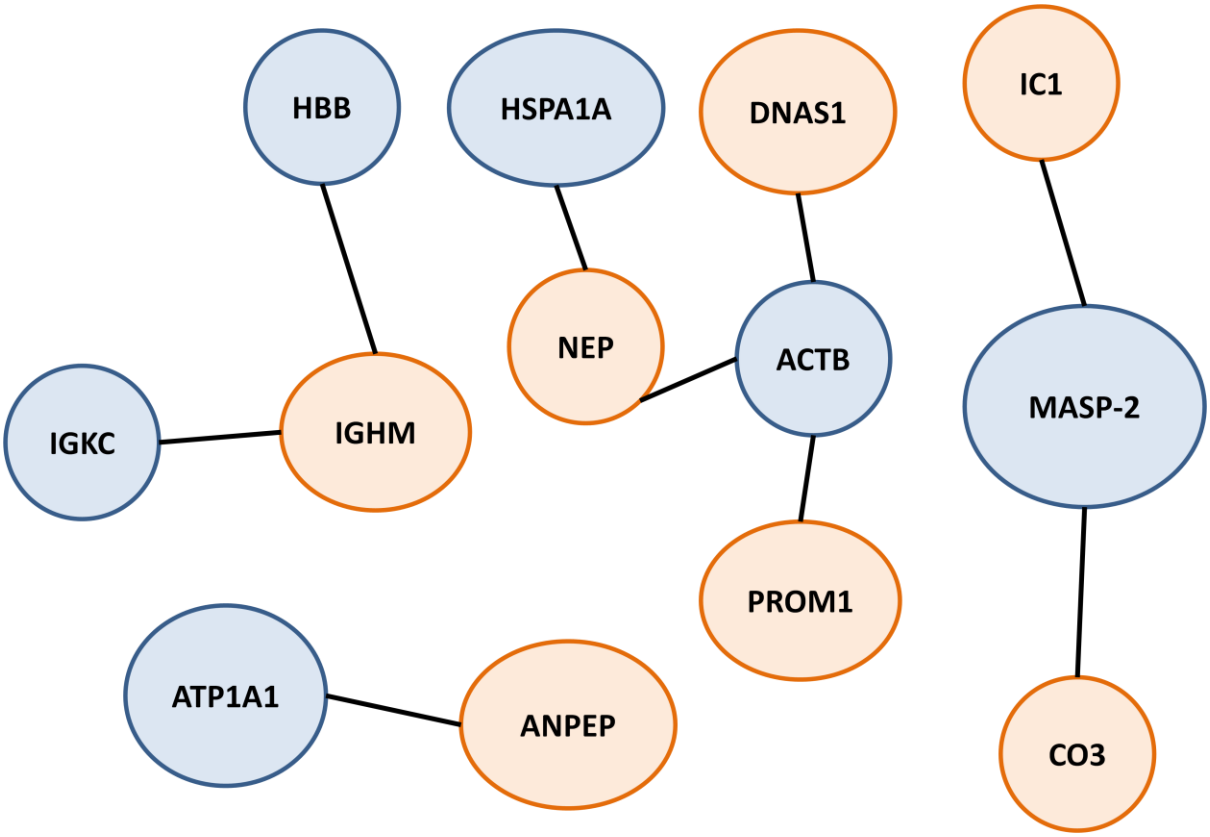

Figure 5.8: Interaction network was manually created between the glycosylated and non-glycosylated proteins using data from I2D and iRefWeb (Turner *et al.*, 2010) databases. The proteins in blue outlines are non-glycosylated proteins and proteins in orange outlines are glycosylated proteins found in hydrazide-enriched fractions. Interacting proteins are connected with thick black lines.

This theory explains the presence of 6 proteins and although they are not glycosylated, they might be bound to the hydrazide column through their glycosylated partners. However, it is surprising that, even after washing the column three times alternatively with 8M urea and 1.5M NaCl, these proteins maintained their association with their interacting partner. The second possibility is that amino acids in non-glycosylated proteins gets oxidised by sodium metaperiodate and subsequently coupled to hydrazide column. Oxidation of proteins and subsequent reaction with hydrazide has been previously reported (Mirzaei & Regnier, 2007). Moreover, in the same study the presence of cross-linked proteins upon oxidation was also

discovered. Cross-linked proteins and reaction of oxidised proteins with hydrazide can explain presence of these 31 proteins in hydrazide-bound fraction. These proteins need not be glycosyated to be present in this fraction.

5.3.3.2 Lectin affinity chromatography for enrichment of glycoproteins from P200,000g and P18,000g

Nine lectins were used for lectin affinity chromatography (LAC) of P18,000g and P200,000g extract (Table 5.4). Extract was prepared using 1% beta-octyl glucoside (because it solubilises lipid rafts (Garner, Smith & Hooper, 2008) and is compatible with LAC) to disrupt the vesicles for efficient enrichment of glycoproteins from both the pellets. Elution was performed on each lectin as described in methods (see elution conditions in Table 5.4). Most of these lectins were chosen because they showed high signal in FLLA (Figure 5.2). While, others like LCA, UEA and MAL-II were chosen because they recognised multiple proteins in the P18,000g and P200,000g in lectin blotting (Figure 5.3).

Table 5.4: Presents nine lectins used in LAC along with their origin and specificity.

Abbreviation	Origin	Nominal sugar binding specificity	Elution conditions
LCA/LcH	*Lens culinaris*	Fucα1-6GlcNAc ; α-Man ; α-Glc	200mM α-methyl mannoside + 200mM α-methyl glucoside
UEA-I	*Ulex europeus*	Fucα1-2LacNAc ; α-Fuc	100mM L-fucose
SNA	*Sambucus nigra*	Siaα2-6Gal ; GalNAc	500mM lactose followed by 500mM lactose in 100mM acetic acid
MAL II	*Maakia amurensis*	Siaα2-3Gal	200mM lactose
RCA^{120}	*Ricinus communis*	Lac ; LacNAc	200mM galactose + 200mM lactose
PHA (E)	*Phaseolus vulgaris*	Bisecting GlcNAc	100mM acetic acid
ConA	*Canavalia ensiformis*	α-Man branched and terminal; terminal GlcNAc	200mM α-methyl mannoside + 200mM α-

			methyl glucoside
WGA	*Triticum vulgaris*	(GlcNAc)n ; multivalent Sia ; β-GlcNAc, GalNAc	
JAC/ Jacalin	*Artocarpus integrifolia*	Gal β 1-3GalNAcα-Thr/Ser (T) ; GalNAcα-Thr/Ser (Tn)	800mM galactose

Based on the total proteins enriched by these lectins, one tenth of the proteins in elution was resolved by SDS-PAGE. This reflects the proteins enriched by corresponding lectins in LAC.

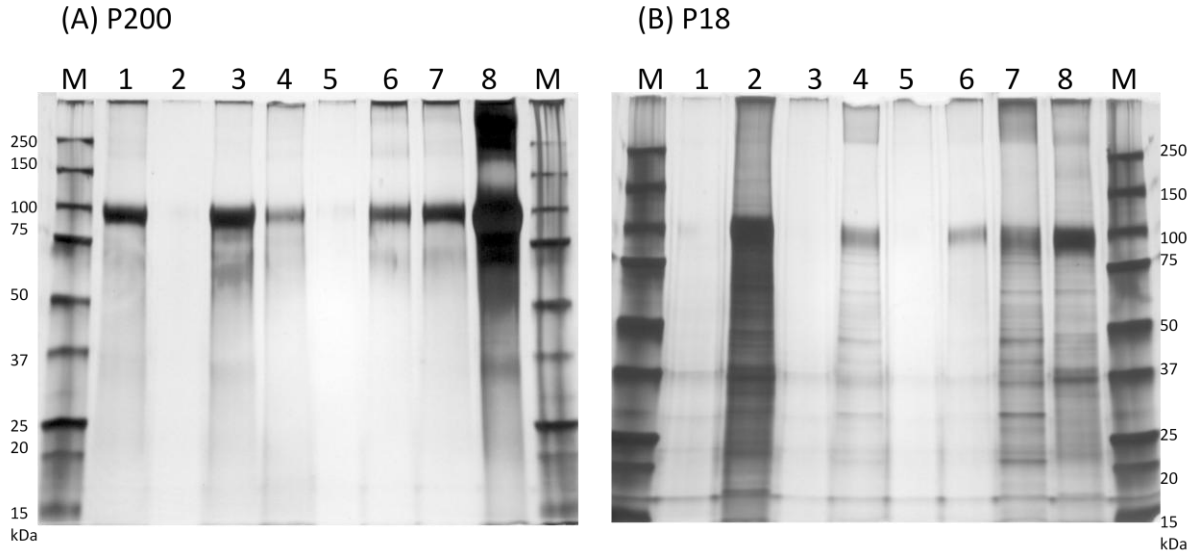

Figure 5.9: Silver stained SDS-PAGE (10%T) profile of proteins enriched by different lectins. Panel (A) (P200,000g): M. Molecular weight marker in kDa. 1. LCA 2. PHA-E 3. RCA 4. WGA 5. MAL-II 6. SNA 7. JAC 8. Con-A; Panel (B) (P18,000g): M. Molecular weight marker in kDa. 1. PHA-E 2. Con-A 3. WGA 4. RCA 5. JAC 6. SNA 7. MAL-II 8. LCA.

We used fixed amount of LAC elutions to run in the gel. Although starting amount of total protein used for each lectin in the LAC was same, but each lectin enriched variable amount of proteins therefore there is discrepancy in the number and amount of proteins shown in Figure 5.9. In LCA gel (Figure 5.9, Panel A, Lane 1) it can be seen that the major protein is THP (~100kDa) as is the case with RCA, WGA, SNA JAC and Con-A. In LCA, some low

molecular weight proteins below the THP band can be faintly seen and two proteins around 200-250 kDa are visible as well. These proteins can also be seen in lectin blot of LCA (Figure 5.4). In P18,000g elution of LCA (Figure 5.9, Panel B, lane 8), multiple proteins and a complex pattern of proteins from low to high molecular weights can be seen. This is however in contrast with FLLA (Figure 5.2) and lectin blot of LCA (Figure 5.4).

In PHA-E and MAL-II (Figure 5.9, Panel A, Lane 2 & 5) the faint band of THP can be seen however there are not many proteins shown in these two lectin elutions. The reason for this could be that amount (in µg) of proteins binding to this lectin was too low to be shown on the gel. Although the signal in FLLA for PHA-E was high (Figure 5.2), it could be seen in Figure 5.4 that the number of proteins recognised by PHA-E is less compared to other lectins. The same is true of PHA-E elution of P18,000g (Figure 5.9, Panel B, Lane 1). MAL-II recognised multiple proteins in P200,000g in lectin blot (Figure 5.4) but the intensity of signal was low. In P18,000g elution of MAL-II (Figure 5.9, Panel B, lane 7) there is a complex pattern of proteins across the range of molecular weight from low to high and it is similar to the lectin blot of MAL-II (Figure 5.4). However, signal in FLLA for MAL-II of P18,000g was low which is in contrast to lectin blot and SDS-PAGE of the elutions.

In case of RCA120 the number of proteins seen in the elution of both P18,000g and P2000,000g is high (Figure 5.9, Panel A, lane 3 and Panel B, lane 4). This is comparable with lectin blotting (Figure 5.4) and FLLA of RCA120 (Figure 5.2) for both these fractions. The number of proteins in P18,000g lectin blot of RCA120 is considerably less compared to elution shown here. In WGA elution of P200,000g (Figure 5.9, Panel A, lane 4) the number of proteins are more compared to P18,000g (Figure 5.9, Panel B, lane 3) which shows negligible signal. THP is the main band in elution of WGA of P200,000g which can be compared to THP signal in FLLA for WGA (Figure 5.2). The signal for Jacalin in elution of P200,000g is mainly a band which could be THP however, in FLLA and lectin blot of

Jacalin the signal for THP is absent. It could be argued that THP here is getting enriched non-specfically. For SNA, the major band seen in elution of P2000,000 (Figure 5.9, Panel A, lane 6) and P18,000g (Figure 5.9, Panel B, lane 6) is THP which could also be seen in lectin blot (Figure 5.4) and the signal for THP in FLLA (Figure 5.2).

For Con-A elution of P200,000g (Figure 5.9, Panel A, lane 8) there are a number of proteins from low to high molecular weight however the major protein is THP which is comparable to THP signal in FLLA (Figure 5.2). In P18,000g of Con-A elution (Figure 5.9, Panel B, Lane 2) there are multiple proteins and THP is not the major protein. This is comparable to signal of Con-A in the P18,000g, seen in FLLA (Figure 5.2). UEA-I elution is not shown in the Figure 5.9 because the amount of protein enriched was visibly small (Size of the dry pellet) which was processed directly to be analysed by LC-MS/MS.

To confirm whether or not all the proteins have been eluted from the lectin columns by our elution buffers, we boiled 50µL beads from lectin columns with SDS-PAGE loading buffer for 15 minutes and loaded the released proteins on the gel.

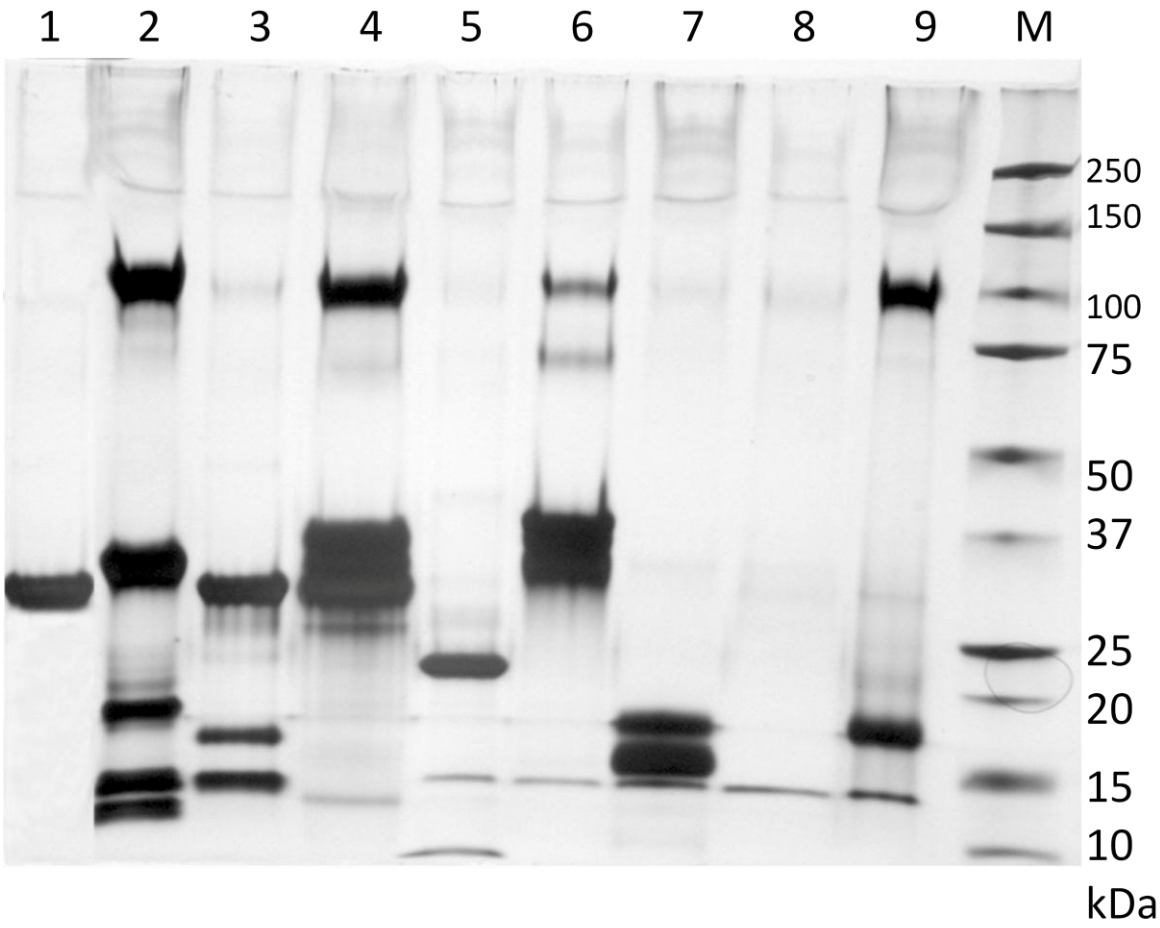

Figure 5.10: Silver stained SDS-PAGE (10%) of the proteins released by beads from lectin affinity columns. 1. UEA-I (31, 32 kDa), 2. PHA-E (33 kDa), 3. Con-A (26,52 kDa), 4. RCA (27, 33, 60 kDa), 5. LCA (8, 17 kDa), 6. SNA (36, 38 kDa), 7. JAC (10, 16 kDa), 8. MAL-II (34, 36 kDa), 9. WGA (18 kDa), M: Molecular weight markers in kDa.

The beads boiled in loading buffer mostly revealed the lectins and their subunits at their molecular weights with the exception of THP. THP was found abundantly in PHA-E, RCA, SNA and WGA columns while in only minor amounts in other lectins. No other major protein was present in the beads suggesting a majority of proteins bound to lectins had been eluted by the elution buffers used in our study. Elution was dialysed and concentrated using speed vacuum and processed by reduction, alkylation and trypsin digestions. Peptides were enriched using Sep-Pak solid phase extraction columns and analysed by LC-MS/MS.

5.3.3.3 Identification of proteins enriched by Lectin affinity chromatography

The nine lectins were used in LAC to enrich various subpopulations of glycoproteins displaying diverse glycan epitopes on their surface. Multiple proteins were identified in elution of every lectin. Excluding keratins, varying number of proteins were identified from LAC of P200,000g with 35 in RCA, 53 in WGA, 78 in JAC, 68 in LCA, 58 in SNA, 66 in PHA-E, 59 in Con-A, 44 in MAL-II and 48 in UEA. In LAC of P18,000g varying number of proteins, with 10 in RCA, 45 in WGA, 13 in MAL-II, 28 in JAC, 20 in LCA, 45 in SNA, 34 in PHA-E, 21 in Con-A and 49 in UEA were identified. All these proteins are shown in supplementary table S5.1 in CD-ROM provided with the thesis. Multiple proteins were common to many lectins. In total 172 non-redundant unique proteins were identified in P200,000g and 97 in P18,000g. A majority of the proteins (73% of total) identified in P18,000g were common to P200,000g (Figure 5.11).

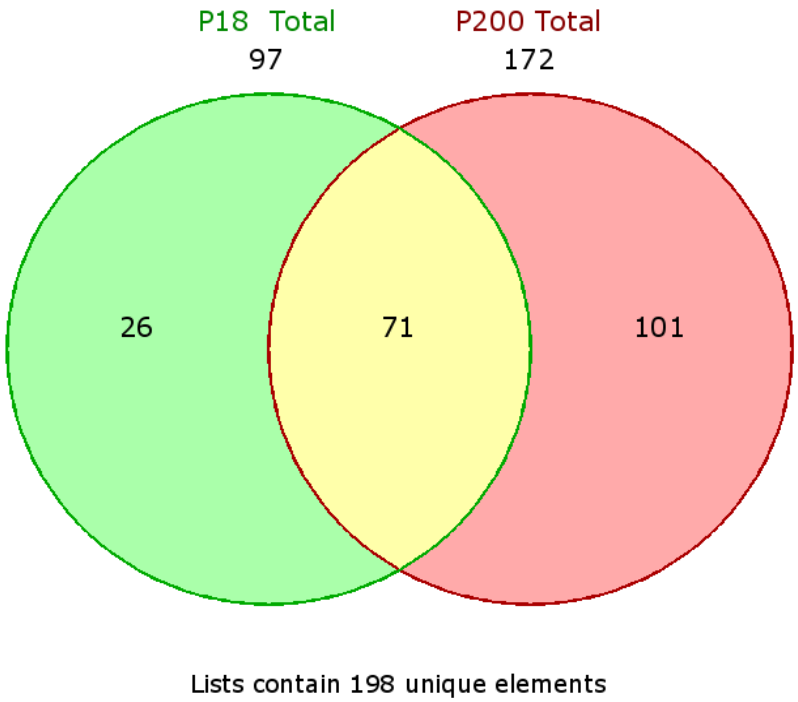

Figure 5.11: Comparison of the proteins identified by LAC in P200,000g and P18,000g.

Our objective was to identify the major glycoprotein population in the exosomal and low speed pellet. We did not take a serial LAC approach (Non-bound of one lectin bound to another one) which could yield a greater number of identifications by removing high abundance proteins in the first few lectin columns. But such an approach would miss out on giving information about type of glycan epitope present on these proteins. We have incubated both the pellet (600µg each) separately with each of these lectins and subsequently identified the bound proteins. Some proteins which are heavily glycosylated at multiple sites with diverse glycans would be expected to be present in high amount in each LAC fraction. By comparing proteins identified in 9 LAC fractions we can identify the proteins which bind to multiple lectins. Crude information about type of glycans present on these proteins would be revealed in the process. Out of 172 proteins identified in P200,000g, 64 proteins (37% of total) were known glycosylated proteins, as annotated by SwissProt. These 64 proteins were distributed in a complex manner among 9 LAC fractions. All of the 9 LAC were compared with each other and the results are presented in table 5.5.

Table 5.5: All the known glycoproteins identified in P200,000 by LAC using 9 lectins are presented. Note that multiple proteins were bound to many different lectins.

UNIPROT ACCESSION	Gene Name	Lectin binding
O43653 P05090	prostate stem cell antigen, Apolipoprotein D	WGA, JAC, LCA, SNA, PHA-E, Con-A, UEA
O75594	peptidoglycan recognition protein 1	WGA, SNA, PHA-E,Con-A, UEA
P00734	coagulation factor II (thrombin)	MAL-II
P00738	haptoglobin-related protein; haptoglobin	WGA, LCA, SNA, MAL-II, UEA
P01009	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1	WGA, JAC, LCA, SNA, MAL-II, Con-A, UEA
P01011 P0C0L4	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3; Complement component 4A (Rodgers blood group)	JAC, LCA, UEA
P01023	alpha-2-macroglobulin	WGA, JAC, LCA,

		Con-A, UEA
P01042	kininogen 1	JAC
P01596 Q12907 P04114 P00450 P16278	Ig kappa chain V-I region CAR; lectin, mannose-binding 2; apolipoprotein B (including Ag (x) antigen); ceruloplasmin (ferroxidase); galactosidase, beta 1	LCA
P01833	polymeric immunoglobulin receptor	JAC, LCA, Con-A, MAL-II, UEA
P01857 P01024 P02787	immunoglobulin heavy constant gamma 1; Complement component 3; transferring	WGA, JAC, LCA, SNA, MAL-II, PHA-E, Con-A, UEA
P01859	immunoglobulin heavy constant gamma 2 (G2m marker)	WGA, JAC, LCA, SNA, PHA-E, Con-A
P01876	immunoglobulin heavy constant alpha 1	WGA, JAC, LCA, SNA, MAL-II, RCA, Con-A, UEA
P02649 Q14624	apolipoprotein E; inter-alpha (globulin) inhibitor H4	JAC, Con-A
P02671	fibrinogen alpha chain	JAC, UEA
P02675 P02679 P01871	fibrinogen beta chain; fibrinogen gamma chain; immunoglobulin heavy constant gamma 1 (G1m marker)	WGA, JAC, LCA,SNA, Con-A, MAL-II, UEA
P02743 P01860 P04004 O60494 P01861	amyloid P component, serum; immunoglobulin heavy constant gamma 1 (G1m marker); vitronectin; cubilin; immunoglobulin heavy constant gamma 4 (G4m marker)	UEA
P02760	alpha-1-microglobulin/bikunin precursor	WGA, JAC, LCA
P02768 P07911	albumin; uromodulin	WGA, JAC, LCA, SNA, MAL-II, PHA-E, RCA, Con-A, UEA
P02790	Hemopexin	JAC, LCA, SNA, Con-A
P04196	histidine-rich glycoprotein	JAC, SNA
P04745	Alpha-amylase 1	WGA, JAC, LCA, PHA-E, RCA, MAL-II
P05155	serpin peptidase inhibitor, clade G (C1 inhibitor), member 1	LCA, UEA
P07339	cathepsin D	PHA-E
P07858	cathepsin B	WGA
P08603 Q8N2E2 P27487	complement factor H; von Willebrand factor D and EGF domains; dipeptidyl-peptidase 4	RCA
P08962 O96009 P01763	CD63 molecule; napsin A aspartic peptidase; Ig heavy chain V-III region WEA	JAC, LCA, Con-A
P09668	cathepsin H	LCA, SNA, PHA-E, Con-A
P12273	prolactin-induced protein	Con-A, MAL-II

P25311	Zinc-alpha-2-glycoprotein	WGA, JAC, LCA, SNA, MAL-II, RCA, Con-A, PHA-E
P41222	prostaglandin D2 synthase	WGA, JAC
P51688	N-sulfoglucosamine sulfohydrolase	LCA, Con-A
P78509	Reelin	SNA
P80188	lipocalin 2	PHA_E, RCA
P98164	low density lipoprotein-related protein 2	JAC, LCA
Q02413	desmoglein 1	WGA, JAC, SNA, PHA-E, RCA, MAL-II
Q08380	lectin, galactoside-binding, soluble, 3 binding protein	WGA, JAC, LCA, SNA, Con-A, UEA
Q08554	desmocollin 1	WG, JAC, LCA, PHA-E
Q12805 P21926 P02763 P02751	EGF-containing fibulin-like extracellular matrix protein 1; CD9 molecule; orosomucoid 1; fibronectin 1	Con-A
Q14574	desmocollin 3	JAC, PHA-E
Q15517	Corneodesmosin	WGA, SNA, PHA-E, Con-A

Most of the proteins supposedly bound specifically to many different lectins and these proteins would be expected to have multiple glycosylation sites with a diverse array of glycans attached to them. For example, prostate stem cell antigen (PSCA), which is a GPI anchored protein, has 4 predicted N-glycosylation sites and it was shown in a study that treatment with PNGase F resulted in only one band of PSCA at 12 kDa. However, protein not treated with PNGase F had three isoforms at 12, 16 and 24 kDa (Reiter *et al.*, 1998). This demonstrates that as much as 40% weight of this protein can be glycan moieties. Thirty seven known glycoproteins were identified in P18,000g LAC with similar pattern in which some proteins were bound to multiple lectins. We compared this 37 known glycoproteins from P18,000g with 64 glycoproteins from P200,000g (Figure 5.12).

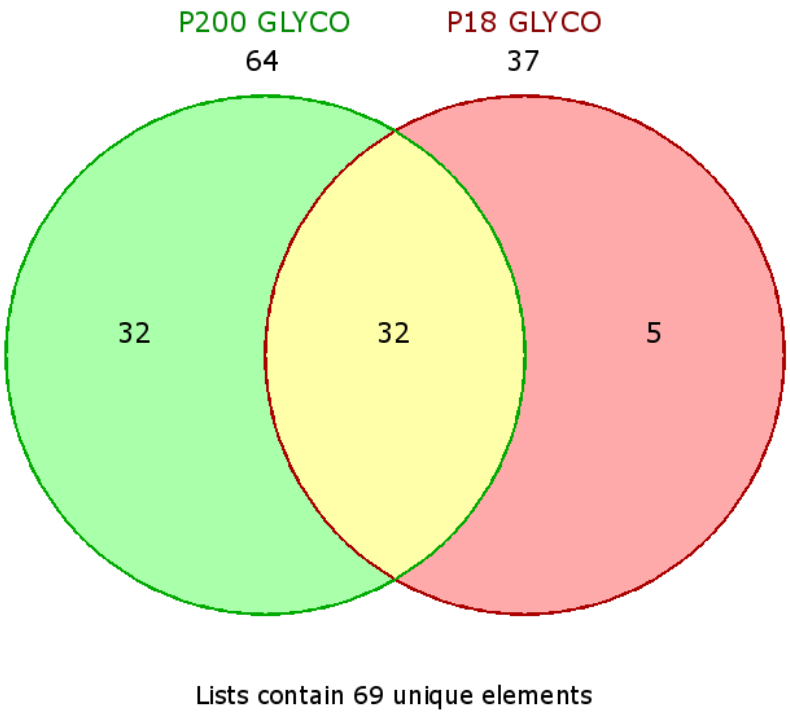

Figure 5.12: Compares the glycoproteins identified in P18,000g and P200,000g by LAC.

As seen in Figure 5.12, only 5 glycoproteins were unique to P18,000g (86% proteins common to P200,000 glycoproteins) therefore proteins identified in P18,000g will not be discussed further and we will focus only on glycoproteins identified in P200,000g.

5.3.3.2 Sialylome of urinary nano-vesicles

Three sialic acid-binding lectins were used in LAC (WGA, SNA, MAL-II). These lectins have different glycan binding specificities. Although WGA can bind to sialic acid in any linkage, it also binds N-acetylglucosamine. However, SNA and MAL-II are linkage-specific and bind α2-6 and α2-3 linked sialic acids respectively. Therefore comparison of proteins identified by LAC using these 3 lectins can reveal the nature of sialic acid linkage on identified proteins. WGA binding can confirm the presence of sialic acids on proteins identified by SNA and MAL-II. A comparison of known glycoproteins identified in LAC of these three lectins is shown in Figure 5.13.

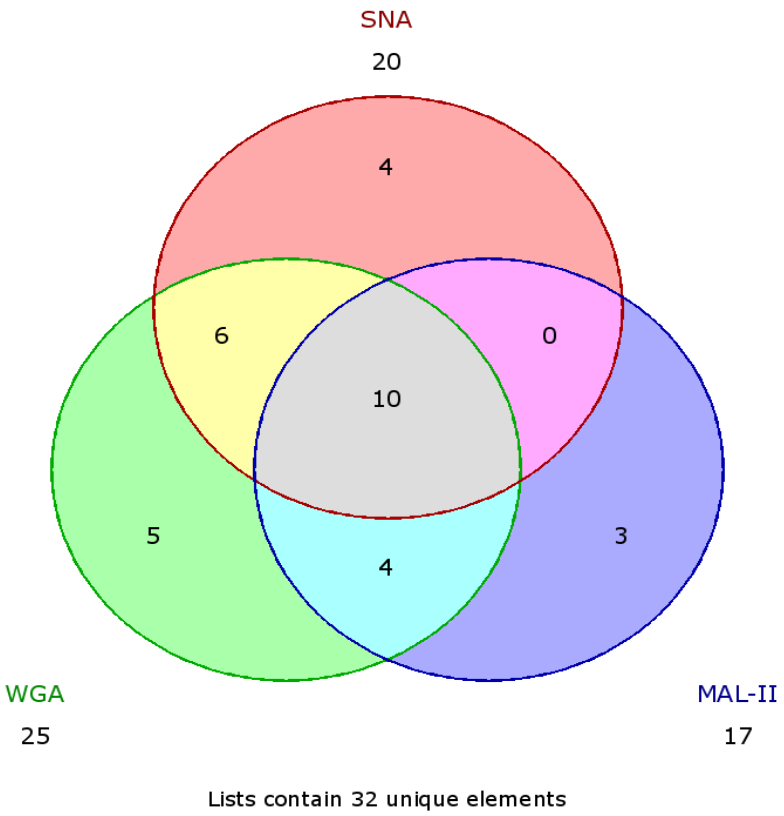

Figure 5.13: Venn diagram showing the comparison of the proteins identified in WGA, SNA and MAL-II LAC.

As can be seen in the Figure 5.13, 10 proteins are common to the three lectins which suggests that these glycoproteins are not only sialylated but also contain both α2-6 and α2-3 linkages of sialic acid. Ten other glycoproteins are present in SNA elution (6 common with WGA and 4 unique to SNA) which would be expected to have only α2-6 linked sialic acid. Seven different proteins are present in MAL-II elution (4 common with WGA and 3 unique to MAL-II) which would be expected to have only α2-3 linked sialic acid (See Table 5.5). We have also described results of glycoproteins identified by hydrazide chemistry in previous section. Periodate oxidation also oxidises sialic acids (Larsen *et al.*, 2007) and the resulting aldehyde is coupled to hydrazide resin. Therefore we compared the 27 non-redundant glycoproteins identified by MAL-II and SNA to proteins identified by hydrazide chemistry (Figure 5.14).

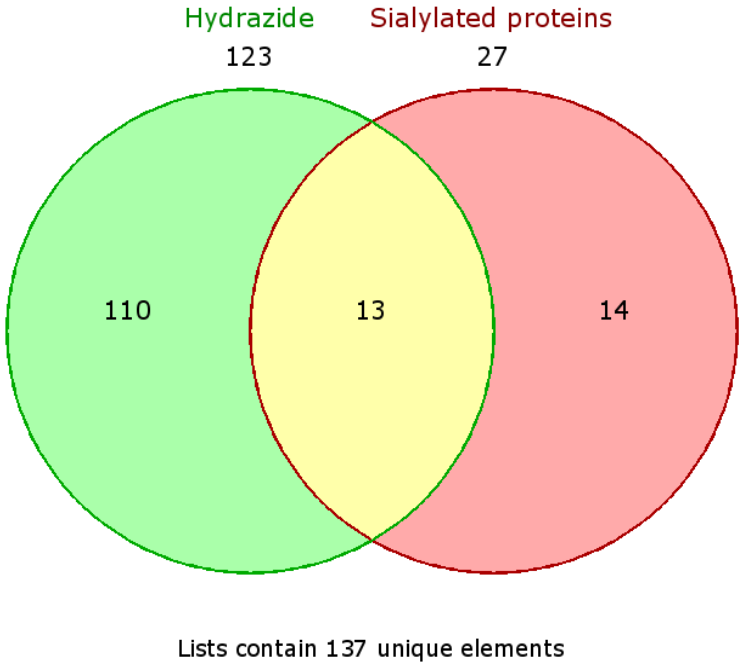

Figure 5.14: Presents comparison of sialylated proteins (identified in LAC using SNA and MAL-II) with proteins identified by hydrazide chemistry.

Thirteen of the 27 proteins (48%) were also identified by hydrazide chemistry method validating the results of LAC. These 27 sialylated proteins are presented in Table 5.5.

Table 5.5: The sialylated proteins identified in our study. Lectin binding and accordingly type of sialic acid linkage, presence of proteins in hydrazide chemistry based enrichment and confirmed sialylation in another study done with human plasma and saliva (Larsen *et al.*, 2007) is indicated in the table.

Uniprot accession	Gene/Protein Name	Lectin Binding	α2-3 or α2-6 linkage	Present in Hydrazide chemistry identifications	Known sialoprotein in Plasma or saliva (Larsen *et al.*, 2007)
P00734	Coagulation factor II (thrombin)	MAL-II	α2-3	Yes	Yes
P01833	Polymeric immunoglobulin receptor	MAL-II	α2-3	Yes	Yes
P12273	Prolactin-induced protein	MAL-II	α2-3	No	Yes
P02790	Hemopexin	SNA	α2-6	No	Yes

P04196	Histidine-rich glycoprotein	SNA	α2-6	Yes	Yes
P09668	Cathepsin H	SNA	α2-6	No	No
P78509	Reelin	SNA	α2-6	No	No
P01871	Ig mu chain C region	WGA, MAL-II	α2-3	Yes	Yes
P02675	Fibrinogen beta chain	WGA, MAL-II	α2-3	No	Yes
P02679	Fibrinogen gamma chain	WGA, MAL-II	α2-3	No	Yes
P04745	Amylase, alpha 1 (salivary)	WGA, MAL-II	α2-3	Yes	Yes
P00738	Haptoglobin	WGA, MAL-II, SNA	Both	No	Yes
P01009	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1	WGA, MAL-II, SNA	Both	Yes	No
P01024	Complement component 3	WGA, MAL-II, SNA	Both	Yes	Yes
P01857	Immunoglobulin heavy constant gamma 1 (G1m marker)	WGA, MAL-II, SNA	Both	No	No
P01876	Immunoglobulin heavy constant alpha 1	WGA, MAL-II, SNA	Both	Yes	Yes
P02768	Albumin	WGA, MAL-II, SNA	Both	Yes	Depleted before sialylated protein enrichment
P02787	Transferrin	WGA, MAL-II, SNA	Both	Yes	Yes
P07911	Uromodulin	WGA, MAL-II, SNA	Both	Yes	No
P25311	Zinc -alpha-2-glycoprotein	WGA, MAL-II, SNA	Both	No	Yes
Q02413	Desmoglein 1	WGA, MAL-II, SNA	Both	No	No
O43653	Prostate stem cell antigen	WGA, SNA	α2-6	No	No
O75594	Peptidoglycan recognition protein 1	WGA, SNA	α2-6	No	No
P01859	Immunoglobulin heavy constant gamma 2 (G2m marker)	WGA, SNA	α2-6	No	Yes
P05090	Apolipoprotein D	WGA, SNA	α2-6	Yes	Yes
Q08380	Galectin-3-binding protein	WGA, SNA	α2-6	Yes	Yes
Q15517	Corneodesmosin	WGA, SNA	α2-6	No	No

A majority of the proteins (63%) present in the table have been previously confirmed as sialoproteins in human plasma and saliva (Larsen *et al.*, 2007).

5.3.3.4 α1-2 and α1-6 (Core) fucosylated proteins in exosomal pellet

Two different lectins LCA and UEA-I were used in LAC for identifying fucosylated proteins. While LCA recognises α1-6 linked L-fucose otherwise known as core fucosylation, UEA-I binds to α1-2 linked L-fucose. Con-A binds biantennary glycans and some of these will have core fucosylation therefore such glycans will be bound by LCA as well. Most glycans which have α1-2 linked fucose on the outer arm (bound by UEA-I) also contain α1-6 core fucosylation (a1-2, 1-6 on one protein). Therefore multiple proteins should be common between Con-A, LCA and UEA-I although some of them would contain only one type of fucosylation either α1-2 or α1-6. The comparison between proteins identified in these three lectins is shown in Figure 5.15.

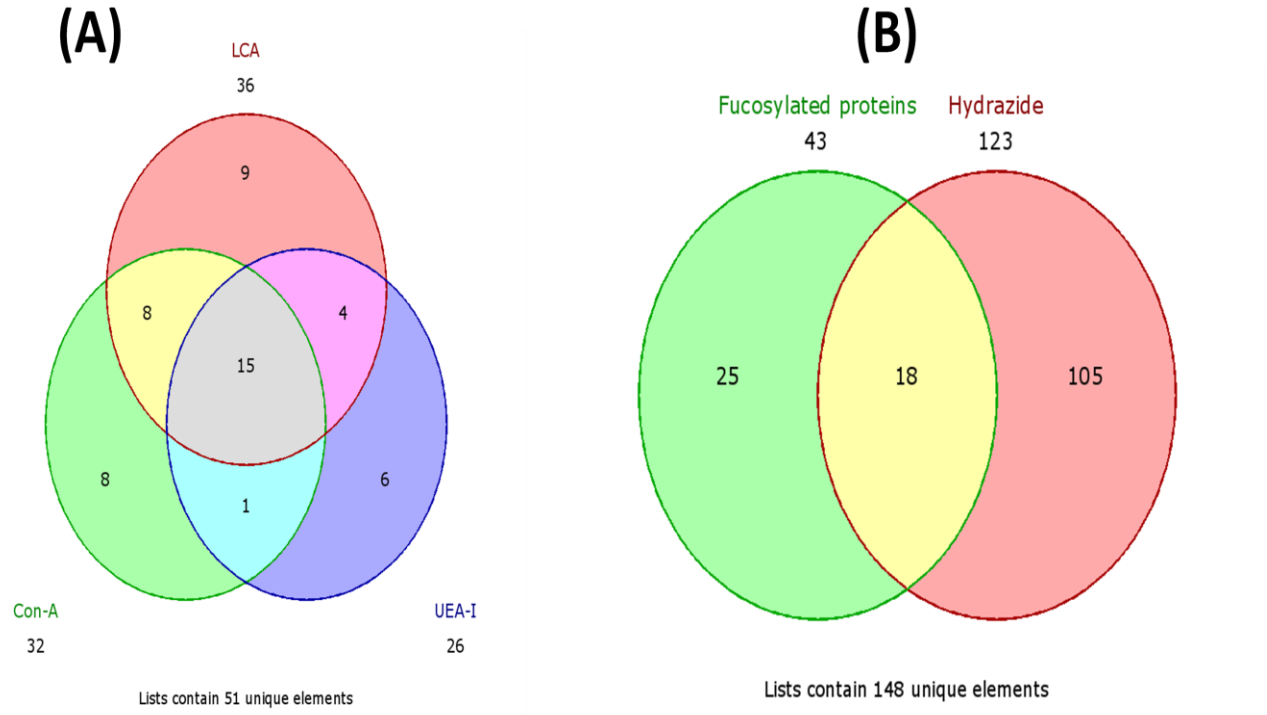

Figure 5.15: Comparison between proteins identified in Con-A, LCA and UEA-I lectin (Panel A) and comparison of total fucosylated proteins with hydrazide method protein identifications (Panel B).

High-mannose type glycans are not known to be a substrate for α1,6-Fucosyltransferase, therefore, the proteins found in Con-A common with LCA should be biantennary hybrid or complex. Seventeen proteins in LCA (9 unique and 8 common with Con-A) are expected to possess only core fucosylation and not α1-2 type (See Table 5.6). Seven proteins (6 Unique to UEA-I and 1 common with Con-A) should contain only α1-2 type fucosylation which is usually found on outer arm of the glycan. However, 19 proteins common between LCA and UEA-I should have both types (α1-6 and α1-2) of fucosylation. These 43 potentially fucosylated proteins are present in table 5.6.

Table 5.6: Potentially fucosylated proteins in urinary membrane vesicles. Uniprot accession, protein name, binding to either LCA or UEA-I and comparison with Con-A identified proteins, potential sialic acid found on same same protein, and identification in Hydrazide method are indicated in the table.

Uniprot	Gene/Protein Name	Lectin binding	Type of glycan	Found in Sialic acid binding Lectins	Present in Hydrazide method
O96009	Napsin A aspartic peptidase	Con-A, LCA	Biantennary, α1-6 Fucose	No	Yes
P01763	Ig heavy chain V-III region WEA	Con-A, LCA	Biantennary, α1-6 Fucose	No	No
P01859	Immunoglobulin heavy constant gamma 2 (G2m marker)	Con-A, LCA	Biantennary, α1-6 Fucose	Yes	No
P02790	Hemopexin	Con-A, LCA	Biantennary, α1-6 Fucose	Yes	No
P08962	CD63 molecule	Con-A, LCA	Biantennary, α1-6 Fucose	No	No
P09668	Cathepsin H	Con-A, LCA	Biantennary, α1-6 Fucose	Yes	No
P25311	Alpha-2-glycoprotein 1, zinc-binding pseudogene 1; alpha-2-glycoprotein 1, zinc-binding	Con-A, LCA	Biantennary, α1-6 Fucose	Yes	No
P51688	N-sulfoglucosamine sulfohydrolase	Con-A, LCA	Biantennary, α1-6 Fucose	No	No
O43653	Prostate stem cell antigen	Con-A, LCA, UEA-I	Biantennary, α1-2, α1-6 Fucose	Yes	No
P01009	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase,	Con-A, LCA,	Biantennary, α1-2, α1-6	Yes	Yes

	antitrypsin), member 1	UEA-I	Fucose		
P01023	Alpha-2-macroglobulin	Con-A, LCA, UEA-I	Biantennary, α1-2, α1-6 Fucose	No	Yes
P01024	complement component 3	Con-A, LCA, UEA-I	Biantennary, α1-2, α1-6 Fucose	Yes	Yes
P01833	Polymeric immunoglobulin receptor	Con-A, LCA, UEA-I	Biantennary, α1-2, α1-6 Fucose	Yes	Yes
P01857	Ig gamma-1 chain C region	Con-A, LCA, UEA-I	Biantennary, α1-2, α1-6 Fucose	Yes	No
P01871	Ig mu chain C region	Con-A, LCA, UEA-I	Biantennary, α1-2, α1-6 Fucose	Yes	Yes
P01876	Immunoglobulin heavy constant alpha 1	Con-A, LCA, UEA-I	Biantennary, α1-2, α1-6 Fucose	Yes	Yes
P02675	Fibrinogen beta chain	Con-A, LCA, UEA-I	Biantennary, α1-2, α1-6 Fucose	Yes	No
P02679	Fibrinogen gamma chain	Con-A, LCA, UEA-I	Biantennary, α1-2, α1-6 Fucose	Yes	No
P02768	Albumin	Con-A, LCA, UEA-I	Biantennary, α1-2, α1-6 Fucose	Yes	Yes
P02787	Transferrin	Con-A, LCA, UEA-I	Biantennary, α1-2, α1-6 Fucose	Yes	Yes
P05090	Apolipoprotein D	Con-A, LCA, UEA-I	Biantennary, α1-2, α1-6 Fucose	Yes	Yes
P07911	Uromodulin	Con-A, LCA, UEA-I	Biantennary, α1-2, α1-6 Fucose	Yes	Yes
Q08380	Galectin-3 binding protein	Con-A, LCA, UEA-I	Biantennary, α1-2, α1-6 Fucose	Yes	Yes
O75594	Peptidoglycan recognition protein 1	Con-A, UEA-I	Biantennary, α1-2 Fucose	Yes	No
P00450	Ceruloplasmin (ferroxidase)	LCA	α1-6 Fucose	No	No
P01596	Ig kappa chain V-I region CAR	LCA	α1-6 Fucose	No	No
P02760	Alpha-1-microglobulin/bikunin precursor	LCA	α1-6 Fucose	No	Yes
P04114	Apolipoprotein B (including Ag (x) antigen)	LCA	α1-6 Fucose	No	No
P04745	Amylase, alpha 1 (salivary)	LCA	α1-6 Fucose	Yes	Yes
P16278	Galactosidase, beta 1	LCA	α1-6 Fucose	No	No
P98164	low density lipoprotein-related protein 2	LCA	α1-6 Fucose	No	Yes
Q08554	Desmocollin 1	LCA	α1-6 Fucose	No	No

Q12907	Lectin, mannose-binding 2	LCA	α1-6 Fucose	No	No
P00738	Haptoglobin	LCA, UEA-I	α1-2, α1-6 Fucose	Yes	No
P01011	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3	LCA, UEA-I	α1-2, α1-6 Fucose	No	Yes
P05155	Serpin peptidase inhibitor, clade G (C1 inhibitor), member 1	LCA, UEA-I	α1-2, α1-6 Fucose	No	Yes
P0C0L4	Complement component 4A (Rodgers blood group)	LCA, UEA-I	α1-2, α1-6 Fucose	No	No
O60494	Cubilin (intrinsic factor-cobalamin receptor)	UEA-I	α1-2 Fucose	No	Yes
P01860	Ig gamma-3 chain C region	UEA-I	α1-2 Fucose	No	No
P01861	Immunoglobulin heavy constant gamma 4 (G4m marker)	UEA-I	α1-2 Fucose	No	No
P02671	Fibrinogen alpha chain	UEA-I	α1-2 Fucose	No	No
P02743	Amyloid P component, serum	UEA-I	α1-2 Fucose	No	No
P04004	Vitronectin	UEA-I	α1-2 Fucose	No	No

Many of these proteins like transferrin, haptoglobin, immunoglobulin G and A and polymeric immunoglobulin receptor have been previously identified in other studies to be α1-6 core fucosylated (Dai *et al.*, 2007; Muinelo-Romay *et al.*, 2011).

5.3.3.5 Galactose bearing, high-mannose type and Complex glycan containing glycoproteins

As described in Table 5.1, Jacalin binds 'T antigen' mostly, on O-glycosylated proteins. However, another study has shown that JAC can also bind oligomannose type glycans (Bourne *et al.*, 2002). This population of proteins should also bind Con-A. Therefore, glycoproteins identified in Jac LAC were compared to those identified in Con-A. The common proteins should bear oligomannose type glycans and those unique to JAC should contain 'T antigen' type of glycan. Moreover, Con-A also binds biantennary types of glycans and it should have some proteins common with PHA-E. RCA, on the other hand binds galactose in N-acetyllactosamine. Some gaalctose-bearing proteins will be bound by JAC as well. For these reasons we compared the proteins identified in these four lectins (JAC, PHA-E, RCA, Con-A) and common and unique proteins are presented in Table 5.7.

Table 5.7: The comparison of proteins identified in JAC, PHA-E, RCA and Con-A LAC. Uniprot accessions, Number of proteins common and unique among different lectins, lectin binding and information inferred from this comparison about type of glycans present in these proteins are indicated in the table.

Uniprot accessions	Number of proteins	Gene/Protein Name	Lectin binding	Type of glycans
P25311 P07911	2	Zinc-alpha-2-glycoprotein; Uromodulin	JAC, PHA-E, RCA, Con-A	Oligomannose, Galactose bearing, complex
Q02413 P04745	2	Desmoglien-1; Alpha-amylase 1	JAC, PHA-E, RCA	Galactose bearing, Complex
P01857 O43653 P01024 P01859 P05090 P02787	6	Immunoglobulin heavy constant gamma 1 (G1m marker); Apolipoprotein D; Immunoglobulin heavy constant gamma 2 (G2m marker); Transferrin; Prostate stem cell antigen; Complement C3	JAC, PHA-E, Con-A	Oligomannose, complex
P01876	1	Ig alpha-1 chain C region	JAC, RCA, Con-A	Galactose bearing, Oligomannose
Q08554 Q14574	2	Desmocollin-1; Desmocollin-3	JAC, PHA-E	Oligomannose, complex
P02790 P02675 Q08380 P02649 P01023 Q14624 P01833 P01009 P02679 P08962 P01871 O96009 P01763	13	Immunoglobulin heavy constant mu; Alpha-1-antitrypsin; Hemopexin; Apolipoprotein E; Polymeric immunoglobulin receptor; Fibrinogen beta chain; Inter-alpha inhibitor H4;Glaectin-3 binding protein; Napsin A; CD63; Alpha-2 macroglobulin; fibrinogen gamma chain; Ig heavy chain V-III region WEA	JAC, Con-A	Oligomannose
P80188	1	Neutrophil gelatinase-associated lipocalin	PHA-E, RCA	Galactose bearing, Complex
O75594 Q15517 P09668	3	Peptidoglycan recognition protein 1; Corneodesmosin; Pro-cathepsin H	PHA-E, Con-A	Biantennary complex
P01042 P01011 P41222 P98164 P02671 P0C0L4 P02760 P04196	8	Alpha-1-antichymotrypsin; Kininogen 1; Fibrinogen alpha chain; Alpha-1-microglobulin; Histidine rich glycoprotein; Complement component 4A; Prostaglandin D2 sysnthase; Megalin	JAC	T antigen

P07339	1	Cathepsin D	PHA-E	Complex
P08603 Q8N2E2 P27487	3	Complement factor H; von Willebrand factor D and EGF domain-containing protein; Dipeptidyl peptidase 4	RCA	Galactose bearing
P12273 P51688 Q12805 P21926 P02763 P02751	6	Prolactin-inducible protein; N-sulphoglucosamine sulphohydrolas; EGF-containing fibulin-like extracellular matrix protein 1; CD9 antigen; Alpha-1-acid glycoprotein 1; Fibronectin	Con-A	Alpha-mannose

5.3.3.6 Non-glycosylation proteins identified in our analysis

We have identified 172 proteins in LAC of P200,000g out of which 64 (37%) are glycosylated. Sixty-three percent (63%) of total proteins identified are non-glycosylated. In a previous report, 70% of proteins identified in LAC using *M. amurensis* lectin were found to be non-glycosylated (McDonald *et al.*, 2009). In the same paper it was shown that in the hydrazide chemistry protocol only 30% of the total identified proteins were glycosylated. compared to that study we have an improved identification of glycosylated proteins (66% glycoproteins in hydrazide method compared to 30% in the previous study and 37% glycoproteins in LAC compared to 30% in the previous study). However, a large number of non-glycosylated proteins are still present. In the same study, the authors had identified the proteins non-specifically adsorbing to the stationary phase matrix. When we compared our identification with the non-specifically adsorbing proteins of HeLa cells we found that although, the two samples are completely different, 14 proteins are present in our identifications as well. These 14 proteins are not known to be glycosylated and must be getting enriched by other means such as non-specific adsorption. These proteins are annexin A1 and A2, actin beta and alpha2, alpha enolase 1, glyceraldehyde 3-phosphate dehydrogenase, phosphogluconate dehydrogenase, histones H2A type 1-A and H4, elongation factor 1-alpha 1, 14-3-3 protein eta, pyruvate kinase isozymes M1/M2 and

peroxiredoxin 1 and 2. These proteins in cellular context are very abundant proteins in HeLa cells. However, quantitative data on their amount in exosomes is missing. Notably six of these are nucleotide binding proteins. The ability of these 14 proteins, adsorbing non-specifically to the stationary phase even when they are present in different samples, could be due to their sticky nature (aggregation) or high amounts. However, a clear answer is not known.

Leaving these proteins apart, there are still a number of non-glycosylated proteins (94 proteins) which were detected in LAC. We hypothesized that a majority of these proteins might be interacting with the glycosylated proteins. For this purpose we screened these non-glycosylated proteins for interactions with the glycosylated proteins identified in our sample using protein-protein ineteraction databases I2D and iRefWeb (Turner *et al.*, 2010). If these proteins form a complex with glycosylated proteins they would be detected in the LAC as well. Forty seven unique interactions were found and it was revealed 34 non-glycosyated proteins identified in LAC of P200,000g are known interacting partners of 18 glycoproteins identified in LAC. These interactions are presented in Table 5.8.

Table 5.8: The non-glycosylated proteins identified in LAC which interact with glycoproteins identified in LAC. Protein name and the source of interaction are given in the table.

Non-glycosylated proteins	Glycosylated interacting partner	Source
CystatinA	Cathepsin B	I2D
Ig kappa chain V-II region Cum	Ig Mu chain	I2D
Ig kappa chain V-III region SIE	Ig alpha-1 chain C region	I2D
Ig lambda chain V-I region HA	Ig alpha-1 chain C region	I2D
Ig heavy chain V-III region VH26;	Ig gamma-1 chain C	I2D

Precursor	region	
Ig kappa chain C region	Ig gamma-1 chain C region	I2D
Cytokeratin-14	Ig gamma-1 chain C region	I2D
Cytokeratin-6A	Ig gamma-1 chain C region	I2D
Apolipoprotein A-I	Ig gamma-1 chain C region	I2D
Apolipoprotein A2	Serum albumin; Precursor	I2D
Apolipoprotein C1	Serum albumin; Precursor	I2D
Cystatin-B	Cathepsin D	I2D
Cytokeratin-6B	Serum albumin; Precursor	I2D
Apolipoprotein A-IV	Serum albumin; Precursor	I2D
Annexin A2	Cathepsin B	I2D
Cytokeratin-16	Ig gamma-1 chain C region	I2D
Cytokeratin-13	Serum albumin; Precursor	I2D
Cytokeratin-5	Serum albumin; Precursor	I2D
Glial fibrillary acidic protein	Serum albumin; Precursor	I2D
Junction plakoglobin	Desmoglein-1	I2D
Desmoplakin	Desmocollin-1	I2D
Cornifin-B	Prolactin-inducible protein	I2D
Trypsin-3	Alpha-1-antitrypsin	I2D
Beta-catenin	Megalin	I2D
Cytokeratin-9	Ig gamma-1 chain C region	I2D
Lysozyme C	Alpha-2-macroglobulin	I2D
Hemoglobin beta chain	Ig Mu chain	I2D
Hemoglobin alpha chain	Cathepsin D	I2D
Dermcidin	Serum albumin; Precursor	I2D
Plakophilin-1	Desmoglein-1	I2D
Trypsinogen C	Serum albumin; Precursor	I2D
Apolipoprotein A2	Apolipoprotein D	iRefWeb
Apolipoprotein A2	Complement C4-A	iRefWeb
Apolipoprotein A1	Apolipoprotein B	iRefWeb
Desmoplakin	Desmocollin-1	iRefWeb
Junction plakoglobin	Desmocollin-1	iRefWeb
Ig kappa chain C region	Ig gamma-1 chain C region	iRefWeb

Cystatin-B	Cathepsin D	iRefWeb
Cystatin-A	Cathepsin B	iRefWeb
Cystatin-A	Pro-cathepsin H	iRefWeb
Junction plakoglobin	Desmoglein-1	iRefWeb
Apolipoprotein A1	Haptoglobin	iRefWeb
Cystatin-B	Pro-cathepsin H	iRefWeb
Cystatin-B	Cathepsin B	iRefWeb
Haptoglobin-related protein	Ig mu chain C region	iRefWeb
Junction plakoglobin	Desmocollin-3	iRefWeb
Glyceraldehyde-3-phosphate dehydrogenase	Cathepsin D	iRefWeb

Using above two approaches, we were able to account for the presence of 44% of the total non-glycosylated proteins detected in LAC. The remaining proteins, we speculate, could be present in our identification due to either new protein-protein interactions (not previously documented) or non-specific adsorption to stationary phase matrix.

5.3.3.7 Gene ontology and bioinformatic analysis of total glycoproteins identified by LAC and hydrazide chemistry

Combined total proteins from LAC and hydrazide method identifications (108 unique glycproteins) were submitted for analysis by IPA software. Proteins were annotated for sub-cellular location and molecular function. Four proteins were not mapped by IPA and they were manually classified into sub-cellular localisation and molecular functions categories. Sub-cellular localisation of these 108 glycoproteins is presented in Figure 5.16. For comparison, the total proteome of P200,000g (Presented in Chapter 3) is shown side-by-side. This comparison gives an idea about which categories are represented heavily in glycoproteins compared to total proteome.

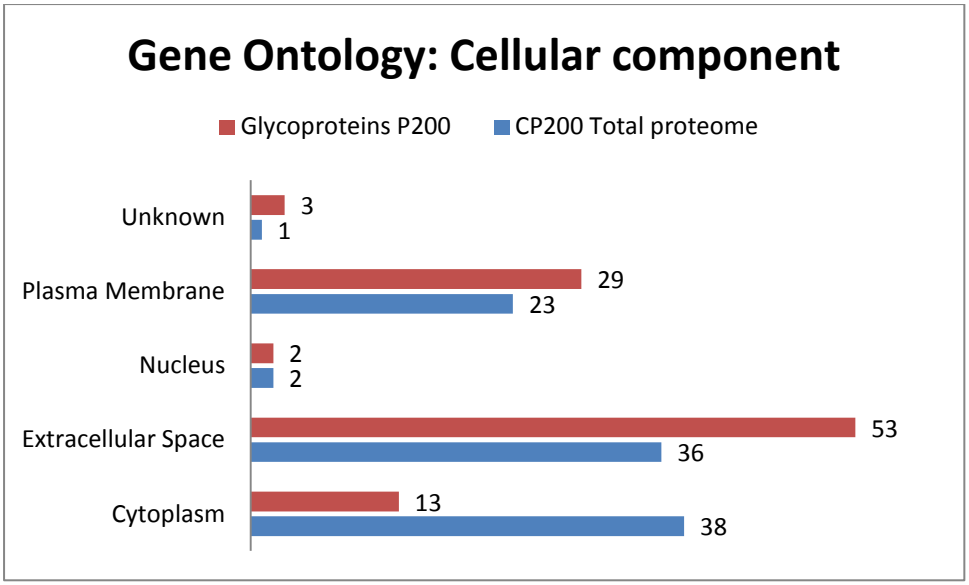

Figure 5.16: Sub-cellular localisation of glycoproteins of exosome and 'exosome-like' vesicles as annotated by IPA software.

The biggest category enriched in the membrane vesicular glycoproteins is extracellular space (53%) and then plasma membrane (29%) and cytoplasm (13%). Extracellular proteins and proteins associated to plasma membrane are over-represented in glycoproteins of membrane vesicles compared to full proteome. However, cytoplasmic proteins in membrane vesicles are heavily under-represented compared to full proteome. This could be due to the fact that intracellular proteins are mostly O-glycosylated which is a small fraction of total glycoproteome where N-glycosylation is the predominant type. Further classification of glycoproteins by molecular functions of the proteins is presented in Figure 5.17.

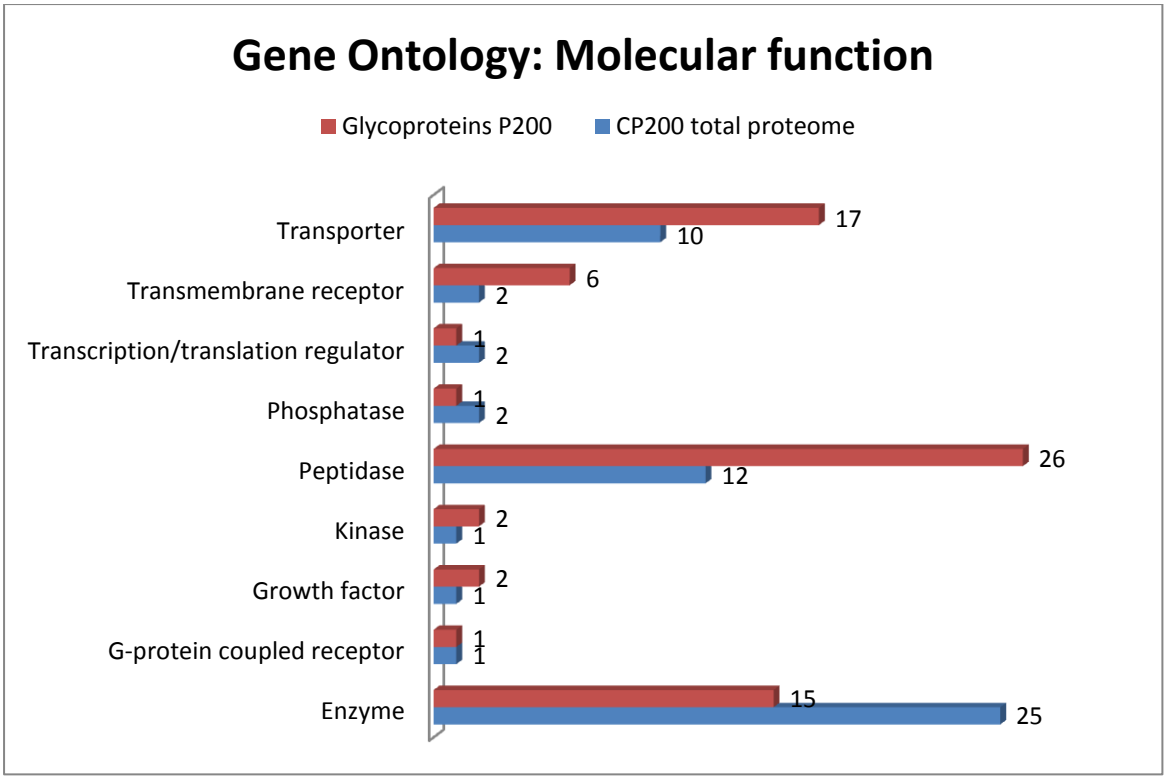

Figure 5.17: The classification of glycoproteins for molecular function categories by IPA.

The biggest category of molecular function is peptidase (26%) followed by transporters (17%). Transmembrane receptors make up 6% of the glycoproteins while enzymes make up another 15%. Compared to total proteome, peptidases (26% compared to 12 in total proteome) and transporter proteins (17% compared to 10% in total proteome) are heavily over-represented in glycoproteins. For example neutrophil gelatinase-associated lipocalin (NGAL) was identified only in glycoprotein study here and not in total proteome in Chapter 3. NGAL is an iron-trafficking protein involved in apoptosis, innate immunity and renal development (Source: Uniprot, accession: P80188). NGAL is also an early biomarker for diabetic nephropathy which appears much earlier than albuminuria (Alter *et al.*, 2012). However, enzymes are under-represented compared to total proteome. Further IPA analysis revealed a set of proteins which are involved in kidney failure and are presented in table 5.9.

Table 5.9: IPA bio-functions analysis revealed a category Kidney failure. Functions annotation, p-Value of enrichment, molecules involved (Gene symbol) and total molecules in the category are indicated.

Category	Functions annotation	p-Value	Molecules	Total molecules
Kidney Failure	Renal failure	1.27 E-05	AGT, ALB, C3, DPP4, F2, IGHG1, KNG1, LCN2, PTGDS	9
Kidney Failure	Chronic renal failure	7.02 E-06	AGT, ALB, C3, DPP4, F2, KNG1	6
Kidney Failure	Failure of kidney	1.58 E-04	AGT, C3, DNASE1, PODXL	4
Kidney Failure	End stage renal disease	2.07 E-04	AGT, ALB, C3, F2, KNG1	5
Kidney Failure	Interstitial fibrosis of kidney	6.53 E-03	AGT, C3, KNG1	3
Kidney Failure	Septic acute kidney injury	2.40 E-03	LCN2	1
Kidney Failure	Acute tubular necrosis	3.85 E-02	LCN2	1
Kidney Failure	End stage renal disease of kidney	3.85 E-02	C3	1
Kidney Failure	Acute renal failure	4.06 E-02	ALB,LCN2	2
Kidney Failure	Ischemic acute renal failure	6.34 E-02	LCN2	1
Kidney Failure	Hepatorenal syndrome	6.95 E-02	ALB	1

Network generation function in the core analysis feature of the IPA generated a network enriched in our glycoprotein identifications. This network enriched in our analysis is involved in neurological diseases, cancer and gastrointestinal diseases. This network is shown in Figure 5.18.

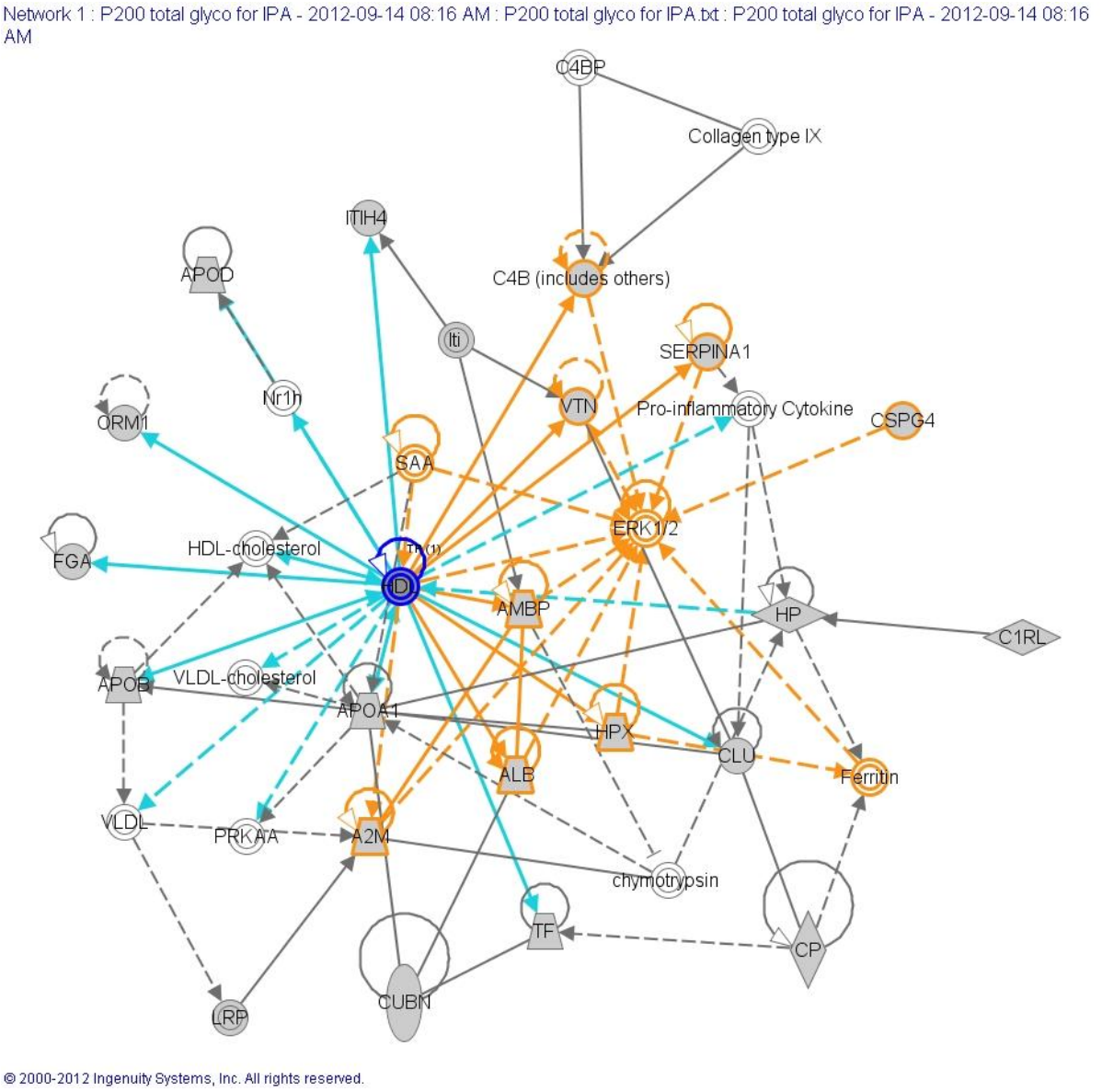

Figure 5.18: Network generated by IPA analysis of exosomal and exosome-like vesicular glycoproteins. Gene symbol of glycoproteins have been used. Proteins in grey shapes are the ones found in our analysis while the ones in white shape are associated with the network but not found in our analysis.

This network shows the proteins which are connected by brown dotted and solid lines are the ones which decrease the phosphorylation of ERK1/2. While the proteins connected by blue lines to the HDL are components of HDL. HDL component proteins are involed in cross-talk with the ERK pathway. It was previously shown that isolated biliary exosomes were able to

induce a decrease of phosphorylated-total ERK1/2 ratio and, subsequently, cholangiocyte proliferation (Masyuk *et al.*, 2010). Therefore, this network enrichment in our dataset could be a platform for future studies validating these targets to establish crosstalk of given exosomal components with the ERK pathway in the recipient cells.

5.3.3.8 Comparison with previous studies on exosome and exosome-like vesicles

We compared our identifications with the previous largest proteomic study on urinary exosomes and exosome-like vesicles (Wang *et al.*, 2011) (Wang). We also combined our identifications from Chapter 2 with Wang and then compared it to the total proteins we identified in glycoprotein analysis.

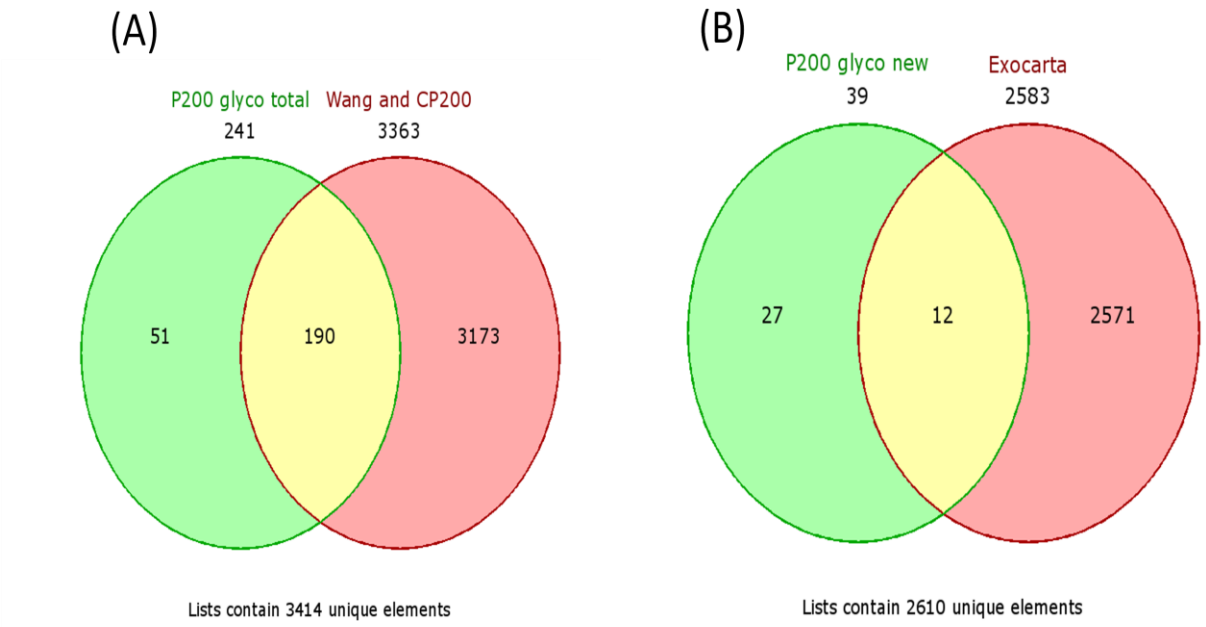

Figure 5.19: Panel A: Comparison of proteins identified in glycoprotein analysis with combined identifications from our Chapter 2 and largest study on exosome and exosome-like vesicles (Wang *et al.*, 2011) (Wang). Panel B: the unique proteins in glycoprotein analysis were compared with Exocarta (Mathivanan *et al.*, 2012) (The biggest database of exosomal proteins from any source, a filter was applied and only human proteins were extracted from Exocarta).

Fifty-one proteins were not found in either our proteomic identifications from chapter 2 or in the Wang *et. al*. study. These 51 proteins were converted from Uniprot accessions to Entrez gene ids which mapped to 39 Ids. These 39 Ids were compared to exocarta human proteins which are in the form of Entrez gene ids. Twenty-seven proteins were found to be unique to our glycoprotein analysis, which have never been described as part of exosomes or 'exosome-like' vesicles. These 27 proteins are presented in Table 5.10.

Table 5.10: The proteins shown for the first time to be part of exosome/'exosome-like' vesicles. Uniprot accessions and gene/protein names are given.

UNIPROT ACCESSION	GENE/PROTEIN NAME	Known/potential glycosylation in Uniprot
P02654	apolipoprotein C-I	No
Q9NZT1	calmodulin-like 5	No
P09668	cathepsin H	Yes
Q9UBR2	cathepsin Z	Yes
Q15517	Corneodesmosin	Yes
P09228	cystatin SA	No
P01037	cystatin SN	No
Q6E0U4	Dermokine	No
P59894	doublecortin domain containing 1	No
O00472	elongation factor, RNA polymerase II, 2	No
P14136	glial fibrillary acidic protein	No
Q68CZ6	HAUS augmin-like complex, subunit 3	No
Q96QV6	histone cluster 1, H2aa	No
B9A064	Immunoglobulin lambda-like polypeptide 5	No
P09914	interferon-induced protein with tetratricopeptide repeats 1	No
Q3SY84	keratin 71	No
Q6ZMR3	lactate dehydrogenase A-like 6A	No
Q7L985	leucine rich repeat and Ig domain containing 2	Yes
Q86SR0	Ly6/neurotoxin 1	No
O60237	protein phosphatase 1, regulatory (inhibitor) subunit 12B	No
Q86SG5	S100 calcium binding protein A7A	No
Q9NQ38	serine peptidase inhibitor, Kazal type 5	No
P22528	small proline-rich protein 1B (cornifin)	No
P35325	small proline-rich protein 2B	No
P22531	small proline-rich protein 2E	No

O94901	unc-84 homolog A (C. elegans)	No
Q8N2E2	von Willebrand factor D and EGF domains	Yes

None of these 27 proteins have been described in exosome or other membrane vesicles to the best of our knowledge. Five proteins out of these 27 are known to be glycosylated which are indicated in the Table 5.10. These proteins, however, could be potential targets to identify their glycosylation status.

5.3.4 Lectin affinity chromatography for isolation of nano-vesicles from urine

If a lectin affinity-based method for enrichment of membrane vesicles from urine is envisaged, one has to assume the possibility of contamination with high abundance proteins and avoid their enrichment. Two major proteins of urine are THP and albumin and THP is highly glycosylated (30% of total weight are glycans) while albumin is usually not glycosylated. Although some glycosylated variants of albumin are known, it can be assumed that albumin should not be enriched in elution of lectin affinity chromatography, as the frequency of glycosylated variants is very low. However, glycosylated variants of human serum albumin such as redhill have been reported (Kragh-Hansen, Donaldson & Jensen, 2001). Based on the FLLA (Figure 5.2) and lectin blotting results (Figure 5.4), it was decided that Jacalin would be used for lectin-affinity isolation of membrane vesicles as THP did not bind to the Jacalin and P200,000g and P18,000g bound strongly to it. Moreover, in MS analysis of P200,000g extract 50% of total glycoproteins identified were membrane proteins as annotated by SOSUI web server (Hirokawa, Boon-Chieng & Mitaku, 1998). Also, the signal in lectin blot and FLLA for SN200,000g was much lower. Biotinylated Jacalin was immobilised on streptavidin agarose and subsequently Jacalin-agarose was used for lectin affinity chromatography (LAC). It was found that THP was still a major protein in the elution

of LAC (Figure 5.20 B). To remove THP, diatomaceous earth filter (DEF) was used to filter urine.

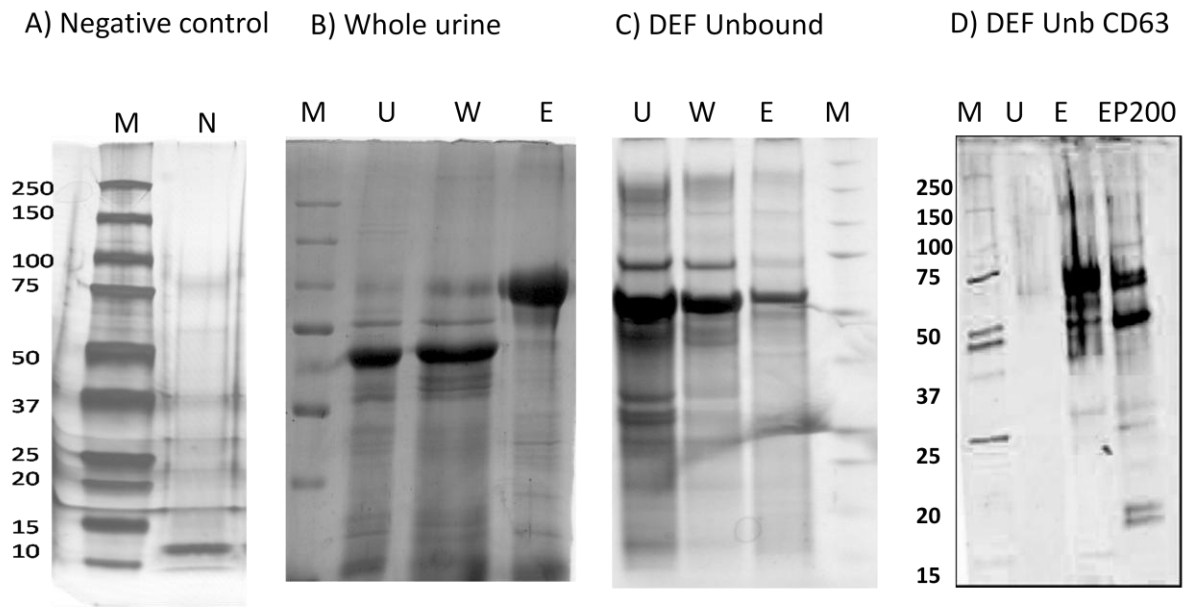

Figure 5.20: SDS-PAGE of non-bound, washing and elution profile of Jacalin affinity chromatography of A) Negative control, B) whole urine, C) DEF non-bound, and Western blot D) of DEF non-bound using anti-CD63 antibodies. M: Molecular weight marker in kDa; N: Negative control, chromatography performed on only streptavidin-agarose without immobilised-Jacalin; U: Non-bound; W: Washing; E: Elution; EP200: Pellet 200,000g of elution of Jacalin-agarose.

DEF non-bound material, mostly free of THP was applied to Jacalin-agarose and eluate was checked for membrane vesicles/exosomal marker CD63. Eluate from Jacalin-agarose was centrifuged at 200,000g to pellet down any membrane vesicles if any in the elution. Whole eluate and P200,000g of eluate had CD63 much more enriched when compared to non-bound fraction. It suggests that Jacalin chromatography of DEF filtered urine can enrich membrane vesicles without much contamination of soluble proteins like THP. To confirm that Jacalin-

agarose enriches membrane vesicles by specific interaction with glycans and to establish a negative control, streptavidin-agarose without immboilised Jacalin was used for chromatography with whole urine. The same protocol as Jacalin-agarose chromatography was followed and the eluate was run on SDS-PAGE. Negative control, when compared to the SDS-PAGE pattern of the Jacalin-agarose chromatography eluate (Figure 5.20, N in Panel A versus E of Panel B & C), demonstrates that enrichment of vesicles is specific to Jacalin binding to the glycans on the surface of membrane vesicles.

However, DEF filtration is a lengthy procedure and THP can entrap membrane vesicles (Fernandez-Llama *et al.*, 2010) resulting in loss of vesicles in the subsequent chromatography step. To avoid this shortcoming and develop a method which is simple with less number of steps, we tried to change the binding buffer of the Jacalin chromatography. HEPES buffer saline was originally used and it can be seen in LAC elution as well that THP is present in the elution (Figure 5.9, Panel A, lane 7). THP is known to polymerise at high ionic strength therefore we changed the binding buffer to 10mM phosphate buffer (pH 7.4) withoutany salt. Polymerisation of THP might be the reason why it was getting enriched in high amounts with Jacalin elution. We were able to eliminate excess THP becoming enriched with fraction eluted from Jacalin-agarose (Figure 5.21).

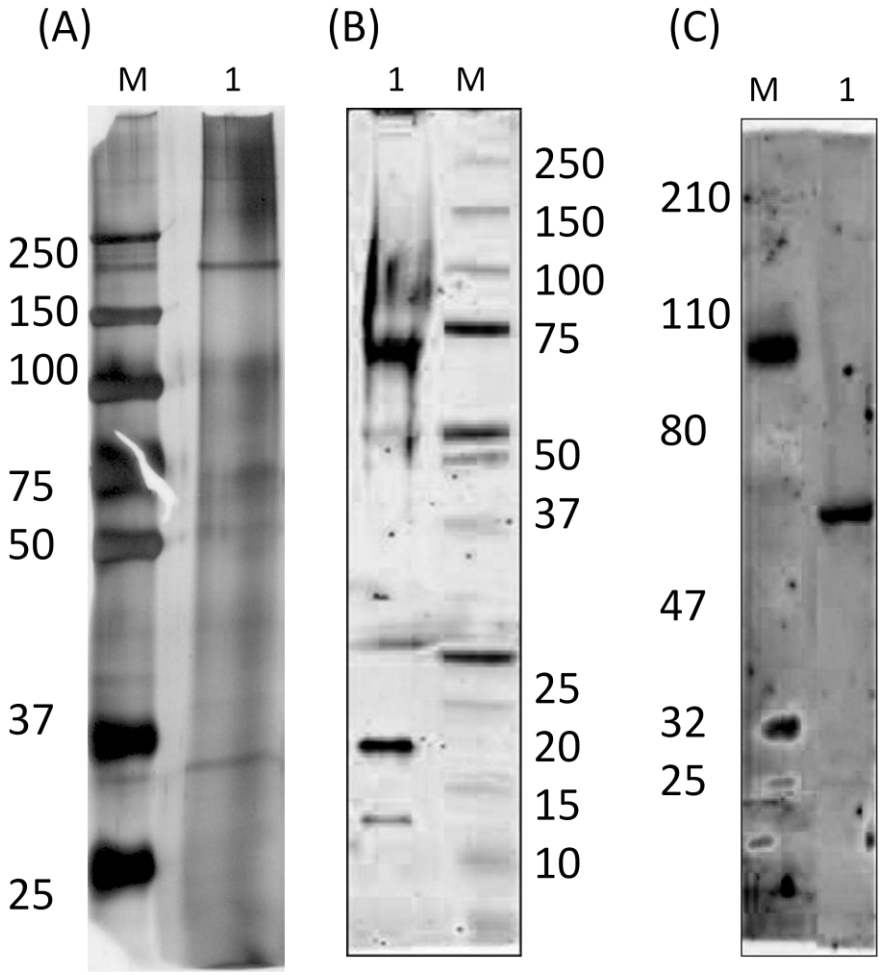

Figure 5.21: SDS-PAGE (Panel A) of Jacalin chromatography using phosphate buffer as the binding buffer without any salt. 1: Jacalin elution M: molecular weight markers in kDa. Western blot of Jacalin elution (Panel B) using anti-CD63 antibodies. 1: Jacalin elution M: molecular weight markers in kDa. Western blot of Jacalin elution (Panel C) using anti-TSG101 antibodies. 1: Elution fraction (vesicles) from the Jacalin-agarose M: molecular weight markers in kDa.

It can be seen in Figure 5.21 that THP (~100 kDa) is not a major contaminant anymore in the Jacalin elution (Panel A) without having any effect on CD63 enrichment (Panel B). TSG101 (intra-vesicular exosomal marker) was also found in Jacalin elution although in very low amounts as the image had to be acquired at high intensity to visualise the band. This elution was further characterised by transmission electron microscopy (TEM) to conclusively prove

that it contains membrane vesicles. Moreover, TEM analysis would also reveal the size and morphology of vesicles enriched by Jacalin from minimally processed urine.

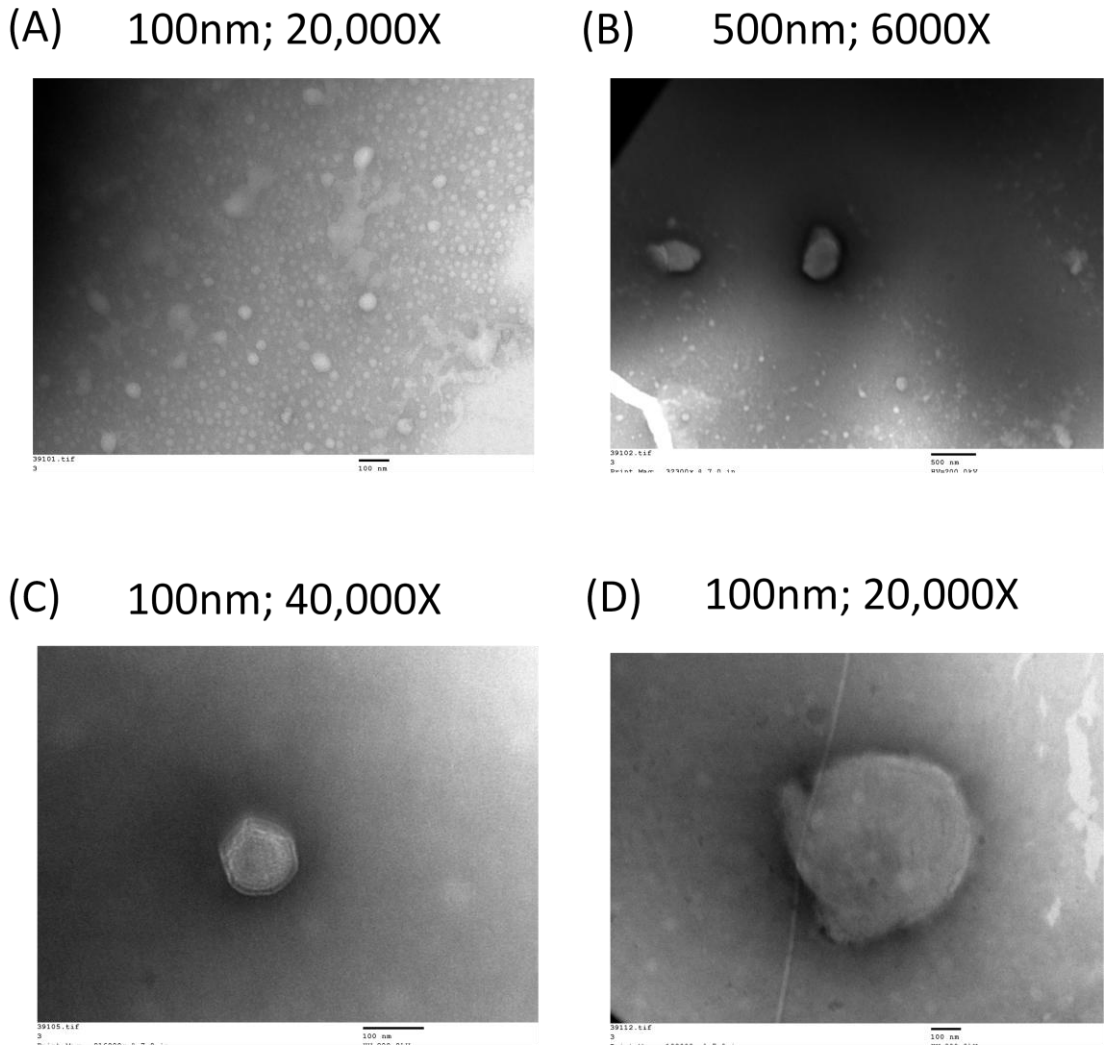

Figure 5.22: TEM analysis of vesicles eluted from Jacalin-agarose. Some representative pictures are presented here. Size of the scale bar and direct magnification is given on top of each panel.

Figure 5.22 panel A shows mainly vesicles which are 50-70 nm while panel B shows bigger vesicles which are 378 and 389nm. Panel C shows a 126nm vesicle while panel D shows a 639nm vesicle. We acquired multiple TEM images and size of the vesicles was calculated by

ImageJ software. We present a size distribution of vesicles present in 16 representative vesicles pictures.

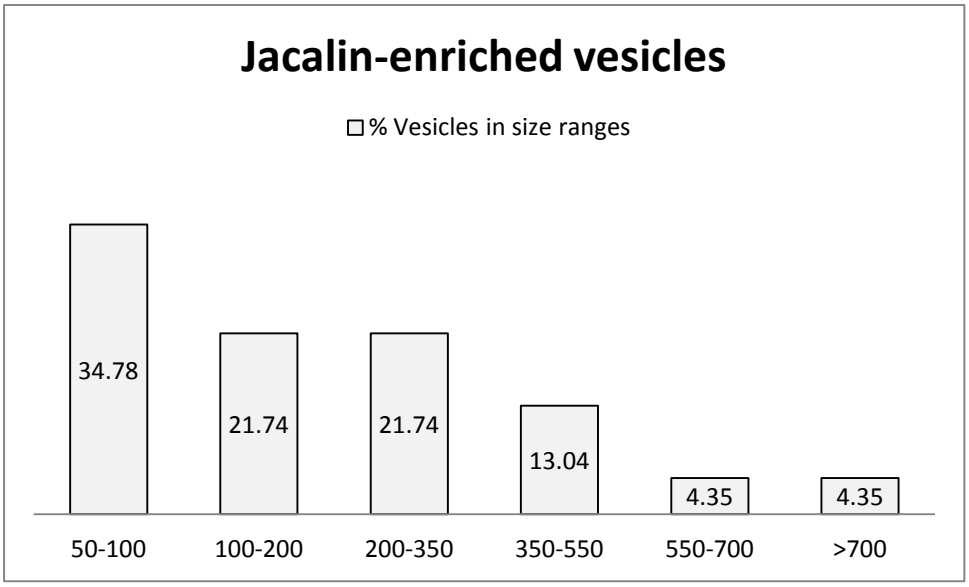

Figure 5.23: Presents the size-distribution of vesicles found in vesicle eluted from Jacalin-agarose. 16 images were analysed by ImageJ software and size of all the vesicles were calculated. X-axis is the size of vesicles in nm and Y-axis is % vesicles in a given size range.

We mostly found vesicles from 50 to 550 nm size but vesicles bigger than 550 were few. It can be said that Jacalin enriched vesicles are mixtures of exosomes and bigger microvesicles.

5.4 Discussion

The purpose of our study was to use robust and exhaustively described methods for novel characterization of urinary vesicles for their glycoconjugates profile. Additionally this information was used to devise a new method to rapidly isolate urinary membrane vesicles. FLLA results showed that urinary membrane vesicles (P200,000g) are enriched in multiple glycan epitopes including complex N-linked glycans (PHA-E, PHA-L), high mannose structures (Con-A, PSA), Sia α 2-6Gal ; GalNAc and multivalent Sia (SNA, WGA), LacNAc (RCA_{120}), GalNAcα-Thr/Ser (Tn) (JAC) and GalNAcαThr/Ser (Tn) ; GalNAcα1-3GalNAc (DBA). The WGA signal, compared to SWGA, was much higher indicating that WGA was binding mostly to sialic acids and not to N-acetylglucosamine. sWGA, as reported by the Vector labs datasheet, binds only to N-acetylglucosamine and not to sialic acids. This is complemented by high signal of SNA which binds Sia α 2-6Gal. The signal for MAL-II was very low suggesting that sialic acids on membrane vesicle surfaces are mainly conjugated by α 2-6.

Many of these glycan epitopes have been previously recorded to be present on membrane vesicles from H9 T-cell line as evidenced by binding of their corresponding lectins (Con-A, PSA, PHA-E, PHA-L, RCA, WGA, SNA) (Krishnamoorthy *et al.,* 2009). It was also shown in the study that these epitopes are enriched in membrane vesicles as compared to parent cell plasma membrane. Another study using a wide variety of cell lines (SkMel-5, HT29, HCT-15, H9, SupT1 and Jurkat-Tat-CCR5 cell lines) and human breast milk as source of membrane vesicles confirmed most of these epitopes to be common to membrane vesicles from different sources (Batista *et al.,* 2011). This suggests that membrane vesicles originate from specialised compartments of cells and glycosylation may play an important role in sorting glycoproteins to these compartments. Accordingly, it is known that, specific glycoforms of prion protein are sorted into exosomes over other forms (Vella *et al.,* 2007).

Our study appears to be the first one undertaken on glycosylation of urinary membrane vesicles. FLLA could be employed to discover disease-specific differences in membrane vesicles glycosylation among different well defined patient samples.

Taking the study further, nine lectins belonging to seven different classes based on their binding specificity were employed on P200,000g and P18,000g extract (prepared using beta-octyl glucoside{BOG) to enrich constituent glycoproteins and identify them using MS analysis. BOG was used to prepare the extract as it dissolves the lipid-rafts (Garner, Smith & Hooper, 2008). One hundred and seventy two proteins excluding keratin were identified in P200,000 and 97 proteins in P18,000g. Out of these 172 proteins in P200,000g, 64 proteins are known glycoproteins (37%). When these 64 proteins were compared with 37 known glycoproteins identified in P18,000g only 5 proteins were unique to P18,000g (13%). High number of proteins common to both pellets suggest that P18,000g also consists of membrane vesicles which due to physico-chemical factors or entrapment in THP fibrils pellet down at low speed. It is to be noted that a similar trend was also noted in Chapter 2 in proteomic analysis of both these pellets. These 5 proteins unique to P18,000g are non-secretory ribonuclease (EDN), clusterin, CD59, mucin 1 and inter-alpha inhibitor H1. EDN was previously found to be present in urine and according to Uniprot, it is N-linked glycosylated at 5 locations and C-type glycosylated at one site. EDN is an anti-microbial proteins present in eosinophil granules and is chemotactic for dendritic cells (Yang *et al.,* 2003). The role of extensive glycosylation of this protein in its biological functions is not known.

Four of these 5 proteins have previously been shown to be secreted with exosomes in urine (Pisitkun, Shen & Knepper, 2004). However, Inter-alpha-trypsin inhibitor heavy chain H1 is not reported be part of exosomes from any source to the best of our knowledge (Absent in Exocarta (Mathivanan *et al.,* 2012)). This protein was also identified in our total proteome analysis in Chapter 3. Total 64 known glycoproteins identified in P200,000g were compared

with each other and a crude information about glycans carried by these proteins was revealed (Table 5.4). When results from the hydrazide method were combined with LAC identifications, a total of 108 known glycoproteins were found. This study would serve as a platform for future studies on glycosylation of exosomal proteins. A number of sialic acid containing glycoproteins and core and α1-2 fucosylated proteins were also identified which are presented in results sections. Changes in sialic acid content in the glomerular capillary walls of the diabetic nephropathy patients compared to healthy tissues were observed previously (Tomino *et al.,* 1988). In this study the binding of WGA decreased with neuraminidase treatment suggesting glycan epitopes contained sialic acids. As previously discussed, WGA signal in FLLA of P200,000 is mainly because of sialic acid content (SWGA signal is much less ruling out N-aceyltglucosamine).

The changes in glycan content or amount in kindey can be reflected on membrane vesicle surface glycosylation which could be detected using FLLA. This is only a hypothesis until confirmed with patient samples. However, if this comes true, it can be further expanded to plate based hybrid lectin-antibody assays after recognising the glycproteins carrying these glycans which might change with disease in kidney and reflected in membrane vesicles. Identification of exosomal surface glycome is a first step towards that which we have performed using hydrazide chemistry as detailed in results section. Thirty-seven membrane glycoproteins including podocalyxcin, GPCR, and solute carrier transporters were identified. These proteins reflect the majority of exosomal surface glycome.

The major category of gene ontology of glycoproteins in cellular component category was extracellular proteins followed by plasma membrane and then cytoplasm. Extracellular proteins peripherally or transiently attach with plasma membrane and many of them are expected to end up in endosomes which might prime them up for secretion with exosomes. Twenty seven proteins were identified to be part of exosome/exosome-like vesicles for the

first time out of which 5 are known glycoproteins. This suggests that lectin enrichment is efficient and can be used to enrich low abundance proteins from urinary membrane vesicles for more thorough proteomic characterization. Lectin affinity chromatography, although an efficient method of glycoprotein enrichment, sometimes catches non-glycosylated proteins which may be enriched as an artefact. For example serum albumin was also identified in our list and this could be due to the fact that albumin, being a transporter, is associated with multiple proteins. However, glycosylated variants of albumin have been reported (Kragh-Hansen, Donaldson & Jensen, 2001). Moreover, albumin was also found in all samples from hydrazide chemistry method to enrich glycproteins. We have used a stringent method using hydrazide chemistry which included washing with 8M urea and 1.5M NaCl three times after the oxidised glycoproteins had bound to the hydrazide resin. This should result in all the non-glycosylated or non-glycated proteins to be washed out. On-column trypsin digestion would result in release of peptides of only glycoproteins. Therefore, a possibility of albumin being either glycosylated or glycated can not be excluded. However, oxidised proteins could bind to hydrazide resin as well (Mirzaei & Regnier, 2007) and these proteins could be non-glycoyslated. Albumin is known to be prone to oxidation at multiple sites and this could be the reason why it may bind to hydrazide resin. Moreover oxidation induced cross-linking of proteins has also been reported and this could present another mechanism by which albumin and other non-glycosylated proteins could be retained on hydrazide resin despite 8M Urea and 1.5M NaCl washes.

Deducing from the FLLA results, we noticed that Jacalin bound strongly to P200,000g and P18,000g but very little to SN200,000g and not at all to purified THP (Figure 2). Also, in lectin blotting Jacalin did not bind much to SN200,000g however some bands in P200,000 were apparent (Figure 5.4). This led us to test whether Jacalin chromatography can be used as a simple method to enrich urinary membrane vesicles. This would not require any

sophisticated instruments like an ultracentrifuge and would also save the labour of setting up tedious density gradient procedures. For such a purpose the method appears to work well. To achieve the results we first centrifuged whole human urine (50 mL) at 2000g and supernatant (SN2000g) was concentrated (using speed vacuum concentrator to 2-3 mL or Vivaspin 300kDa MWCO) and used for the chromatography step. We are aware that while using the Vivaspin, we might lose some membrane vesicles on the filter surface or entrapped in THP polymers sticking to the filter. For removing the interference of THP poylmerisation at this step, no buffer was added and whole urine was concentrated. This would minimise the THP polymerization. And low speed used in the centrifugation for Vivapsin concentration (5000rpm) would mean most vesicles would not pellet down at this speed and remain in solution. Introduction of Vivaspin at this step was done to minimise 50mL urine to a manageable amount as the resin volume was small (1mL). Centrifugation (2000g) before Vivaspin concentration would be expected to get rid of any dead or shed cells which may be present in urine.

It was found that, although Jacalin did not show any binding to THP in FLLA, THP was still a major protein in the Jacalin elution. It could be due to the reason that THP polymerises in the binding buffer and these polymers interact non-specifically with the agarose beads. THP polymerization has been previously reported (SERAFINICESSI *et al.,* 1989). To remove contamination of THP, SN2000g was filtered using DEF. DEF removed the majority of THP from the urine and non-bound of DEF was applied to chromatography with Jacalin-agarose. This removed the problem of THP contamination and enrichment of the membrane vesicles marker CD63 was observed in elution fraction compared to non-bound fraction where it was almost undetectable. Presence of membrane vesicles in the elution of Jacalin was confirmed by centrifuging elution fraction of Jacalin at 200,000g and membrane vesicles/exosomal marker CD63 was noticed to be pelleting down at this speed. However, DEF filtration is a

lengthy procedure and impractical for clinical settings. Moreover, the DEF filtration method relies on polymerization of THP and it has been reported that membrane vesicles are entrapped in THP polymers (Fernandez-Llama *et al.,* 2010). This would result in potential loss of vesicles during filtration on DEF. We looked at other ways of removing THP contamination and it was found that getting rid of the salt from binding buffer (changed from HEPES buffer saline to phosphate buffer {pH7.4) solves this problem and THP was not found in elution anymore. TSG101 in small amounts was also found in the elution. TEM analysis revealed that Jacalin enriches a mixture of vesicles including smaller vesicles (same size as exosomes) and bigger microvesicles. This vesicle population might be carrying specific glycan epitopes on their surface. In the future a use for this specific fraction of membrane vesicles can be found. This method can be performed with inexpensive laboratory instruments.

Finally, 108 glycoproteins were identified in this study including the surface glycoproteins of exosomes and microvesicles. IPA analysis revealed that multiple proteins are involved in kidney size and morphology maintenance as well as some involved in kidney failure and other pathologies. Secretion of such clinically important proteins with membrane vesicles is of interest. This suggests the potential clinical application of FLLA method as well as glycoprotein enrichment protocols although much further work would be needed for such an application. Moreover, a simple method for microvesicle enrichment was also developed which does not require expensive instrumentation or specialised skills.

5.6 Reference

Adachi, J., Kumar, C., Zhang, Y., Olsen, J.V. & Mann, M. (2006), "The human urinary proteome contains more than 1500 proteins, including a large proportion of membrane proteins", *GENOME BIOLOGY*, vol. 7, no. 9, pp. R80.

Alter, M.L., Kretschmer, A., Von Websky, K., Tsuprykov', O., Reichetzeder, C., Simon, A., Stasch, J. & Hocher, B. (2012), "Early Urinary and Plasma Biomarkers for Experimental Diabetic Nephropathy", *CLINICAL LABORATORY*, vol. 58, no. 7-8, pp. 659-671.

Barres, C., Blanc, L., Bette-Bobillo, P., Andre, S., Mamoun, R., Gabius, H. & Vidal, M. (2010), "Galectin-5 is bound onto the surface of rat reticulocyte exosomes and modulates vesicle uptake by macrophages", *BLOOD*, vol. 115, no. 3, pp. 696-705.

Batista, B.S., Eng, W.S., Pilobello, K.T., Hendricks-Munoz, K.D. & Mahal, L.K. (2011), "Identification of a Conserved Glycan Signature for Microvesicles", *JOURNAL OF PROTEOME RESEARCH*, vol. 10, no. 10, pp. 4624-4633.

Bhatia, P. & Mukhopadhyay, A. (1999), "Protein glycosylation: Implications for in vivo functions and therapeutic applications", *ADVANCES IN BIOCHEMICAL ENGINEERING/BIOTECHNOLOGY*, vol. 64, pp. 155-201.

Bourne, Y., Astoul, C., Zamboni, V., Peumans, W., Menu-Bouaouiche, L., Van Damme, E., Barre, A. & Rouge, P. (2002), "Structural basis for the unusual carbohydrate-binding specificity of jacalin towards galactose and mannose", *BIOCHEMICAL JOURNAL*, vol. 364, no. Part 1, pp. 173-180.

Cirio, M.C., Hui, Z., Haldin, C.E., Cosentino, C.C., Stuckenholz, C., Chen, X., Hong, S., Dawid, I.B. & Hukriede, N.A. (2011), "Lhx1 Is Required for Specification of the Renal Progenitor Cell Field", *PLOS ONE*, vol. 6, no. 4.

Dai, Z., Fan, J., Liu, Y., Zhou, J., Bai, D., Tan, C., Guo, K., Zhang, Y., Zhao, Y. & Yang, P. (2007), "Identification and analysis of alpha 1,6-fucosylated proteins in human normal liver tissues by a target glycoproteomic approach", *ELECTROPHORESIS*, vol. 28, no. 23, SI, pp. 4382-4391.

Debray, H., Decout, D., Strecker, G., Spik, G. & Montreuil, J. (1981), "Specificity of Twelve Lectins Towards Oligosaccharides and Glycopeptides Related to N-Glycosylproteins", *EUROPEAN JOURNAL OF BIOCHEMISTRY*, vol. 117, no. 1, pp. 41-51.

Escrevente, C., Keller, S., Altevogt, P. & Costa, J. (2011), "Interaction and uptake of exosomes by ovarian cancer cells", *BMC CANCER*, vol. 11, pp. 108.

Fernandez-Llama, P., Khositseth, S., Gonzales, P.A., Star, R.A., Pisitkun, T. & Knepper, M.A. (2010), "Tamm-Horsfall protein and urinary exosome isolation", *JOURNAL OF HYPERTENSION*, vol. 28, no. A, pp. E164.

Gabius, H.J. & Gabius, S. (1997), *Glycosciences: status and perspectives*, Chapman \& Hall.

Garner, A.E., Smith, D.A. & Hooper, N.M. (2008), "Visualization of detergent solubilization of membranes: Implications for the isolation of rafts", *BIOPHYSICAL JOURNAL*, vol. 94, no. 4, pp. 1326-1340.

Halim, A., Nilsson, J., Ruetschi, U., Hesse, C. & Larson, G. (2012), "Human Urinary Glycoproteomics; Attachment Site Specific Analysis of N- and O-Linked Glycosylations by CID and ECD", *MOLECULAR \& CELLULAR PROTEOMICS*, vol. 11, no. 4, pp. M111.

Hao, S., Bai, O., Li, F., Yuan, J., Laferte, S. & Xiang, J. (2007), "Mature dendritic cells pulsed with exosomes stimulate efficient cytotoxic T-lymphocyte responses and antitumour immunity", *IMMUNOLOGY*, vol. 120, no. 1, pp. 90-102.

Hirokawa, T., Boon-Chieng, S. & Mitaku, S. (1998), "SOSUI: classification and secondary structure prediction system for membrane proteins", *BIOINFORMATICS*, vol. 14, no. 4, pp. 378-379.

Hsu, K. & Mahal, L.K. (2006), "A lectin microarray approach for the rapid analysis of bacterial glycans", *NATURE PROTOCOLS*, vol. 1, no. 2, pp. 543-549.

Huet, G., Gouyer, V., Delacour, D., Richet, C., Zanetta, J., Delannoy, P. & Degand, P. (2003), "Involvement of glycosylation in the intracellular trafficking of glycoproteins in polarized epithelial cells", *BIOCHIMIE*, vol. 85, no. 3-4, pp. 323-330.

Kragh-Hansen, U., Donaldson, D. & Jensen, P. (2001), "The glycan structure of albumin Redhill, a glycosylated variant of human serum albumin", *BIOCHIMICA ET BIOPHYSICA ACTA-PROTEIN STRUCTURE AND MOLECULAR ENZYMOLOGY*, vol. 1550, no. 1, pp. 20-26.

Krishnamoorthy, L., Bess, J.,Julian W., Preston, A.B., Nagashima, K. & Mahal, L.K. (2009), "HIV-1 and microvesicles from T cells share a common glycome, arguing for a common origin", *NATURE CHEMICAL BIOLOGY*, vol. 5, no. 4, pp. 244-250.

Larsen, M.R., Jensen, S.S., Jakobsen, L.A. & Heegaard, N.H.H. (2007), "Exploring the sialiome using titanium dioxide chromatography and mass spectrometry", *MOLECULAR \& CELLULAR PROTEOMICS*, vol. 6, no. 10, pp. 1778-1787.

Liener, I.E., Sharon, N. & Goldstein, I.J. (1986), *The Lectins: properties, functions, and applications in biology and medicine*, Academic Press.

Masyuk, A.I., Huang, B.Q., Ward, C.J., Gradilone, S.A., Banales, J.M., Masyuk, T.V., Radtke, B., Splinter, P.L. & LaRusso, N.F. (2010), "Biliary exosomes influence cholangiocyte regulatory mechanisms and proliferation through interaction with primary cilia", *AMERICAN JOURNAL OF PHYSIOLOGY-GASTROINTESTINAL AND LIVER PHYSIOLOGY*, vol. 299, no. 4, pp. G990-G999.

Mathivanan, S., Fahner, C.J., Reid, G.E. & Simpson, R.J. (2012), "ExoCarta 2012: database of exosomal proteins, RNA and lipids", *NUCLEIC ACIDS RESEARCH*, vol. 40, no. D1, pp. D1241-D1244.

McDonald, C.A., Yang, J.Y., Marathe, V., Yen, T. & Macher, B.A. (2009), "Combining Results from Lectin Affinity Chromatography and Glycocapture Approaches Substantially Improves the Coverage of the Glycoproteome", *MOLECULAR \& CELLULAR PROTEOMICS*, vol. 8, no. 2, pp. 287-301.

Mirzaei, H. & Regnier, F. (2007), "Identification of yeast oxidized proteins - Chromatographic top-down approach for identification of carbonylated, fragmented and cross-linked proteins in yeast", *JOURNAL OF CHROMATOGRAPHY A*, vol. 1141, no. 1, pp. 22-31.

Muinelo-Romay, L., Villar-Portela, S., Cuevas, E., Gil-Martin, E. & Fernandez-Briera, A. (2011), "Identification of alpha(1,6) fucosylated proteins differentially expressed in human colorectal cancer", *BMC CANCER*, vol. 11, pp. 508.

Musante, L., Saraswat, M., Duriez, E., Byrne, B., Ravida, A., Domon, B. & Holthofer, H. (2012), "Biochemical and Physical Characterisation of Urinary Nanovesicles following CHAPS Treatment", *PLOS ONE*, vol. 7, no. 7, pp. e37279.

Nieuwland, R. & Sturk, A. (2010), "Why do cells release vesicles?", *THROMBOSIS RESEARCH*, vol. 125, Supplement 1, no. 0, pp. S49.

Nilsson, J., Ruetschi, U., Halim, A., Hesse, C., Carlsohn, E., Brinkmalm, G. & Larson, G. (2009), "Enrichment of glycopeptides for glycan structure and attachment site identification", *NATURE METHODS*, vol. 6, no. 11, pp. 809-U26.

Nishioka, J., Ning, M., Hayashi, T. & Suzuki, K. (1998), "Protein C inhibitor secreted from activated platelets efficiently inhibits activated protein C on phosphatidylethanolamine of platelet membrane and microvesicles", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 273, no. 18, pp. 11281-11287.

Pilcibello, K.T., Slawek, D.E. & Mahal, L.K. (2007), "A ratiometric lectin microarray approach to analysis of the dynamic mammalian glycome", *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, vol. 104, no. 28, pp. 11534-11539.

Pisitkun, T., Shen, R. & Knepper, M. (2004), "Identification and proteomic profiling of exosomes in human urine", *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, vol. 101, no. 36, pp. 13368-13373.

Pisitkun, T., Johnstone, R. & Knepper, M.A. (2006), "Discovery of urinary biomarkers", *MOLECULAR \& CELLULAR PROTEOMICS*, vol. 5, no. 10, pp. 1760-1771.

Reiter, R., Gu, Z., Watabe, T., Thomas, G., Szigeti, K., Davis, E., Wahl, M., Nisitani, S., Yamashiro, J., Le Beau, M., Loda, M. & Witte, O. (1998), "Prostate stem cell antigen: A cell surface marker overexpressed in prostate cancer", *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, vol. 95, no. 4, pp. 1735-1740.

Serafinicessi, F., Bellabarba, G., Malagolini, N. & Dallolio, F. (1989), "Rapid isolation of Tamm-Horsfall glycoprotein (uromodulin) from human-urine", *JOURNAL OF IMMUNOLOGICAL METHODS*, vol. 120, no. 2, pp. 185-189.

Shipp, M. & Look, A. (1993), "Hematopoietic differentiation antigens that are membrane-associated enzymes - cutting is the key", *BLOOD*, vol. 82, no. 4, pp. 1052-1070.

Tomino, Y., Inoue, W., Watanabe, S., Yagame, M., Eguchi, K., Nomoto, Y. & Sakai, H. (1988), "Detection of glomerular sialic acids in patients with diabetic nephropathy", *AMERICAN JOURNAL OF NEPHROLOGY*, vol. 8, no. 1, pp. 21-26.

Turner, B., Razick, S., Turinsky, A.L., Vlasblom, J., Crowdy, E.K., Cho, E., Morrison, K., Donaldson, I.M. & Wodak, S.J. (2010), "iRefWeb: interactive analysis of consolidated protein interaction data and their supporting evidence", *DATABASE-THE JOURNAL OF BIOLOGICAL DATABASES AND CURATION*, vol. 2010, pp. baq023.

Ueno, T., Linder, S., Na, C., Rice, W., Johansson, J. & Weaver, T. (2004), "Processing of pulmonary surfactant protein B by napsin and cathepsin H", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 279, no. 16, pp. 16178-16184.

Vella, L.J., Sharples, R.A., Lawson, V.A., Masters, C.L., Cappai, R. & Hill, A.F. (2007), "Packaging of prions into exosomes is associated with a novel pathway of PrP processing", *JOURNAL OF PATHOLOGY*, vol. 211, no. 5, pp. 582-590.

Wang, L., Li, F., Sun, W., Wu, S., Wang, X., Zhang, L., Zheng, D., Wang, J. & Gao, Y. (2006), "Concanavalin A-captured glycoproteins in healthy human urine", *MOLECULAR \& CELLULAR PROTEOMICS*, vol. 5, no. 3, pp. 560-562.

Wang, Y., Hare, T., Won, B., Stowell, C., Scanlin, T., Glick, M., Hard, K., Vankuik, J. & Vliegenthart, J. (1990), "Additional fucosyl residues on membrane-glycoproteins but not a secreted glycoprotein from cystic-fibrosis fibroblasts", *CLINICA CHIMICA ACTA*, vol. 188, no. 3, pp. 193-210.

Wang, Z., Hill, S., Luther, J.M., Hachey, D.L. & Schey, K.L. (2011), "Proteomic analysis of urine exosomes by multidimensional protein identification technology (MudPIT)", *PROTEOMICS, vol. 12, no. 2*, pp. 329-338.

Wu, T., Hsieh, S., Li, K., Wu, C., Yu, C., Yang, A. & Tsai, C. (2008), "Altered glycosylation of Tamm-Horsfall glycoprotein derived from renal allograft recipients leads to changes in its biological function", *TRANSPLANT IMMUNOLOGY*, vol. 18, no. 3, pp. 237-245.

Yang, D., Rosenberg, H., Chen, Q., Dyer, K., Kurosaka, K. & Oppenheim, J. (2003), "Eosinophil-derived neurotoxin (EDN), an antimicrobial protein with chemotactic activities for dendritic cells", *BLOOD*, vol. 102, no. 9, pp. 3396-3403.

Yaylayan, V. (2003), "Recent advances in the chemistry of Strecker degradation and Amadori rearrangement: Implications to aroma and color formation", *FOOD SCIENCE AND TECHNOLOGY RESEARCH*, vol. 9, no. 1, pp. 1-6.

Zhang, Q., Tang, N., Brock, J.W.C., Mottaz, H.M., Ames, J.M., Baynes, J.W., Smith, R.D. & Metz, T.O. (2007), "Enrichment and analysis of nonenzymatically glycated peptides: Boronate affinity chromatography coupled with electron-transfer dissociation mass spectrometry", *JOURNAL OF PROTEOME RESEARCH*, vol. 6, no. 6, pp. 2323-2330.

CHAPTER 6

PURIFICATION AND IDENTIFICATION OF PALMITOYLATED PROTEINS IN HUMAN URINARY MEMBRANE VESICLES

6.1 Introduction

Post-translational modification (PTM) of proteins can dictate their localization or function or both (Walsh, Garneau-Tsodikova & Gatto, 2005). Lipidation of cysteine residues has emerged as previously less appreciated but recently found to be a widespread PTM. Arguably, the most widespread of the lipidation PTM is palmitoylation of specific cysteine residues. This requires C_{16} fatty-acyl CoA as donors for acyltransferases which transfer the C_{16} to the Cys residue (S-palmitoylation) (Bijlmakers & Marsh, 2003). The identity of enzymes responsible for S-palmitoylation has not been fully defined in contrast to the N-myristoylation machinery. Non-enzyme dependent palmitoylation has also been found. Palmitoylated proteins such as Src family kinases, eNOS and Gα subunits have also been observed to be enriched in caveolae (Shenoyscaria *et al.,* 1994; Robbins, Quintrell & Bishop, 1995). It was recently shown that blocking the protein palmitoylation in a murine model of vascular injury results in inhibition of platelet aggregation and decreased incorporation into thrombi (Sim, Dilks & Flaumenhaft, 2007). Following this study, 215 proteins were identified in another study as candidate palmitoylated proteins in platelets and 51 of them were previously known while 103 were new putative palmitoylated proteins (Dowal *et al.,* 2011). Realising the importance of this PTM, several studies on various cell and tissue types have been conducted and hundreds of palmitoylated proteins have been identified (Kang *et al.,* 2008; Yang *et al.,* 2010; Dowal *et al.,* 2011; Martin *et al.,* 2012). A picture which has emerged for roles of S-palmitoylation is that it can influence and regulate the localisation of modified proteins (membrane tethering or lipid raft targetting) and therefore affect their function spatially. Moreover palmitoylation can be reversible and dynamic (Kang *et al.,* 2008) therefore can affect protein's functions and temporal signal transmission as well.

It has previously been shown that plasma membrane association can target a highly oligomerized cytoplasmic protein to endosome-like domains and into the membrane vesicles

like exosomes and microvesicles (Fang *et al.,* 2007). This targeting is independent of class E vacuolar protein sorting (VPS) pathways of sorting cargo to multivesicular body (MVB). Building on this study it was recently found that plasma membrane anchors like, N-myristoylation and S-palmitoylation, can target oligomeric proteins to the site of vesicle budding and into the exosomes and microvesicles (Shen *et al.,* 2011). The biogenesis of MVB and the mechanism of protein sorting to the exosomes/microvesicles are incompletely understood. Therefore, we proposed to identify the S-palmitoylated proteins present in exosomes isolated from the urine of healthy individuals using traditional ultracentrifugation based method. This will help increase our understanding of the protein sorting into these exosomes and shed light on the acylome of exosome/microvesicles. For enrichment of S-palmitoylated proteins we have used the biotin-acyl exchange method which has proven very effective in the recent past for identifying palmitoylated proteins (Kang *et al.,* 2008). Following purification/enrichment we have identified all the proteins using MudPIT analysis.

6.2 Materials and methods

6.2.1 Exosome/microvesicle isolation from human urine

Exosomal fraction from urine was purified as described in methods section of Chapter 2. Crude P200,000g pellet was used here as starting material.

6.2.2 Biotin-acyl exchange method for enrichment of S-palmitoylated proteins

All materials, unless specified otherwise were purchased from Sigma Chemical Company, St. Louis, MO, USA. For acyl-biotinyl exchange a previously published protocol was followed (Roth *et al.,* 2006) with some modifications. Briefly, P200,000g (2mg) extract (prepared by incubating with 1% beta-octylglucoside overnight at +4°C) was adjusted to buffer A (Tris 50mM pH 7.4, 5mM EDTA 4% SDS) and 2mM N-ethyl maleimide (NEM) was added followed by the incubation at +4°C with rotation. The following day, sample was dialysed (MWCO 3500 Da) to remove the NEM. Then 1M hydroxylamine (HA) was added with 1mM N-[6- (Biotinamido)hexyl]-3'- (2'-pyridyldithio)propionamide biotin (HPDP-biotin) and incubated for 1 hour at +25°C (HA+ fraction). For the minus hydroxylamine fraction (HA- fraction) same volume of Tris 50mM pH 7.4 was added instead of HA. HA was removed from HA+ fraction by dialysis and samples were adjusted to buffer A with 1mM HPDP-biotin and incubated at +25°C for 1 hour. HPDP-biotin was removed by dialysis (MWCO 3500 Da). Samples were adjusted to tris pH7.4 and incubated with 100ul streptavididn agarose at +25°C for 90 Minutes. Non-bound proteins were removed by washing with tris buffer (50mM, pH7.4) 4 times and finally bound proteins were eluted by incubating the resin with 1% (v/v) beta-mercaptoethanol at 37C for 20 minutes. Eluate was dialysed (MWCO 3500 Da) to remove beta-mercaptoethanol and converted to a dry pellet using speed-vacuum centrifugation.

6.2.3 SDS-PAGE

SDS-PAGE was carried out as described in methods section of Chapter 2. The gels were stained with silver staining for protein detection, as previously described (Shevchenko *et al.*, 1996).

6.2.4 Bioinformatic analysis and Gene ontology

The HA+ protein list was analysed with IPA software (Ingenuity systems, Redwood city, CA.) and DAVID Bioinformatics resources 6.7 (Huang, Sherman & Lempicki, 2009a; Huang, Sherman & Lempicki, 2009b) (NIH, USA). The graph for biological processes was generated with Blast2Go software (Conesa *et al.*, 2005). The CSS PALM 3.0 server (http://csspalm.biocuckoo.org/online.php) was used for high stringent prediction of palmitoylation of given proteins (Ren *et al.*, 2008).

6.2.5 LC-MS/MS analysis

It was carried out as described in the methods section of Chapter 5.

6.3 Results

6.3.1 Purification of S-palmitoylated proteins

The membrane vesicle pellet (P200,000g) was obtained as described in Chapter 2. This pellet, which is enriched in exosomal markers and contains mainly 40-100 nm vesicles (Chapter 2 Figure 2.2), was used for this study. The acyl-biotin exchange method was used for enriching the palmitoylated proteins from membrane vesicle fractions. This method is schematically presented in Figure 6.1.

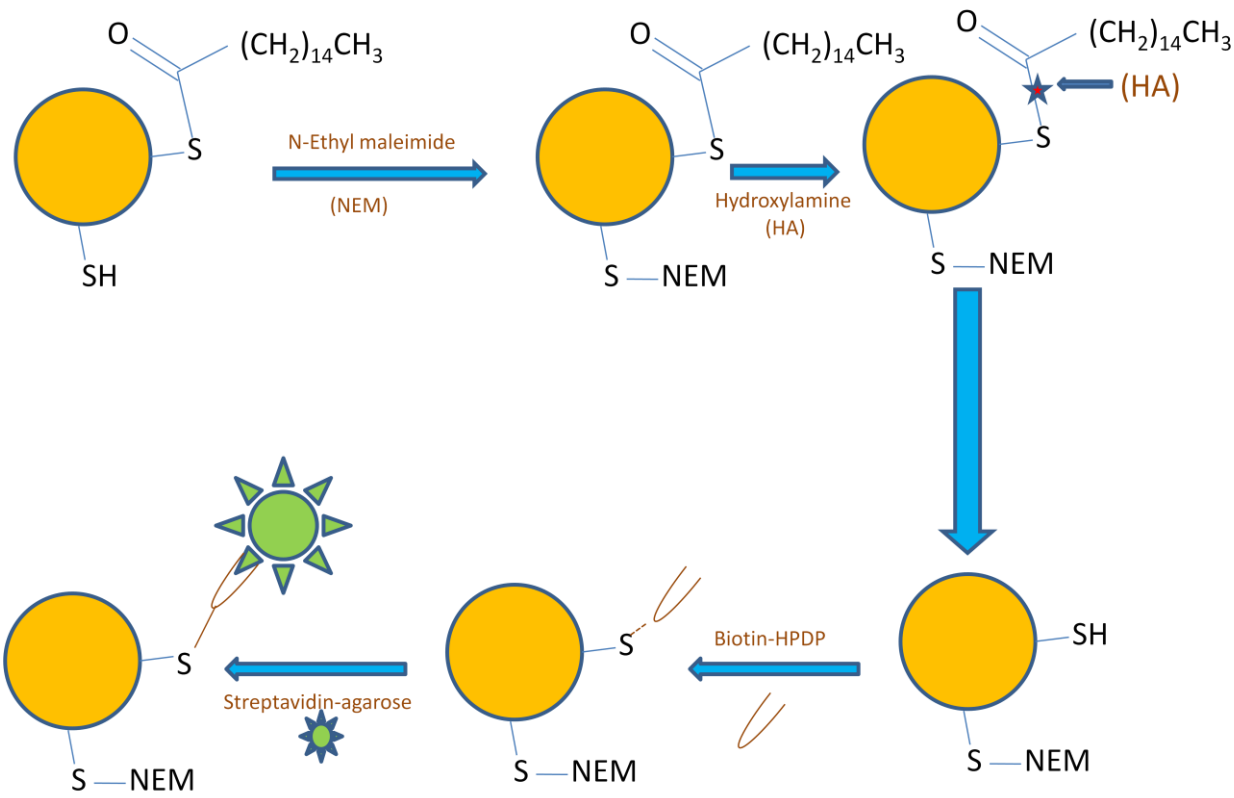

Figure 6.1: Schematic representation for purification of palmitoylated proteins using acyl-biotin exchange chemistry.

After blocking the free thiol groups of cysteine which may be present in the sample, HA was used to cleave the thioester bond of Cys-palmitate. This newly generated free thiol was then

biotinylated with HPDP-biotin. Two mg of total protein was divided into two parts and processed in parallell. One was treated with HA and the other was just incubated in buffer without HA. These two samples HA+ and HA- were then enriched with streptavidin-agarose following biotinylation. HPDP-biotin can be cleaved with reducing agents. Therefore, β-mercaptoethanol was used for eluting the proteins bound to streptavidin-agarose. This method yields a complex mixture of proteins in both the samples as shown in Figure 6.2.

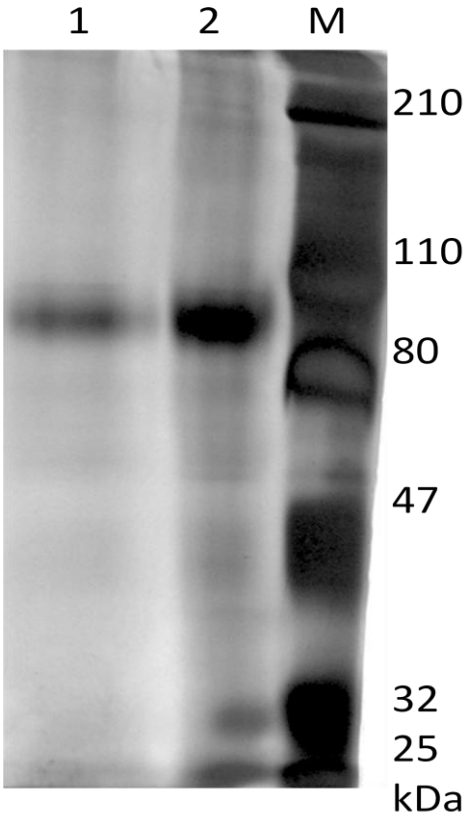

Figure 6.2: SDS-PAGE (T, 10%) of HA+ and HA- fractions enriched by acyl-biotin exchange method. 1: HA- 2. HA+ M. Molecular weight marker (in kDa).

However, the pattern looks similar but there is a visible enrichment of multiple proteins in HA+ fraction. Proteins in both the samples were identified by in solution LC-MS/MS.

6.3.2 Protein identification by LC-MS/MS

Excluding keratins, hundred and seventy two (172) proteins were identified in HA+ fraction (Supplementary Table 6.1) while 57 proteins were identified in HA- fraction (Supplementary Table 6.2). Out of this, 128 proteins were unique to HA+ fraction and 13 were unique to HA- fraction while 44 proteins were common (Figure 6.3).

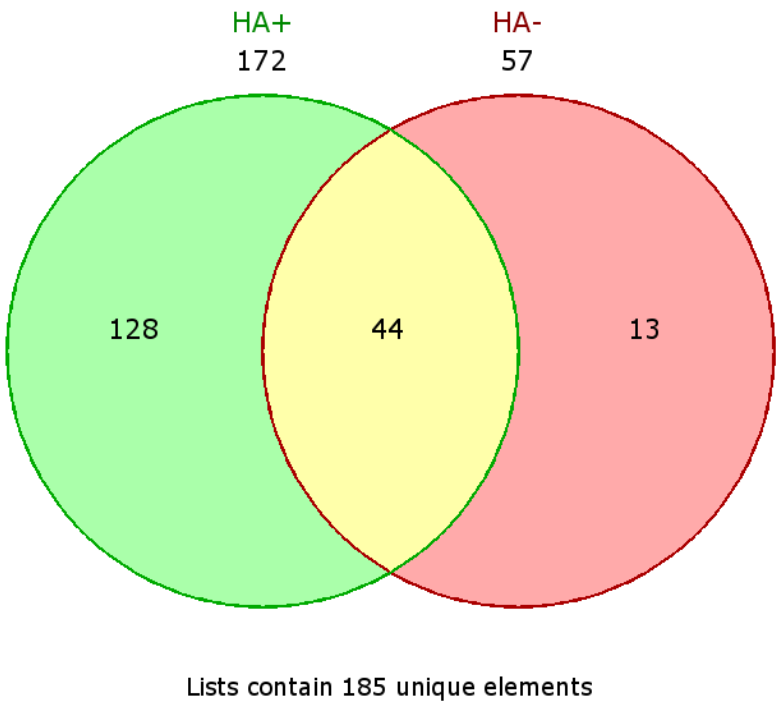

Figure 6.3: Comparison of proteins identified in the fraction HA+ vs HA-.

The HA- fraction is mainly composed of contaminant proteins getting enriched non-specifically and/or endogenously biotinylated proteins. These 128 unique proteins in HA+ fraction were taken as candidate palmitoylated proteins for further analysis. These 128 proteins will be referred to as HA+ proteins from here on. Sequences of these proteins were retrieved from Batch retrieval-Uniprot service (http://www.uniprot.org/) and submitted to CSS Palm 3.0 web server (Ren *et al.*, 2008) (http://csspalm.biocuckoo.org/online.php) which

predicts the potential palmitoylated proteins. The threshold was set to *high* which returned 78 potentially palmitoylated proteins. These proteins are listed in Supplementary Table 6.A.

6.3.3 Comparison with previous studies

There have been a number of studies identifying candidate palmitoylated proteins (High and medium confidence) on various tissue and cell-line types. Proteins from yeast, human platelets, rat neurons and DU145 cell lines have been identified (Roth *et al.,* 2006; Kang *et al.*, 2008; Yang *et al.*, 2010; Dowal *et al.*, 2011; Martin *et al.*, 2012). Table 6.1 presents 12 proteins from our dataset which have previously been identified as palmitoylated proteins. These 12 proteins were part of 78 proteins predicted by CSS Palm3.0 to be palmitoylated. These include a number of RAS family members, stomatin and tetraspanin CD9.

Table 6.1: The known palmitoylated proteins (platelet palm and S-acylation MCP) identified in our study. Sub-cellular location, function or type of molecule and biomarker applications (IPA, Ingenuity systems), if any, are indicated.

Serial No	Uniprot Accession	Symbol	Entrez gene name	Location	Function/type	Biomarker application	Reference for Biormaker	Reference
1	P04746	AMY2A	Amylase, alpha 2A (Pancreatic)	Extra-cellular space	Enzyme	N/A	N/A	Human Platelet (Dowal *et al.*, 2011)
2	Q13510	ASAH1	N-acylsphingosine amidohydrolase (acid ceramidase) 1	Cytoplasm	Enzyme	Diagnosis, Cancer	http://www.iqac.csic.es/images/stories/IQAC-KT/leaflet%20im_002.pdf	Human Platelet (Dowal *et al.*, 2011), DU145 cells (Yang *et al.*, 2010)
3	P15291	B4GALT1	UDP-GAL:BetaGlcNAc beta 1,4 galactosyltransferase,	Cytoplasm	Enzyme	Diagnosis, Cancer	IPA, Ingenuity	DU145 cells, human (Yang *et al.*, 2010)

			polypeptide 1					
4	P21926	CD9	CD9 molecule	Plasma membrane	Other	Efficacy, GEFITINIB	IPA, Ingenuity	DU145 cells, human (Yang *et al.*, 2010)
5	P25325	MPST	Mercaptopyruvate sulfurtransferase	Cytoplasm	Enzyme	N/A	N/A	(Martin *et al.*, 2012)
6	P60900	PSMA6	Proteasome (prosome, macropain) subunit, alpha type, 6	Cytoplasm	Peptidase	N/A	N/A	DU145 cells, human (Yang *et al.*, 2010)
7	P61026	RAB10	RAB10, member RAS oncogene family	Cytoplasm	Enzyme	N/A	N/A	DU145 cells, human (Yang *et al.*, 2010)
8	Q9NRW1	RAB6B	RAB6B, member RAS oncogene family	Cytoplasm	Enzyme	N/A	N/A	DU145 cells, human (Yang *et al.*, 2010)
9	P62979	RPS27A	Ribosomal protein S27a	Cytoplasm	Enzyme	N/A	N/A	DU145 cells, human (Yang *et al.*, 2010)
10	P62070	RRAS2	Related RAS viral (r-ras) oncogene homolog 2	Plasma membrane	Enzyme	N/A	N/A	BW5147-derived mouse T-cell hybridoma (Martin *et al.*, 2012)
11	P27105	STOM	Stomatin	Plasma membrane	Other	N/A	N/A	Human Platelet (Dowal *et al.*, 2011), DU145 cells

								(Yang *et al.*, 2010)
12	P27348	YWHAQ	Tyrosine 3-monooxegenase/tryptophan 5-monooxegenase activation protein, theta	Cytoplasm	Other	N/A	N/A	Human Platelet (Dowal *et al.*, 2011), DU145 cells (Yang *et al.*, 2010)

Most of these proteins's subcellular location is cytoplasm while stomatin and RRAS2 localise to plasma membrane. It is to be noted that most of the known palmitoylated proteins fall into the enzyme category which is expected because membrane localization is essential for correct functions of many enzymes. ASAH1 activity and B4GALT1 expression has previously been used as a biomarker for diagnosis of various types of cancers. CD9 has been used as a biomarker for efficacy of Gefitinib which is a selective inhibitor of epidermal growth factor receptor's tyrosine kinase domain. This drug has been developed for non-small cell lung cancer. CD9, a tetraspanin, is an exosomal marker proteins and in a previous study it has been shown that palmitoylation of CD9 was not necessary for its sorting into exosomes *in vitro* (Abache *et al.,* 2007). It suggests another mechanism is at play for CD9 sorting into exosomes.

Other than these 12 known palmitoylated proteins, 66 proteins from our dataset were also predicted to be palmitoylated by CSS PALM 3.0. These proteins were classified by SOSUI web server (Hirokawa, Boon-Chieng & Mitaku, 1998) and 28 proteins (42%) were found to be membrane proteins. Some of these transmembrane proteins are presented in Table 6.2

Table 6.2: Transmembrane proteins found in candidate palmitoylated proteins from our dataset. These proteins were predicted to be palmitoylated by CSS PALM 3.0. Uniprot accession, gene name and number of transmembrane domains are indicated in the table.

Uniprot accession	Gene/Protein Name	Transmembrane domains
Q15375	Ephrin type-A receptor 7	2
Q9NZH0	G-protein coupled receptor family C group 5 member B	7
Q9NQ84	G-protein coupled receptor family C group 5 member C	8
Q969X1	Protein lifeguard 3	7
Q13621	Solute carrier family 12 member 1	11
P22732	Solute carrier family 2, facilitated glucose transporter member 5	6
Q92673	Sortilin-related receptor	2
O60635	Tetraspanin-1	5
O75264	Transmembrane protein C19orf77	2
O00322	Uroplakin-1a	4

Palmitoylation sites are frequently found in cytoplasmic regions flanking transmembrane domains or within these domains (Salaun, Greaves & Chamberlain, 2010). Among the transmembrane proteins present in our dataset there are two G-protein coupled receptors, GPRC5B and GPRC5C. These GPCRs have not been reported to be palmitoylated in the literature. However, palmitoylation of GPCR and their cognate G-proteins has emerged as a general feature of this family of proteins and multiple GPCRs are already known to be palmitoylated (Qanbar & Bouvier 2003; Zheng *et al.*, 2012) and the list is growing by the day. GPRC5B and GPRC5C identified in our dataset are potential candidates for palmitoylation among other candidate proteins.

HA+ proteins were also compared to Exocarta database (Mathivanan *et al.,* 2012) and previous published studies on urinary exosomes and exosome-like vesicles ((Pisitkun, Shen & Knepper, 2004; Gonzales *et al.,* 2009; Hogan *et al.,* 2009; Wang *et al.,* 2011) Knepper, Hogan, Wang). As biotin-acyl exchange is an enrichment method for palmitoylated proteins

it might lead to identification of some low-abundance proteins which are part of exosomes but not otherwise identified in other studies because of the abundant proteins. All protein identifiers were converted to Unigene identifiers to facilitate the comparison.

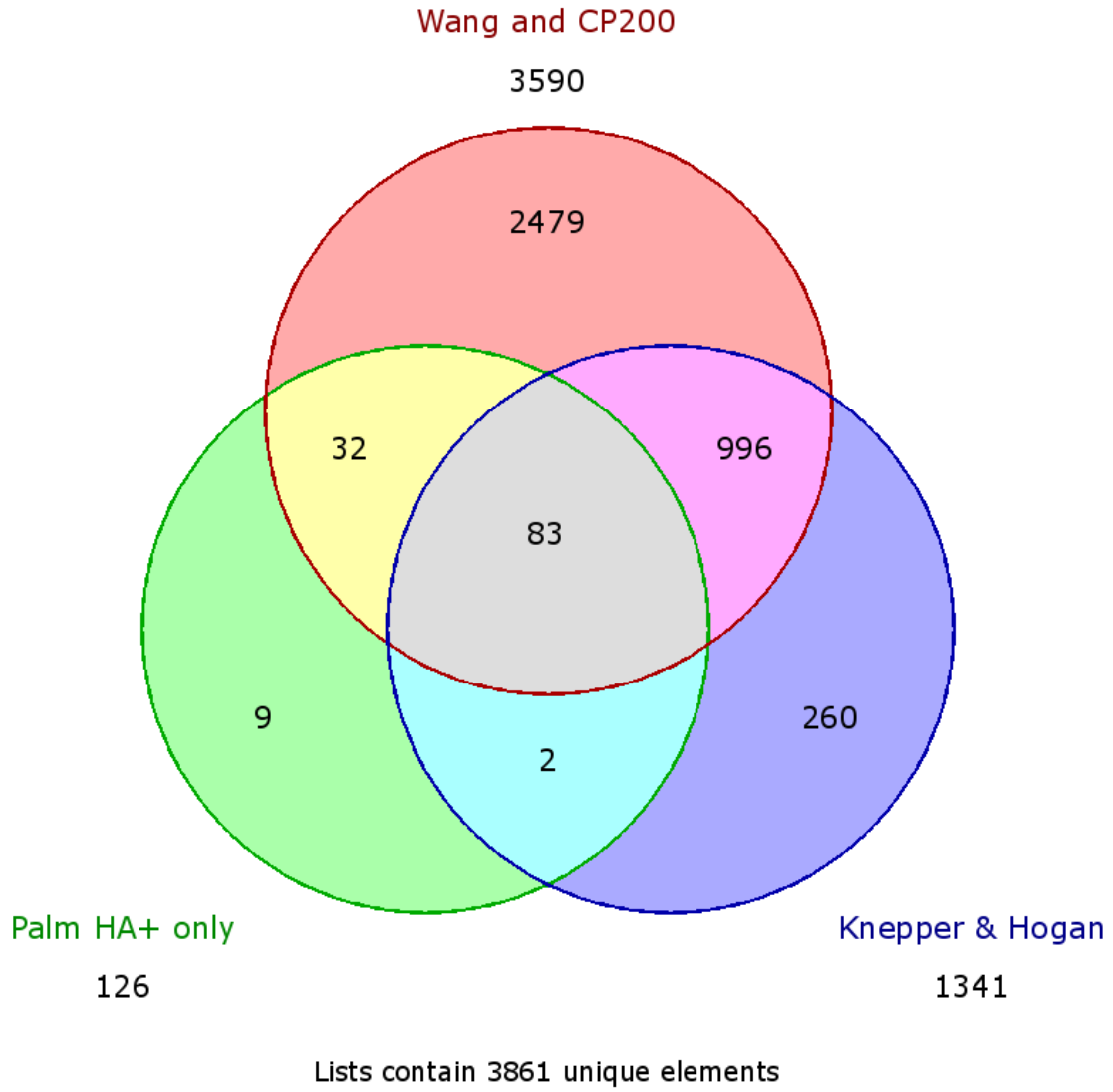

Figure 6.4: Our protein list from Chapter 2 (CP200) and other studies on membrane vesicles (Knepper, Hogan and Wang (Pisitkun, Shen & Knepper, 2004; Gonzales *et al.*, 2009; Hogan *et al.*, 2009; Wang *et al.*, 2011)) were compared to proteins Unique to HA+ fraction. Unigene identifiers were used for all of these studies.

Nine proteins were found to be unique to our dataset which were subsequently compared with Exocarta (Mathivanan *et al.,* 2012), the biggest database for membrane vesicle proteins and mRNA. Three proteins were returned which are unique to our analysis. These proteins include plasma membrane proteins like EPH receptor A7 and Copine-9 and are shown in Table 6.3. They could potentially be a part of exosomes and exosome-like vesicles.

Table 6.3: Some of the proteins identified for the first time as being part of exosomes. Gene name for the proteins, presence in HA+ or HA- fraction, sub-cellular location, Mascot score of their identification and number of protein matches are indicated in the table.

Uniprot accession	Gene Name	Found in fraction	Sub-cellular location	Mascot score	Protein Matches
Q8IYJ1	Copine-9	HA+ only	Plasma membrane	40	1
Q9UBR2	Cathepsin Z	HA+ only	Extracellular space	60	1
Q15375	EPH receptor A7	HA+ only	Plasma membrane	86	1

6.3.4 Bioinformatic analysis and gene ontology

IPA software from Ingenuity systems, DAVID bioinformatics resources (Huang, Sherman & Lempicki, 2009a; Huang, Sherman & Lempicki, 2009b) (NIH, USA) and Blast2Go software (Conesa *et al.,* 2005) were used for detailed bioinfomatic analysis of the HA+ proteins. An annotated list from IPA software produced 111mapped Ids, classified according to their sub-cellular location and function/type. From this data figures were generated using MS Excel for the percentage of proteins localised to various sub-cellular compartments or their functional category. Figure 6.5 presents proteins classified according to their sub-cellular localisation.

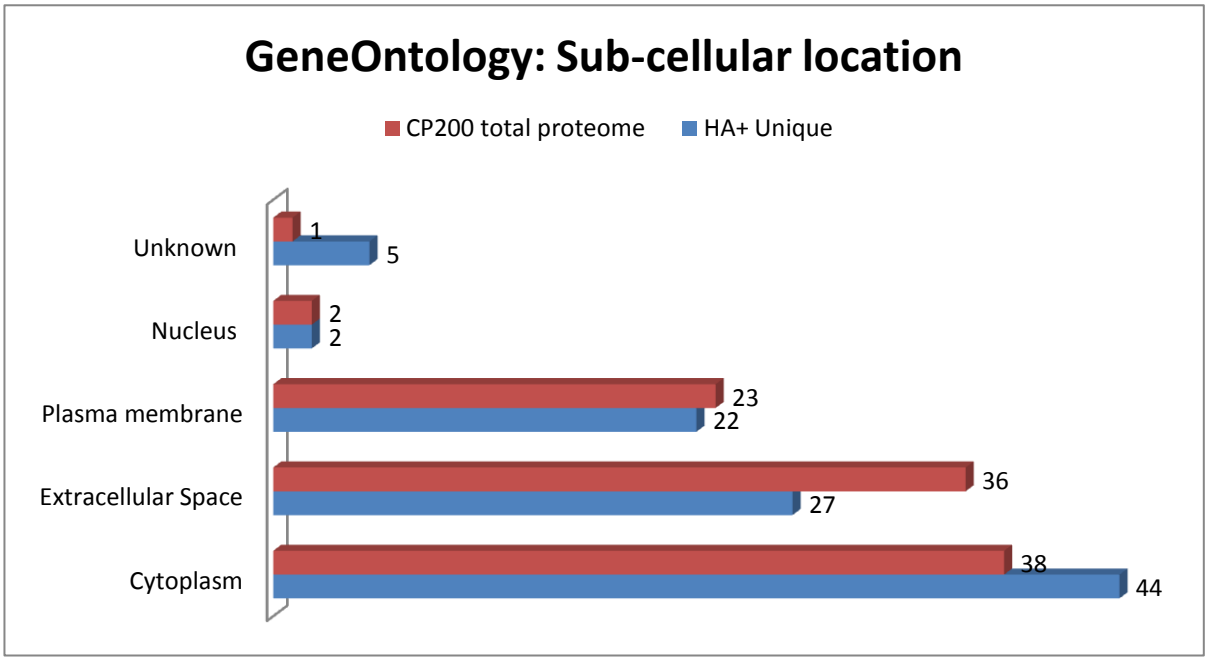

Figure 6.5: Sub-cellular localisation of proteins unique to HA+ fraction. The annotation was done with IPA software from Ingenuity systems, USA.

The biggest category in localisation is cytoplasmic proteins forming 44% of the total palmitoyled candidate proteins. This is expected as many cytoplasmic proteins are palmitoyled sub-cellularly which changes their location to the membrane where they exert their functions (Shenoyscaria *et al.,* 1994). The cytoplasmic proteins are over-represented in HA+ fraction compared to total proteome. The next biggest category is extracellular space followed by plasma membrane proteins. Only two nuclear proteins, cofillin-2 and nuclear transport factor 2 were identified and 6 proteins were classified as of unknown location. Four of these proteins of unknown location are annotated by Uniprot as membrane proteins or involved in membrane trafficking (Copine family member IX, Matrix-remodeling-associated protein 8, Chromosome 19 open reading frame 77 and Tetraspanin 1) while one is a cytoplasmic protein (5'-nucleotidase, cytosolic III-like) and another being signalling protein of unknown location (POTE ankyrin domain family, member I).

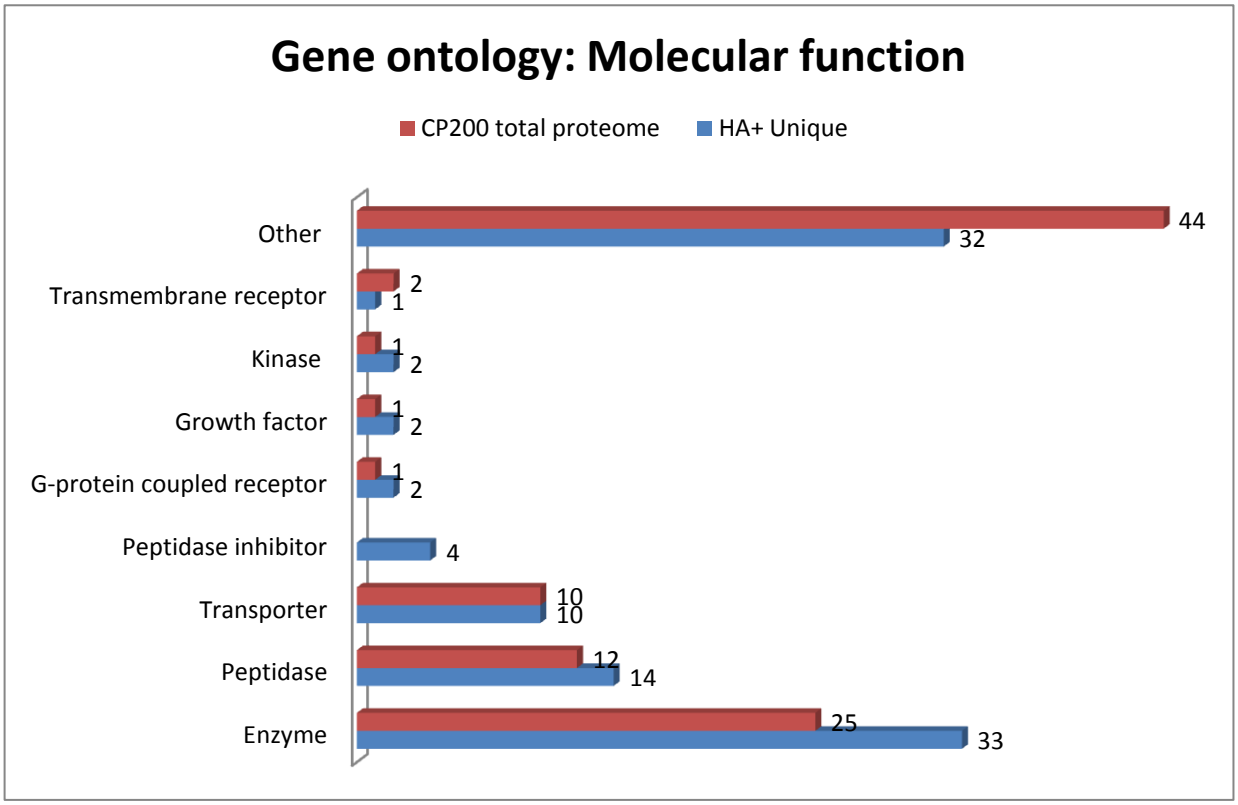

Figure 6.6: Protiens unique to HA+ fraction classified according to molecular function. The annotation was done with IPA software from Ingenuity systems, USA.

On classifying the HA+ proteins by molecular function/type (Figure 6.6) the biggest category is enzymes (33%). This trend of discovering enzymes as being palmitoylated has been seen in previous studies as well. The trafficking of enzyme to membrane domains where their activities are needed for normal cell signalling is important and palmitoylation is one of the signals for membrane localisation (Shenoyscaria *et al.*, 1994). When compared to total proteins, the enzyme category is over-represented in HA+ fraction. Proteases and protease inhibitors form another 18% of the list while G-protein coupled receptors (GPRC5B and GPRC5C) and kinases, including pyruvate kinase, are also present.

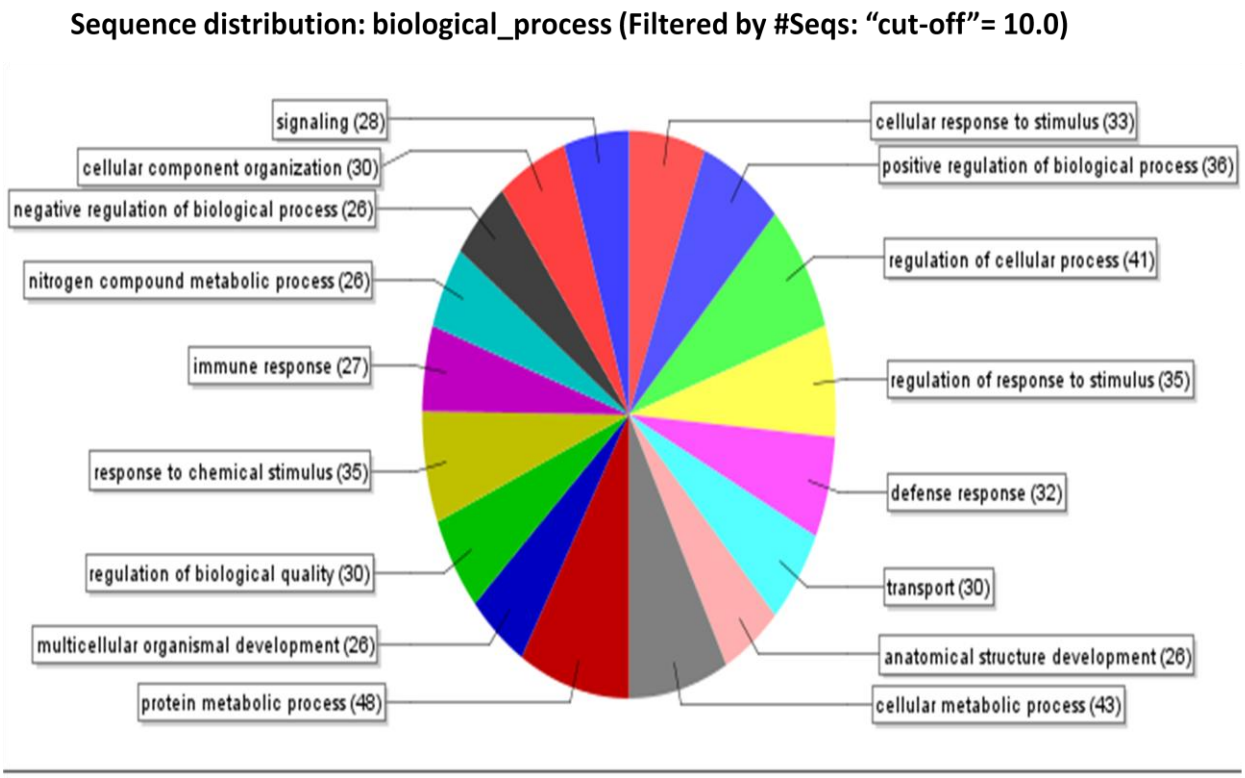

Figure 6.7: Blast2Go software was used to annotate proteins according to the biological processes they are involved in. The 'cut-off' was kept at 10 sequences and only those categories having at least 10 proteins in it were considered.

Blast2Go software was used for classifying proteins according to the biological processes they are involved in (Figure 6.7). The biggest categories involving multiple proteins were signalling, cellular response to stimulus and its regulation and metabolic processes. Palmitoylation seems to selectively target proteins involved in cell signalling and regulation. DAVID bioniformatics resources found some proteins in the HA+ fraction which are transporters of proteins and small molecules. These proteins are listed in Table 6.4.

Table 6.4: Proteins annotated by DAVID as being involved in protein and small molecules transport.

ID	Gene Name
P54920	N-ethylmaleimide-sensitive factor attachment protein, alpha
P61026	RAB10, member RAS oncogene family
P59190	RAB15, member RAS onocogene family
Q9H082	RAB33B, member RAS oncogene family
Q9NRW1	RAB6B, member RAS oncogene family
P00450	ceruloplasmin (ferroxidase)
P68871	hemoglobin, beta
P02790	Hemopexin
P02649	hypothetical LOC100129500; apolipoprotein E
P61970	nuclear transport factor 2
Q8WUM4	programmed cell death 6 interacting protein
P41222	prostaglandin D2 synthase, hematopoietic; prostaglandin D2 synthase 21kDa (brain)
Q13621	solute carrier family 12 (sodium/potassium/chloride transporters), member 1
P22732	solute carrier family 2 (facilitated glucose/fructose transporter), member 5
Q92673	sortilin-related receptor, L (DLR class) A repeats-containing
Q9UK41	vacuolar protein sorting 28 homolog (S. cerevisiae)
Q9UN37	vacuolar protein sorting 4 homolog A (S. cerevisiae)

Among these transport proteins RAB10 and RAB6B are already known to be palimotylated (Yang *et al.,* 2010) while RAB15 and RAB33B are very promising candidates. This class also includes programmed cell death 6 interacting proteins (Alix) and VPS family proteins which are members of endosomal sorting complex required for transport (ESCRT) that is involved in sorting proteins into exosomes (Babst *et al.,* 2002a; Babst *et al.,* 2002b; Katzmann, Babst & Emr, 2001).

For further analysis of HA+ proteins, IPA was used in biomarker analysis and 29 proteins were returned among the list which has been used as biomarkers of various diseases (Table 6.5). Different types of cancers and cardiovascular diseases were among the top diseases in which these biomarkers are used.

Table 6.5: IPA software was used for biomarker analysis of the HA+ proteins. These biomarkers are listed according to their applications.

Symbol	Entrez Gene Name	Drug (s)	GenPept /UniProt /Swiss-Prot Accession	Plasma/Serum	Urine	Biomarker Application (s)
ACE	angiotensin I converting enzyme (peptidyl-dipeptidase A) 1	pentopril, perindoprilat, amlodipine/benazepril, lisinopril/hydrochlorothiazide, benazepril, enalapril, perindopril, captopril, enalapril/felodipine, hydrochlorothiazide/moexipril, benazepril/hydrochlorothiazide, hydrochlorothiazide/quinapril, fosinopril/hydrochlorothiazide, captopril/hydrochlorothiazide, enalapril/hydrochlorothiazide, ramipril, moexipril, quinapril, lisinopril, enalaprilat, trandolapril, trandolapril/verapamil, diltiazem/enalapril, fosinopril	P12821	X	x	Diagnosis, Efficacy,Prognosis,Safety
AGT	angiotensinogen (serpin peptidase inhibitor, clade A, member 8)		P01019	X	X	Efficacy
ANXA1	annexin A1		P04083	X	X	Diagnosis, Prognosis, Unspecified Application
ANXA5	annexin A5		P08758	X	X	Diagnosis
APOE	apolipoprotein E		P02649	X	X	Diagnosis, Efficacy,Prognosis,Unspecified Application
B4GALT1	UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase,		P15291		X	Diagnosis

	polypeptide 1					
CD9	CD9 molecule		P21926		X	Efficacy
CLU	Clusterin		P10909	X	x	Diagnosis, Efficacy,Unspecified Application
CP	ceruloplasmin (ferroxidase)		P00450	X	X	Disease Progression ,Efficacy
CTSD	cathepsin D		P07339	X	X	Diagnosis, Unspecified Application
DPP4	dipeptidyl-peptidase 4	saxagliptin, talabostat, SYR-322, sitagliptin, linagliptin, metformin/saxagliptin, simvastatin/sitagliptin, metformin/sitagliptin, linagliptin/metformin	P27487	X	X	Prognosis
EZR	Ezrin		P15311	X	X	Prognosis
FGA	fibrinogen alpha chain	F2	P02671	X	X	Diagnosis
FUCA1	fucosidase, alpha-L- 1, tissue		P04066		X	Diagnosis
GSN	Gelsolin		P06396	X	X	Disease Progression ,Efficacy
GSTM3	glutathione S-transferase mu 3 (brain)		P21266	X	X	Diagnosis, Disease Progression ,Prognosis
GSTP1	glutathione S-transferase pi 1		P09211	X	X	Diagnosis, Efficacy,Prognosis,Response to Therapy,Safety,Unspecified Application
HSPB1	heat shock 27kDa protein 1		P04792		X	Diagnosis
KLK3	kallikrein-related peptidase 3		P07288	X	X	Diagnosis, Disease Progression ,Efficacy,Safety,Unspecified Application
LGALS 3BP	lectin, galactoside-binding, soluble, 3 binding protein		Q08380	X	x	Prognosis
MME	membrane metallo-endopeptidase		P08473		X	Diagnosis, Efficacy,U

						nspecified Application
PKM2	pyruvate kinase, muscle		P14618	X	X	Diagnosis, Unspecified Application
PLAU	plasminogen activator, urokinase		P00749	X	X	Disease Progression,Efficacy,Prognosis,Response to Therapy
PRDX1	peroxiredoxin 1		Q06830	X	X	Diagnosis
PTGDS	prostaglandin D2 synthase 21kDa (brain)		P41222	X	X	Efficacy
SERPINA1	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1		P01009	X	X	Diagnosis, Disease Progression,Unspecified Application
THBS1	thrombospondin 1		P07996	X	X	Diagnosis, Efficacy,Prognosis
TPI1	triosephosphate isomerase 1		P60174	X	X	Diagnosis, Unspecified Application
YWHAZ	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide		P63104		X	Diagnosis

Renal toxicity analysis was also carried out and multiple proteins involved in renal failure and renal injury etc. were found. These proteins and the diseases they are involved in are listed in Table 6.6.

Table 6.6: Toxicity function analysis of HA+ proteins by IPA software. Various diseases and the proteins involved in these diseases are listed along with p-value for their enrichment.

Category	Functions Annotation	p-Value	Molecules
Glomerular Injury	Glomerulosclerosis	1.54E-03	AGT,APOE,CLU,GAS6
Glomerular Injury	lipoprotein glomerulopathy	6.71E-03	APOE
Glomerular Injury	renal fibrosis	8.39E-02	AGT
Glomerular Injury	focal segmental glomerulosclerosis	1.02E-01	CLU
Increased Levels of Albumin	increases flux of albumin	1.34E-02	THBS1
Increased Levels of Albumin	increases excretion of albumin	3.31E-02	AGT
Increased Levels of Creatinine	increases clearance of creatinine	9.00E-02	AGT
Increased Levels of Creatinine	increases quantity of creatinine	1.66E-01	AGT
Increased Levels of Potassium	increases quantity of K+	6.51E-02	SLC12A1
Increased Levels of Potassium	increases excretion of K+	9.61E-02	AGT
Kidney Failure	renal failure	6.98E-05	ACE,AGT,AMY2A,GSTP1,PTGDS,SLC12A1,THBS1
Kidney Failure	acute renal failure	2.25E-03	ACE,GSTP1,SLC12A1
Kidney Failure	failure of kidney	3.23E-02	AGT,SLC12A1
Kidney Failure	end stage renal disease	9.86E-02	ACE,AGT
Kidney Failure	tubulo-interstitial fibrosis	2.36E-01	AGT
Nephrosis	minimal change nephrotic syndrome	5.88E-02	CLU
Renal Atrophy	atrophy of kidney	3.58E-04	ACE,AGT,EFEMP1
Renal Damage	damage of tubulointerstitium	3.32E-03	ACE,APOE
Renal Damage	damage of kidney	3.40E-03	ACE,AGT,HPX,THBS1
Renal Damage	injury of kidney	6.88E-02	AGT,THBS1
Renal Damage	injury of renal glomerulus	9.00E-02	AGT
Renal Damage	reperfusion injury of kidney	1.02E-01	THBS1
Renal Damage	damage of renal tubule	1.44E-01	THBS1
Renal Degeneration	degeneration of kidney	7.14E-02	CLU
Renal Dilation	vasodilation of kidney	1.34E-02	AGT
Renal Dysfunction	dysfunction of kidney	1.08E-01	AGT
Renal Necrosis/Cell Death	cell death of kidney cells	1.70E-03	AGT,APOE,CLU,GSTP1,HSPB1,PTGDS,VTN,YWHAQ
Renal Necrosis/Cell Death	cell death of kidney cell lines	1.18E-02	AGT,GSTP1,HSPB1,PTGDS,VTN,YWHAQ
Renal Necrosis/Cell Death	necrosis of renal glomerulus	2.00E-02	PLAU

Renal Nephritis	membranous glomerulonephritis	9.00E-02	CLU
Renal Nephritis	focal glomerulonephritis	9.61E-02	ACE
Renal Nephritis	Nephritis	1.04E-01	ACE,AGT,APCS,CLU
Renal Nephritis	Glomerulonephritis	1.64E-01	ACE,APCS,CLU
Renal Tubule Injury	proximal tubular toxicity	3.75E-05	CP,FGA,FGB (includes EG:110135),GSTP1,HPX,HSPB 1
Renal Tubule Injury	damage of tubulointerstitium	3.32E-03	ACE,APOE
Renal Tubule Injury	damage of renal tubule	1.44E-01	THBS1

These proteins listed in Table 6.6 are involved in signalling processes in various kidney pathologies. These would make good drug targets and biomarker candidates for these diseases. Another interesting analysis carried out by IPA was finding the upstream regulators of HA+ proteins. This will likely shed light on what pathways these proteins are involved in. These upstream regulators are GPCRs, growth factors and cytokines among others. Selected parts of this analysis are presented in table 6.7.

Table 6.7: Upstream regulators of the HA+ proteins. Multiple proteins are regulated by these regulators listed in column 1.

Upstream Regulator	Molecule Type	p-value of overlap	Target molecules in dataset
P2RY2	G-protein coupled receptor	1.17E-03	SLC12A1,VTN
CXCR4	G-protein coupled receptor	1.57E-02	CD9,DPP4
AGTR1	G-protein coupled receptor	1.66E-02	AGT,CP
CCR5	G-protein coupled receptor	1.75E-02	CD9,SERPINA1
GPR39	G-protein coupled receptor	2.30E-02	CLU
ADORA2A	G-protein coupled	2.68E-02	COL6A1,NAPA (includes EG:108124),VPS4A

	receptor		
NFkB (complex)	Complex	8.54E-06	AGT,APOE,B4GALT1,CLU,FTH1 (includes EG:14319),FUCA1,GAS6,HSPB1,IGHG3,IGKC
IL6	cytokine	3.63E-06	AGT,ANPEP,ANXA1,APCS,APOE,CLU,CP,FGA,FGB (includes EG:110135),KLK3
IFNG (includes EG:15978)	Cytokine	3.81E-06	AGRN,AGT,AZGP1,CP,CTSC,CTSD,CTSZ,DPP4,FTH1 (includes EG:14319),GAS6
IL1B	Cytokine	1.24E-05	ANXA1,APCS,APOE,CP,CTSZ,DPP4,FGB (includes EG:110135),GAS6,GSTA1,GUSB
TNF	Cytokine	3.12E-05	ACE,AGT,ANPEP,APCS,APOE,B4GALT1,CLU,CP,CTSC,CTSZ
GH1	Cytokine	4.85E-05	AGT,APOE,CLU,FTH1 (includes EG:14319),GSTP1,HBB,PRDX1,YWHAZ
OSM	Cytokine	6.26E-04	ACE,ANXA1,ASAH1,COL6A1,FGA,FGB (includes EG:110135),PLAU,QSOX1,SERPINA1,UPK1A
IL13	Cytokine	6.77E-04	CTSC,EZR,GAS6,GSN,PLAU,QSOX1,SERPINA1,THBS1
EDN1	Cytokine	2.34E-03	ANXA1,ANXA5,EZR,PLAU,THBS1
CSF2	Cytokine	2.94E-03	ABP1 (includes EG:26),ACE,CTSC,HBB,MME,QSOX1
IL5	Cytokine	3.51E-03	ABP1 (includes EG:26),CTSC,PKM2,QSOX1,RRAS2,TPI1
PRL	Cytokine	5.12E-03	ANXA5,CLU,CTSA,CTSD,MME
IL1A	Cytokine	2.39E-02	DPP4,FTH1 (includes EG:14319),PLAU,SERPINA1
IL4 (includes EG:16189)	Cytokine	3.51E-02	APOE,CTSC,DPP4,IGHG3,KLK3,LGALS3BP,PLAU,PSMA6
IL8	Cytokine	3.53E-02	ANXA1,KLK3
EGF (includes EG:13645)	growth factor	3.04E-06	ANPEP,B4GALT1,CLU,CTSD,DPP4,EZR,GSTP1,IDH1,KLK3,PLAU
FGF19	growth factor	5.20E-06	AMY2A,APOE,MME,PTGDS,SERPINA1,SLC2A5
AGT	growth factor	5.06E-04	ACE,AGT,COL6A1,CP,EFEMP1,HSPB1,IDH1,PLAU,SOD3
TGFB1 (includes EG:21803)	growth factor	5.50E-03	ACE,ANPEP,APOE,CLU,COL6A1,CTSC,CTSD,FTH1 (includes EG:14319),GSN,GUSB
FGF2	growth factor	6.78E-03	ACE,ANPEP,FTH1 (includes EG:14319),PLAU,THBS1,YWHAZ
VEGFA	growth factor	7.01E-03	ACE,ANPEP,CSTB,PLAU,THBS1
KITLG	growth factor	9.18E-03	FTH1 (includes EG:14319),GAS6,HBB,MME,PKM2
IGF1	growth factor	9.84E-03	ANPEP,CLU,CTSD,HBB,PLAU,THBS1

6.4 Discussion

Palmitoylation is a PTM which is reversible and dynamically regulated (Kang *et al.,* 2008; Martin *et al.,* 2012) although mechanisms involved in the turnover of palmitoylated proteins are not clearly understood. While protein acyltransferases and protein acylthioesterases are known to regulated enzymatic recycling of palmitoylation in many proteins (Resh, 2006; Smotrys & Linder, 2004) controlling membrane association in a reversible manner, many palmitoylated proteins are not regulated by them (Resh, 2006; Linder & Deschenes, 2007). Palmitoylation of proteins was not understood better for many years mainly because of the lack of method available to detect the palmitoylation and identification of modified proteins. The traditional method for detecting palmitoylation was ^{3}H palmitate labelling of cells (Drisdel *et al.,* 2006). Following the labelling the proteins can be purified and analysed by SDS-PAGE. However other than the disadvantage of radioactive material handling, it has two distinct disadvantages. First, the method usually requires very long labelling and exposure times and second, it is limited to the analysis of live cells. Another method is the metabolic labelling of cells with 17-Octadecayonic acid (17-ODYA) (Martin *et al.,* 2012) but this too is limited for analysing living cells.

Recently a new method has been described which exchanges the acyl group with biotin in three simple steps (Roth *et al.,* 2006). This method is called acyl-biotin exchange (ABE). Biotinylated proteins can then be purified using streptavidin-agarose and identified by LC-MS/MS. This is the only method which can be applied to any type of samples to detect and identify palmitoylated proteins. We have used the ABE method here and purified candidate palmitoylated proteins from urinary exosomes and identified them by LC-MS/MS. The method is described in Figure 6.1. The two fractions HA+ and HA- allowed identification of 172 and 57 proteins, respectively. Multiple proteins are shared among the two fractions. Theoretically there should be no overlap between the two fractions but endogenously

biotinylated proteins, incomplete blocking of free thiols by NEM, non-specific biotinylation and non-specific binding of proteins to streptavidin-agarose matrix lead to a significant number of overlaps between the two groups. For example the maleimide group, although specific to thiols at low concentration is known to react with non-thiol goups at high concentrations (Tyagarajan, Pretzer & Wiktorowicz, 2003).

Therefore we have only considered proteins which are unique to HA+ fraction as candidate palmitoylated proteins and proteins found in HA- fractions are not considered. Of the 128 proteins unique to HA+ fraction we have identified 12 proteins which were previously identified as being biotinylated in other studies (Dowal *et al.,* 2011; Yang *et al.,* 2010; Martin *et al.,* 2012). These proteins include RAS family members RAB10, RAB6B and RRAS2, tetraspanin CD9 and proteasome subunit PSMA6. RAS family members RABs are involved in membrane trafficking and proteins containing ciliary targeting signals are transported to preicentriolar recycling endosomes with the aid of RAB10 (Nachury, Seeley & Jin, 2010). The endosomal system trafficking including exosome release is controlled by RAB proteins (Hendrix & Hume, 2011). Palmitoylation of RAB proteins therefore seems logical and it would be expected to control their membrane association in a reversible manner. R-RAS2, on the other hand can transform cells if its activity is aletred (Graham *et al.,* 1994) and its overexpression may contribute to development of human breast cancers (Clark *et al.,* 1996). This protein may contribute to the effect of exosomes uptake by recipient cells. Tetraspanin CD9 is a part of tetraspanin web domains on plasma membrane and takes part in cell adhesion, cell motility and tumor metastasis (Ikeyama *et al.,* 1993; Masellissmith & Shaw, 1994). Acylation of this protein will help in its association with specialised plasma membrane domains.

Another protein identified in the current study and known to be palmitoylated is Stomatin (Snyers, Umlauf & Prohaska, 1999) which is known to be enriched in lipid-rafts (Mairhofer

et al., 2002). Raft-like domains are known to be enriched in exosomes and palmitoylation of this protein will help its association with rafts and subsequent incorporation into exosomes. Proteasome subunit alpha type-6 was also detected in our analysis and it is previously known to be a candidate palmitoylated protein (Yang *et al.,* 2010). All 7 alpha and 7 beta chains of 20S proteasomes have been previously detected in mesenchymal exosomes and it was also shown that a functional 20S proteasome co-purifies with exosomes in blood (Lai *et al.,* 2012). This has implications for recipient cells of these exosomes as proteasome acquisition will help degrade misfolded and oxidised proteins. It is to be noted that the ABE method will also result in enrichment of proteins which uses thioester linkages for attachment of other groups like lipoic acid. Currently no method exists to differentiate between the two but not many lipoic acid-modified proteins are known. Moreover, many of the known palmitoylated proteins detected by ABE have been validated in many studies by metabolic labelling which is specific for palmitoylation (Martin *et al.,* 2012) confirming that ABE mostly detects palmitoylated proteins.

ABE enrichment also helped identify some of the low-abundance proteins which were previously not known to be part of exosomes. These 18 proteins were not found in Exocarta which is the biggest database of exosomal proteins and RNA species. But when compared to our list from chapter 2 and other membrane vesicles studies, three proteins were found to be unique to our dataset (Table 6.3). One of these proteins is Ephrin type-A receptor 7 which is a receptor tyrosine kinase which binds GPI-anchored ephrin-A family ligands and modulates cell-cell adhesion and repulsion. It is typically a brain protein but it is also expressed in kidney cells (Genecards.org). These are new candidate exosomal proteins and expand the coverage of exosomal proteins.

Annotation of HA+ proteins by DAVID returned a category of proteins which are involved in transport of proteins and small molecules (Table 6.3). It is previously known that palmitoyl-

CoA which is a lipid palmitate donor supports budding and fusion of vesicles thereby inducing golgi transport assay. This effect is thought to be due to resulting palmitoylation of a protein upon supply of palmitoyl-CoA (Glick & Rothman, 1987; Pfanner *et al.,* 1990). Therefore these transporters are good candidate for palmitoylation with many of them being intracellular transporters.

IPA revealed 29 biomarkers present in HA+ fractions as well as multiple proteins involved in pathogenesis of various kidney pathologies (Table 6.4 & 6.5). These findings demonstrate the power of urinary exosome analysis for understanding pathologies and finding sensitive biomarkers of various diseases. Another analysis performed by IPA was the finding of upstream regulators for multiple proteins form the HA+ fractions (Table 6.6). Two of these regulators are interleukin-6 (IL-6) and tumor necrosis factor-α (TNF- α) which regulate multiple proteins found in our analysis in the HA+ fractions. It is previously known that IL-6 and TNF- α may contribute to Th cell imbalance, cardiovascular disease and wasting in patients with end stage kidney disease (ESRD) in uremic millieu (Stenvinkel *et al.,* 2005). Il-6 is elevated in most of the ESRD patients and reduced renal function affects TNF- α clearance in rats (Bemelmans, Gouma & Buurman, 1993). Anti-cytokine therapies have been proposed for ESRD patients (Ridker *et al.,* 1999; Huang *et al.,* 2003; Zhao & Zhang, 2003; Stenvinkel *et al.,* 2005) and the downstream targets of these cytokines would be excellent candidates for monitoring biomarkers of such therapies.

In conclusion, we have identified 128 proteins which are unique to HA+ fractions. Seventy nine proteins from these 128 were predicted as high-confidence palmitoylated proteins by CSS-PALM 3.0 online server. Twelve previously known palmitoylated proteins were detected which confirmed their PTM status in urinary exosomes. Sixty six candidate palmitoylated proteins were established by their presence in only HA+ fraction. This study serves as a platform for the future studies on human urinary exosome and exosome-like

vesicles to verify the presence of palmitoylated proteins and further validate these results. IPA analysis and DAVID annotation detected the presence of several biomarker candidates as well as molecular targets for diseases.

6.5 Reference

Abache, T., Le Naour, F., Planchon, S., Harper, F., Boucheix, C. & Rubinstein, E. (2007), "The transferrin receptor and the tetraspanin web molecules CD9, CD81, and CD9P-1 are differentially sorted into exosomes after TPA treatment of K562 cells", *JOURNAL OF CELLULAR BIOCHEMISTRY*, vol. 102, no. 3, pp. 650-664.

Babst, M., Katzmann, D., Estepa-Sabal, E., Meerloo, T. & Emr, S. (2002 a), "ESCRT-III: An endosome-associated heterooligomeric protein complex required for MVB sorting", *DEVELOPMENTAL CELL*, vol. 3, no. 2, pp. 271-282.

Babst, M., Katzmann, D., Snyder, W., Wendland, B. & Emr, S. (2002 b), "Endosome-associated complex, ESCRT-II, recruits transport machinery for protein sorting at the multivesicular body", *DEVELOPMENTAL CELL*, vol. 3, no. 2, pp. 283-289.

Bemelmans, M., Gouma, D. & Buurman, W. (1993), "Influence of nephrectomy on tumor-necrosis-factor clearance in a murine model", *JOURNAL OF IMMUNOLOGY*, vol. 150, no. 5, pp. 2007-2017.

Bijlmakers, M. & Marsh, M. (2003), "The on-off story of protein palmitoylation", *TRENDS IN CELL BIOLOGY*, vol. 13, no. 1, pp. 32-42.

Clark, G., Kinch, M., Gilmer, T., Burridge, K. & Der, C. (1996), "Overexpression of the Ras-related TC21/R-Ras2 protein may contribute to the development of human breast cancers", *ONCOGENE*, vol. 12, no. 1, pp. 169-176.

Conesa, A., Gotz, S., Garcia-Gomez, J., Terol, J., Talon, M. & Robles, M. (2005), "Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research", *BIOINFORMATICS*, vol. 21, no. 18, pp. 3674-3676.

Dowal, L., Yang, W., Freeman, M.R., Steen, H. & Flaumenhaft, R. (2011), "Proteomic analysis of palmitoylated platelet proteins", *BLOOD*, vol. 118, no. 13, pp. E62-E73.

Drisdel, R.C., Alexander, J.K., Sayeed, A. & Green, W.N. (2006), "Assays of protein palmitoylation", *METHODS*, vol. 40, no. 2, pp. 127-134.

Fang, Y., Wu, N., Gan, X., Yan, W., Morrell, J.C. & Gould, S.J. (2007), "Higher-order oligomerization targets plasma membrane proteins and HIV gag to exosomes", *PLOS BIOLOGY*, vol. 5, no. 6, pp. 1267-1283.

GLICK, B. & ROTHMAN, J. (1987), "Possible role for fatty acyl-coenzyme-A in intracellular protein-transport", *NATURE*, vol. 326, no. 6110, pp. 309-312.

Gonzales, P.A., Pisitkun, T., Hoffert, J.D., Tchapyjnikov, D., Star, R.A., Kleta, R., Wang, N.S. & Knepper, M.A. (2009), "Large-Scale Proteomics and Phosphoproteomics of Urinary Exosomes", *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY*, vol. 20, no. 2, pp. 363-379.

Graham, S., Cox, A., Drivas, G., Rush, M., Deustachio, P. & Der, C. (1994), "Aberrant function of the RAS-related protein TC21/R-RAS2 triggers malignant transformation", *MOLECULAR AND CELLULAR BIOLOGY*, vol. 14, no. 6, pp. 4108-4115.

Hendrix, A. & Hume, A.N. (2011), "Exosome signaling in mammary gland development and cancer", *INTERNATIONAL JOURNAL OF DEVELOPMENTAL BIOLOGY*, vol. 55, no. 7-9, SI, pp. 879-887.

Hirokawa, T., Boon-Chieng, S. & Mitaku, S. (1998), "SOSUI: classification and secondary structure prediction system for membrane proteins", *BIOINFORMATICS*, vol. 14, no. 4, pp. 378-379.

Hogan, M.C., Manganelli, L., Woollard, J.R., Masyuk, A.I., Masyuk, T.V., Tammachote, R., Huang, B.Q., Leontovich, A.A., Beito, T.G., Madden, B.J., Charlesworth, M.C., Torres, V.E., LaRusso, N.F., Harris, P.C. & Ward, C.J. (2009), "Characterization of PKD Protein-Positive Exosome-Like Vesicles", *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY*, vol. 20, no. 2, pp. 278-288.

Huang, D.W., Sherman, B.T. & Lempicki, R.A. (2009 a), "Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists", *NUCLEIC ACIDS RESEARCH*, vol. 37, no. 1, pp. 1-13.

Huang, D.W., Sherman, B.T. & Lempicki, R.A. (2009 b), "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources", *NATURE PROTOCOLS*, vol. 4, no. 1, pp. 44-57.

Huang, K., Chen, C., Chen, J. & Lin, W. (2003), "Statins induce suppressor of cytokine signaling-3 in macrophages", *FEBS LETTERS*, vol. 555, no. 2, pp. 385-389.

Ikeyama, S., Koyama, M., Yamaoko, M., Sasada, R. & Miyake, M. (1993), "Suppression of cell motility and metastasis by transfection with human motility-related protein (MRP-1/CD9) DNA", *JOURNAL OF EXPERIMENTAL MEDICINE*, vol. 177, no. 5, pp. 1231-1237.

Kang, R., Wan, J., Arstikaitis, P., Takahashi, H., Huang, K., Bailey, A.O., Thompson, J.X., Roth, A.F., Drisdel, R.C., Mastro, R., Green, W.N., Yates, I.,John R., Davis, N.G. & El-Husseini, A. (2008), "Neural palmitoyl-proteomics reveals dynamic synaptic palmitoylation", *NATURE*, vol. 456, no. 7224, pp. 904-909.

Katzmann, D., Babst, M. & Emr, S. (2001), "Ubiquitin-dependent sorting into the multivesicular body pathway requires the function of a conserved endosomal protein sorting complex, ESCRT-I", *CELL*, vol. 106, no. 2, pp. 145-155.

Lai, R., Tan, S., Teh, B., Sze, S., Arslan, F., de Kleijn, D., Choo, A. & Lim, S. (2012), "Proteolytic potential of the MSC exosome proteome: implications for an exosome-mediated delivery of therapeutic proteasome.", *INTERNATIONAL JOURNAL OF PROTEOMICS*, vol. 2012, pp. 971907.

Linder, M.E. & Deschenes, R.J. (2007), "Palmitoylation: policing protein stability and traffic", *NATURE REVIEWS MOLECULAR CELL BIOLOGY*, vol. 8, no. 1, pp. 74-84.

Mairhofer, M., Steiner, M., Mosgoeller, W., Prohaska, R. & Salzer, U. (2002), "Stomatin is a major lipid-raft component of platelet alpha granules", *BLOOD*, vol. 100, no. 3, pp. 897-904.

Martin, B.R., Wang, C., Adibekian, A., Tully, S.E. & Cravatt, B.F. (2012), "Global profiling of dynamic protein palmitoylation", *NATURE METHODS*, vol. 9, no. 1, pp. 84-U205.

Masellissmith, A. & Shaw, A. (1994), "CD9-regulated adhesion - Anti-CD9 monoclonal-antibody induce pre-B-cell adhesion to bone-marrow fibroblasts through de-novo recognition of fibronectin", *JOURNAL OF IMMUNOLOGY*, vol. 152, no. 6, pp. 2768-2777.

Mathivanan, S., Fahner, C.J., Reid, G.E. & Simpson, R.J. (2012), "ExoCarta 2012: database of exosomal proteins, RNA and lipids", *NUCLEIC ACIDS RESEARCH*, vol. 40, no. D1, pp. D1241-D1244.

Nachury, M.V., Seeley, E.S. & Jin, H. (2010), "Trafficking to the Ciliary Membrane: How to Get Across the Periciliary Diffusion Barrier?" in *ANNUAL REVIEW OF CELL AND DEVELOPMENTAL BIOLOGY, VOL 26*, eds. R. Schekman, L. Goldstein & R. Lehmann, pp. 59-87.

Pfanner, N., Glick, B., Arden, S. & Rothman, J. (1990), "Fatty acylation promotes fusion of transport vesicles with golgi cisternae", *JOURNAL OF CELL BIOLOGY*, vol. 110, no. 4, pp. 955-961.

Pisitkun, T., Shen, R. & Knepper, M. (2004), "Identification and proteomic profiling of exosomes in human urine", *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, vol. 101, no. 36, pp. 13368-13373.

Qanbar, R. & Bouvier, M. (2003), "Role of palmitoylation/depalmitoylation reactions in G-protein-coupled receptor function", *PHARMACOLOGY \& THERAPEUTICS*, vol. 97, no. 1, pp. 1-33.

Ren, J., Wen, L., Gao, X., Jin, C., Xue, Y. & Yao, X. (2008), "CSS-Palm 2.0: an updated software for palmitoylation sites prediction", *PROTEIN ENGINEERING DESIGN \& SELECTION*, vol. 21, no. 11, pp. 639-644.

Resh, M.D. 2006, "Palmitoylation of Ligands, Receptors, and Intracellular Signaling Molecules", *SCI.STKE*, vol. 2006, no. 359, pp. re14.

Ridker, P., Rifai, N., Pfeffer, M., Sacks, F., Braunwald, E. (1999), "Long-term effects of pravastatin on plasma concentration of C-reactive protein. Cholesterol \& Recurrent Events CARE Investigators ", *CIRCULATION*, vol. 100, no. 3, pp. 230-235.

Robbins, S., Quintrell, N. & Bishop, J. (1995), "Myristoylation and differential palmitoylation of the HCK protein-tyrosine kinases govern their attachment to membranes and association with caveolae", *MOLECULAR AND CELLULAR BIOLOGY*, vol. 15, no. 7, pp. 3507-3515.

Roth, A., Wan, J., Bailey, A., Sun, B., Kuchar, J., Green, W., Phinney, B., Yates, J. & Davis, N. (2006), "Global analysis of protein palmitoylation in yeast", *CELL*, vol. 125, no. 5, pp. 1003-1013.

Salaun, C., Greaves, J. & Chamberlain, L.H. (2010), "The intracellular dynamic of protein palmitoylation", *JOURNAL OF CELL BIOLOGY*, vol. 191, no. 7, pp. 1229-1238.

Shen, B., Wu, N., Yang, J. & Gould, S.J. (2011), "Protein Targeting to Exosomes/Microvesicles by Plasma Membrane Anchors", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 286, no. 16, pp. 14383-14395.

Shenoyscaria, A., Dietzen, D., Kwong, J., Link, D. & Lublin, D. (1994), "Cysteine(3) of SRC family protein-tyrosine kinases determines palmitoylation and localization in caveolae", *JOURNAL OF CELL BIOLOGY*, vol. 126, no. 2, pp. 353-363.

Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. (1996), "Mass spectrometric sequencing of proteins from silver stained polyacrylamide gels", *ANALYTICAL CHEMISTRY*, vol. 68, no. 5, pp. 850-858.

Sim, D.S., Dilks, J.R. & Flaumenhaft, R. (2007), "Platelets possess and require an active protein palmitoylation pathway for agonist-mediated activation and in vivo thrombus formation", *ARTERIOSCLEROSIS THROMBOSIS AND VASCULAR BIOLOGY*, vol. 27, no. 6, pp. 1478-1485.

Smotrys, J. & Linder, M. (2004), "Palmitoylation of intracellular signaling proteins: Regulation and function", *ANNUAL REVIEW OF BIOCHEMISTRY*, vol. 73, pp. 559-587.

Snyers, L., Umlauf, E. & Prohaska, R. (1999), "Cysteine 29 is the major palmitoylation site on stomatin", *FEBS LETTERS*, vol. 449, no. 2-3, pp. 101-104.

Stenvinkel, P., Ketteler, M., Johnson, R., Lindholm, B., Pecoits-Filho, R., Riella, M., Heimburger, O., Cederholm, T. & Girndt, M. (2005), "IL-10, IL-6, and TNF-alpha: Central factors in the altered cytokine network of uremia - The good, the bad, and the ugly", *KIDNEY INTERNATIONAL*, vol. 67, no. 4, pp. 1216-1233.

Tyagarajan, K., Pretzer, E. & Wiktorowicz, J. (2003), "Thiol-reactive dyes for fluorescence labeling of proteomic samples", *ELECTROPHORESIS*, vol. 24, no. 14, pp. 2348-2358.

Walsh, C., Garneau-Tsodikova, S. & Gatto, G. (2005), "Protein posttranslational modifications: The chemistry of proteome diversifications", *ANGEWANDTE CHEMIE-INTERNATIONAL EDITION*, vol. 44, no. 45, pp. 7342-7372.

Wang, Z., Hill, S., Luther, J.M., Hachey, D.L. & Schey, K.L. (2011), "Proteomic analysis of urine exosomes by multidimensional protein identification technology (MudPIT)", *PROTEOMICS*, vol.12, no. 2, pp. 329-338.

Yang, W., Di Vizio, D., Kirchner, M., Steen, H. & Freeman, M.R. (2010), "Proteome scale characterization of human S-acylated proteins in lipid raft-enriched and non-raft membranes", *MOLECULAR \& CELLULAR PROTEOMICS*, vol. 9, no. 1, pp. 54-70.

Zhao, S. & Zhang, D. (2003), "Atorvastatin reduces interleukin-6 plasma concentration and adipocyte secretion of hypercholesterolemic rabbits", *CLINICA CHIMICA ACTA*, vol. 336, no. 1-2, pp. 103-108.

Zheng, H., Pearsall, E.A., Hurst, D.P., Zhang, Y., Chu, J., Zhou, Y., Reggio, P.H., Loh, H.H. & Law, P. (2012), "Palmitoylation and membrane cholesterol stabilize mu-opioid receptor homodimerization and G protein coupling", *BMC CELL BIOLOGY*, vol. 13, pp. 6.

Supplementary Table 6.1: Proteins identified in HA+ fraction. Uniprot accession, protein description, score, molecular weight (in Dalton), number of protein matches and prediction of palmitoylation or known status are indictaed in the table.

Accession	Description	score	Mol wt.	Prot matches	Known palmitoylated or predicted (CSS PALM3.0)
P27348	14-3-3 protein theta OS=Homo sapiens GN=YWHAQ PE=1 SV=1	52	28032	1	Known & Predicted
P63104	14-3-3 protein zeta/delta OS=Homo sapiens GN=YWHAZ PE=1 SV=1	87	27899	2	None
Q9BUT1	3-hydroxybutyrate dehydrogenase type 2 OS=Homo sapiens GN=BDH2 PE=1 SV=2	58	27049	1	Predicted
P25325	3-mercaptopyruvate sulfurtransferase OS=Homo sapiens GN=MPST PE=1 SV=3	41	33443	1	Known & Predicted
Q13510	Acid ceramidase OS=Homo sapiens GN=ASAH1 PE=1 SV=5	73	45087	2	Known & Predicted
O00468	Agrin OS=Homo sapiens GN=AGRN PE=1 SV=4	104	222861	2	Predicted
P01009	Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3	101	46878	2	Predicted
P02765	Alpha-2-HS-glycoprotein OS=Homo sapiens GN=AHSG PE=1 SV=1	51	40098	1	Predicted
P54802	Alpha-N-acetylglucosaminidase OS=Homo sapiens GN=NAGLU PE=1 SV=2	213	82670	5	Common with HA-
P54920	Alpha-soluble NSF attachment protein OS=Homo sapiens GN=NAPA PE=1 SV=3	56	33667	1	None
P19801	Amiloride-sensitive amine oxidase [copper-containing] OS=Homo sapiens GN=ABP1 PE=1 SV=4	45	85723	1	None
P15144	Aminopeptidase N OS=Homo sapiens GN=ANPEP PE=1 SV=4	335	109870	10	None
P12821	Angiotensin-converting enzyme OS=Homo sapiens GN=ACE PE=1 SV=1	45	150418	1	None
P01019	Angiotensinogen OS=Homo sapiens GN=AGT PE=1 SV=1	54	53406	1	None
P04083	Annexin A1 OS=Homo sapiens GN=ANXA1 PE=1 SV=2	143	38918	2	Predicted
P08758	Annexin A5 OS=Homo sapiens GN=ANXA5 PE=1 SV=2	109	35971	3	Predicted
P05090	Apolipoprotein D OS=Homo sapiens GN=APOD PE=1 SV=1	134	21547	3	Common with HA-
P02649	Apolipoprotein E OS=Homo sapiens GN=APOE PE=1 SV=1	225	36246	6	Predicted
P15291	Beta-1,4-galactosyltransferase 1 OS=Homo sapiens GN=B4GALT1 PE=1 SV=5	42	44291	1	Known & Predicted
P16278	Beta-galactosidase OS=Homo sapiens GN=GLB1 PE=1 SV=2	128	76483	4	None
P08236	Beta-glucuronidase OS=Homo sapiens GN=GUSB PE=1 SV=2	53	75027	1	Predicted
Q93088	Betaine--homocysteine S-methyltransferase 1 OS=Homo sapiens GN=BHMT PE=1 SV=2	44	45426	1	Predicted
Q9Y6W3	Calpain-7 OS=Homo sapiens GN=CAPN7 PE=1 SV=1	41	93335	1	Predicted
P31944	Caspase-14 OS=Homo sapiens GN=CASP14 PE=1 SV=2	80	27947	2	Common with HA-
P07339	Cathepsin D OS=Homo sapiens GN=CTSD PE=1 SV=1	46	45037	1	Predicted
Q9UBR2	Cathepsin Z OS=Homo sapiens GN=CTSZ PE=1 SV=1	60	34530	1	None
P21926	CD9 antigen OS=Homo sapiens GN=CD9 PE=1 SV=4	47	25969	1	Known & Predicted
P00450	Ceruloplasmin OS=Homo sapiens GN=CP PE=1 SV=1	68	122983	2	None

P10909	Clusterin OS=Homo sapiens GN=CLU PE=1 SV=1	88	53031	2	None
Q9Y281	Cofilin-2 OS=Homo sapiens GN=CFL2 PE=1 SV=1	63	18839	1	None
P12109	Collagen alpha-1(VI) chain OS=Homo sapiens GN=COL6A1 PE=1 SV=3	71	109602	2	Predicted
P39059	Collagen alpha-1(XV) chain OS=Homo sapiens GN=COL15A1 PE=1 SV=2	79	142202	2	Common with HA-
P01024	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2	113	188569	3	Common with HA-
Q8IYJ1	Copine-9 OS=Homo sapiens GN=CPNE9 PE=1 SV=3	40	62281	1	None
O60494	Cubilin OS=Homo sapiens GN=CUBN PE=1 SV=5	610	407262	17	Common with HA-
P04080	Cystatin-B OS=Homo sapiens GN=CSTB PE=1 SV=2	72	11190	2	Predicted
Q969T7	Cytosolic 5~-nucleotidase III-like protein OS=Homo sapiens GN=NT5C3L PE=1 SV=3	41	33792	1	Predicted
P81605	Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2	71	11391	1	Common with HA-
Q08554	Desmocollin-1 OS=Homo sapiens GN=DSC1 PE=1 SV=2	116	101406	3	Common with HA-
Q02413	Desmoglein-1 OS=Homo sapiens GN=DSG1 PE=1 SV=2	272	114702	7	Common with HA-
P15924	Desmoplakin OS=Homo sapiens GN=DSP PE=1 SV=3	210	334021	5	Common with HA-
P16444	Dipeptidase 1 OS=Homo sapiens GN=DPEP1 PE=1 SV=3	141	46101	4	Predicted
P53634	Dipeptidyl peptidase 1 OS=Homo sapiens GN=CTSC PE=1 SV=2	48	52619	1	None
P27487	Dipeptidyl peptidase 4 OS=Homo sapiens GN=DPP4 PE=1 SV=2	94	88907	4	Predicted
Q03001	Dystonin OS=Homo sapiens GN=DST PE=1 SV=4	42	865259	1	Predicted
Q12805	EGF-containing fibulin-like extracellular matrix protein 1 OS=Homo sapiens GN=EFEMP1 PE=1 SV=2	242	56885	4	Predicted
Q9H223	EH domain-containing protein 4 OS=Homo sapiens GN=EHD4 PE=1 SV=1	46	61365	1	None
Q15375	Ephrin type-A receptor 7 OS=Homo sapiens GN=EPHA7 PE=1 SV=3	86	113735	1	Predicted
P27105	Erythrocyte band 7 integral membrane protein OS=Homo sapiens GN=STOM PE=1 SV=3	119	31882	2	Known & Predicted
P08294	Extracellular superoxide dismutase [Cu-Zn] OS=Homo sapiens GN=SOD3 PE=1 SV=2	53	26291	1	Predicted
P15311	Ezrin OS=Homo sapiens GN=EZR PE=1 SV=4	79	69484	2	None
Q01469	Fatty acid-binding protein, epidermal OS=Homo sapiens GN=FABP5 PE=1 SV=3	45	15497	1	Common with HA-
P02794	Ferritin heavy chain OS=Homo sapiens GN=FTH1 PE=1 SV=2	48	21383	1	None
P02671	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	58	95656	1	Predicted
P02675	Fibrinogen beta chain OS=Homo sapiens GN=FGB PE=1 SV=2	47	56577	1	Predicted
Q5D862	Filaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1	67	249296	1	Common with HA-
P05062	Fructose-bisphosphate aldolase B OS=Homo sapiens GN=ALDOB PE=1 SV=2	81	39961	2	Predicted
Q08380	Galectin-3-binding protein OS=Homo sapiens GN=LGALS3BP PE=1 SV=1	96	66202	3	None
Q92820	Gamma-glutamyl hydrolase OS=Homo sapiens GN=GGH PE=1 SV=2	77	36340	2	Predicted
P06396	Gelsolin OS=Homo sapiens GN=GSN PE=1 SV=1	79	86043	2	Common with HA-
P08263	Glutathione S-transferase A1 OS=Homo sapiens GN=GSTA1 PE=1 SV=3	83	25672	2	None
P21266	Glutathione S-transferase Mu 3 OS=Homo sapiens GN=GSTM3 PE=1 SV=3	104	26998	3	Predicted

P09211	Glutathione S-transferase P OS=Homo sapiens GN=GSTP1 PE=1 SV=2	79	23569	1	Predicted
Q9NZH0	G-protein coupled receptor family C group 5 member B OS=Homo sapiens GN=GPRC5B PE=2 SV=2	53	45279	1	Predicted
Q9NQ84	G-protein coupled receptor family C group 5 member C OS=Homo sapiens GN=GPRC5C PE=1 SV=2	96	48732	2	Predicted
Q14393	Growth arrest-specific protein 6 OS=Homo sapiens GN=GAS6 PE=1 SV=2	114	81678	2	Predicted
P62873	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1 OS=Homo sapiens GN=GNB1 PE=1 SV=3	55	38151	1	Predicted
P00738	Haptoglobin OS=Homo sapiens GN=HP PE=1 SV=1	246	45861	7	Common with HA-
P08107	Heat shock 70 kDa protein 1A/1B OS=Homo sapiens GN=HSPA1A PE=1 SV=5	67	70294	2	Predicted
P04792	Heat shock protein beta-1 OS=Homo sapiens GN=HSPB1 PE=1 SV=2	42	22826	1	None
P68871	Hemoglobin subunit beta OS=Homo sapiens GN=HBB PE=1 SV=2	53	16102	1	None
P02790	Hemopexin OS=Homo sapiens GN=HPX PE=1 SV=2	45	52385	1	Predicted
P04196	Histidine-rich glycoprotein OS=Homo sapiens GN=HRG PE=1 SV=1	41	60510	1	Predicted
Q86YZ3	Hornerin OS=Homo sapiens GN=HRNR PE=1 SV=2	75	283140	2	Common with HA-
P01876	Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=2	141	38486	3	Common with HA-
P01859	Ig gamma-2 chain C region OS=Homo sapiens GN=IGHG2 PE=1 SV=2	98	36505	3	Common with HA-
P01860	Ig gamma-3 chain C region OS=Homo sapiens GN=IGHG3 PE=1 SV=2	184	42287	5	None
P01825	Ig heavy chain V-II region NEWM OS=Homo sapiens PE=1 SV=1	43	12953	1	None
P01767	Ig heavy chain V-III region BUT OS=Homo sapiens PE=1 SV=1	96	12485	3	None
P01768	Ig heavy chain V-III region CAM OS=Homo sapiens PE=1 SV=1	95	13773	2	None
P01781	Ig heavy chain V-III region GAL OS=Homo sapiens PE=1 SV=1	84	12836	2	None
P01765	Ig heavy chain V-III region TIL OS=Homo sapiens PE=1 SV=1	85	12462	2	None
P01779	Ig heavy chain V-III region TUR OS=Homo sapiens PE=1 SV=1	90	12537	2	None
P01764	Ig heavy chain V-III region VH26 OS=Homo sapiens PE=1 SV=1	87	12745	2	None
P01834	Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 SV=1	73	11773	1	None
P01593	Ig kappa chain V-I region AG OS=Homo sapiens PE=1 SV=1	96	12099	1	None
P01598	Ig kappa chain V-I region EU OS=Homo sapiens PE=1 SV=1	169	11895	3	Predicted
P01613	Ig kappa chain V-I region Ni OS=Homo sapiens PE=1 SV=1	48	12352	1	Predicted
P01611	Ig kappa chain V-I region Wes OS=Homo sapiens PE=1 SV=1	65	11715	1	Predicted
P01614	Ig kappa chain V-II region Cum OS=Homo sapiens PE=1 SV=1	81	12782	1	Common with HA-
P01620	Ig kappa chain V-III region SIE OS=Homo sapiens PE=1 SV=1	89	11882	1	Predicted
P01625	Ig kappa chain V-IV region Len OS=Homo sapiens PE=1 SV=2	47	12746	1	Predicted
P80748	Ig lambda chain V-III region LOI OS=Homo sapiens PE=1 SV=1	96	12042	1	None
P01717	Ig lambda chain V-IV region Hil OS=Homo sapiens PE=1 SV=1	94	11624	2	None
P0CG05	Ig lambda-2 chain C regions OS=Homo sapiens GN=IGLC2 PE=1 SV=1	41	11458	1	None

P01871	Ig mu chain C region OS=Homo sapiens GN=IGHM PE=1 SV=3	115	49960	3	Common with HA-
Q9Y6R7	IgGFc-binding protein OS=Homo sapiens GN=FCGBP PE=1 SV=3	315	596443	9	Predicted
P01591	Immunoglobulin J chain OS=Homo sapiens GN=IGJ PE=1 SV=4	142	18543	4	Common with HA-
Q14624	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	229	103521	7	Common with HA-
O75874	Isocitrate dehydrogenase [NADP] cytoplasmic OS=Homo sapiens GN=IDH1 PE=1 SV=2	63	46915	1	None
P53990	IST1 homolog OS=Homo sapiens GN=IST1 PE=1 SV=1	79	39897	2	None
P14923	Junction plakoglobin OS=Homo sapiens GN=JUP PE=1 SV=3	97	82434	3	Common with HA-
P13645	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6	746	59020	17	Common with HA-
P13646	Keratin, type I cytoskeletal 13 OS=Homo sapiens GN=KRT13 PE=1 SV=4	344	49900	10	Common with HA-
P02533	Keratin, type I cytoskeletal 14 OS=Homo sapiens GN=KRT14 PE=1 SV=4	533	51872	14	Common with HA-
P19012	Keratin, type I cytoskeletal 15 OS=Homo sapiens GN=KRT15 PE=1 SV=3	202	49409	6	Common with HA-
P08779	Keratin, type I cytoskeletal 16 OS=Homo sapiens GN=KRT16 PE=1 SV=4	491	51578	14	Common with HA-
Q04695	Keratin, type I cytoskeletal 17 OS=Homo sapiens GN=KRT17 PE=1 SV=2	215	48361	7	Common with HA-
P08727	Keratin, type I cytoskeletal 19 OS=Homo sapiens GN=KRT19 PE=1 SV=4	97	44079	3	Common with HA-
P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3	348	62255	9	Common with HA-
P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6	938	66170	25	Common with HA-
P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2	727	65678	20	Common with HA-
P19013	Keratin, type II cytoskeletal 4 OS=Homo sapiens GN=KRT4 PE=1 SV=4	116	57649	3	Common with HA-
P13647	Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 PE=1 SV=3	361	62568	12	Common with HA-
P02538	Keratin, type II cytoskeletal 6A OS=Homo sapiens GN=KRT6A PE=1 SV=3	482	60293	15	Common with HA-
P04259	Keratin, type II cytoskeletal 6B OS=Homo sapiens GN=KRT6B PE=1 SV=5	397	60315	14	Common with HA-
P48668	Keratin, type II cytoskeletal 6C OS=Homo sapiens GN=KRT6C PE=1 SV=3	478	60273	15	Common with HA-
Q8N1N4	Keratin, type II cytoskeletal 78 OS=Homo sapiens GN=KRT78 PE=2 SV=2	65	57629	2	Common with HA-
P05787	Keratin, type II cytoskeletal 8 OS=Homo sapiens GN=KRT8 PE=1 SV=7	99	53671	3	Common with HA-
P01042	Kininogen-1 OS=Homo sapiens GN=KNG1 PE=1 SV=2	157	72996	5	Common with HA-
P07195	L-lactate dehydrogenase B chain OS=Homo sapiens GN=LDHB PE=1 SV=2	126	36900	4	None
P98164	Low-density lipoprotein receptor-related protein 2 OS=Homo sapiens GN=LRP2 PE=1 SV=3	756	540376	21	Common with HA-
P10619	Lysosomal protective protein OS=Homo sapiens GN=CTSA PE=1 SV=2	77	54944	1	Predicted
O43451	Maltase-glucoamylase, intestinal OS=Homo sapiens GN=MGAM PE=1 SV=5	195	211031	6	Common with HA-
O00187	Mannan-binding lectin serine protease 2 OS=Homo sapiens GN=MASP2 PE=1 SV=4	129	77193	3	None
P33908	Mannosyl-oligosaccharide 1,2-alpha-mannosidase IA OS=Homo sapiens GN=MAN1A1 PE=1 SV=3	69	73150	2	None
Q9BRK3	Matrix-remodeling-associated protein 8 OS=Homo sapiens GN=MXRA8 PE=1 SV=1	75	49500	2	None
Q6W4X9	Mucin-6 OS=Homo sapiens GN=MUC6 PE=1 SV=3	47	263159	1	Predicted

Q9H8L6	Multimerin-2 OS=Homo sapiens GN=MMRN2 PE=1 SV=2	70	105028	1	Predicted
O96009	Napsin-A OS=Homo sapiens GN=NAPSA PE=1 SV=1	74	45700	2	Common with HA-
P08473	Neprilysin OS=Homo sapiens GN=MME PE=1 SV=2	198	86144	5	Predicted
P59665	Neutrophil defensin 1 OS=Homo sapiens GN=DEFA1 PE=1 SV=1	92	10536	2	Common with HA-
P61970	Nuclear transport factor 2 OS=Homo sapiens GN=NUTF2 PE=1 SV=1	94	14640	2	Predicted
Q6UX06	Olfactomedin-4 OS=Homo sapiens GN=OLFM4 PE=1 SV=1	261	57529	5	Common with HA-
P04746	Pancreatic alpha-amylase OS=Homo sapiens GN=AMY2A PE=1 SV=2	178	58354	4	Known & Predicted
Q06830	Peroxiredoxin-1 OS=Homo sapiens GN=PRDX1 PE=1 SV=1	51	22324	1	None
P30041	Peroxiredoxin-6 OS=Homo sapiens GN=PRDX6 PE=1 SV=3	69	25133	2	None
P30086	Phosphatidylethanolamine-binding protein 1 OS=Homo sapiens GN=PEBP1 PE=1 SV=3	126	21158	2	None
P05154	Plasma serine protease inhibitor OS=Homo sapiens GN=SERPINA5 PE=1 SV=3	201	45760	6	Common with HA-
P01833	Polymeric immunoglobulin receptor OS=Homo sapiens GN=PIGR PE=1 SV=4	235	84429	5	Common with HA-
P0CG38	POTE ankyrin domain family member I OS=Homo sapiens GN=POTEI PE=3 SV=1	50	122858	1	Predicted
P01133	Pro-epidermal growth factor OS=Homo sapiens GN=EGF PE=1 SV=2	398	137613	11	Common with HA-
Q8WUM4	Programmed cell death 6-interacting protein OS=Homo sapiens GN=PDCD6IP PE=1 SV=1	199	96590	6	None
P12273	Prolactin-inducible protein OS=Homo sapiens GN=PIP PE=1 SV=1	57	16847	1	Common with HA-
P41222	Prostaglandin-H2 D-isomerase OS=Homo sapiens GN=PTGDS PE=1 SV=1	49	21243	1	Predicted
P07288	Prostate-specific antigen OS=Homo sapiens GN=KLK3 PE=1 SV=2	48	29293	1	None
P15309	Prostatic acid phosphatase OS=Homo sapiens GN=ACPP PE=1 SV=3	41	44880	1	Common with HA-
P60900	Proteasome subunit alpha type-6 OS=Homo sapiens GN=PSMA6 PE=1 SV=1	47	27838	1	Known & Predicted
P02760	Protein AMBP OS=Homo sapiens GN=AMBP PE=1 SV=1	165	39886	3	Common with HA-
Q969X1	Protein lifeguard 3 OS=Homo sapiens GN=TMBIM1 PE=1 SV=2	67	34927	1	Predicted
Q8WVV4	Protein POF1B OS=Homo sapiens GN=POF1B PE=1 SV=3	49	68878	1	Predicted
P31151	Protein S100-A7 OS=Homo sapiens GN=S100A7 PE=1 SV=4	54	11578	1	Common with HA-
P05109	Protein S100-A8 OS=Homo sapiens GN=S100A8 PE=1 SV=1	143	10885	4	Common with HA-
P06702	Protein S100-A9 OS=Homo sapiens GN=S100A9 PE=1 SV=1	124	13291	2	Common with HA-
P49221	Protein-glutamine gamma-glutamyltransferase 4 OS=Homo sapiens GN=TGM4 PE=1 SV=2	134	77951	3	Predicted
A6NMY6	Putative annexin A2-like protein OS=Homo sapiens GN=ANXA2P2 PE=5 SV=2	128	38806	3	None
Q5VTE0	Putative elongation factor 1-alpha-like 3 OS=Homo sapiens GN=EEF1A1P5 PE=5 SV=1	66	50495	2	Predicted
P14618	Pyruvate kinase isozymes M1/M2 OS=Homo sapiens GN=PKM2 PE=1 SV=4	92	58470	2	Predicted
P61026	Ras-related protein Rab-10 OS=Homo sapiens GN=RAB10 PE=1 SV=1	71	22755	2	Known & Predicted
P59190	Ras-related protein Rab-15 OS=Homo sapiens GN=RAB15 PE=1 SV=1	51	24660	1	None
Q9H082	Ras-related protein Rab-33B OS=Homo sapiens GN=RAB33B PE=1 SV=1	51	26043	1	Predicted

Q9NRW1	Ras-related protein Rab-6B OS=Homo sapiens GN=RAB6B PE=1 SV=1	51	23561	1	Known & Predicted
P62070	Ras-related protein R-Ras2 OS=Homo sapiens GN=RRAS2 PE=1 SV=1	56	23613	1	Known & Predicted
P04279	Semenogelin-1 OS=Homo sapiens GN=SEMG1 PE=1 SV=2	138	52157	3	Common with HA-
P02787	Serotransferrin OS=Homo sapiens GN=TF PE=1 SV=3	486	79294	16	Common with HA-
P29508	Serpin B3 OS=Homo sapiens GN=SERPINB3 PE=1 SV=2	83	44594	2	Common with HA-
P02768	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2	446	71317	13	Common with HA-
P02743	Serum amyloid P-component OS=Homo sapiens GN=APCS PE=1 SV=2	56	25485	1	None
Q13621	Solute carrier family 12 member 1 OS=Homo sapiens GN=SLC12A1 PE=1 SV=2	53	122627	1	Predicted
P22732	Solute carrier family 2, facilitated glucose transporter member 5 OS=Homo sapiens GN=SLC2A5 PE=1 SV=1	83	55394	2	Predicted
Q92673	Sortilin-related receptor OS=Homo sapiens GN=SORL1 PE=1 SV=2	59	253798	1	Predicted
O00391	Sulfhydryl oxidase 1 OS=Homo sapiens GN=QSOX1 PE=1 SV=3	41	83324	1	Predicted
O60635	Tetraspanin-1 OS=Homo sapiens GN=TSPAN1 PE=1 SV=2	47	26910	1	Predicted
P10599	Thioredoxin OS=Homo sapiens GN=TXN PE=1 SV=3	79	12015	1	Common with HA-
P07996	Thrombospondin-1 OS=Homo sapiens GN=THBS1 PE=1 SV=2	184	133291	5	Predicted
Q9UKU6	Thyrotropin-releasing hormone-degrading ectoenzyme OS=Homo sapiens GN=TRHDE PE=2 SV=1	76	117439	2	Predicted
P04066	Tissue alpha-L-fucosidase OS=Homo sapiens GN=FUCA1 PE=1 SV=4	76	53940	1	None
O75264	Transmembrane protein C19orf77 OS=Homo sapiens GN=C19orf77 PE=2 SV=2	59	15012	1	Predicted
P60174	Triosephosphate isomerase OS=Homo sapiens GN=TPI1 PE=1 SV=3	50	31057	1	Predicted
O14773	Tripeptidyl-peptidase 1 OS=Homo sapiens GN=TPP1 PE=1 SV=2	108	61723	2	Predicted
P62979	Ubiquitin-40S ribosomal protein S27a OS=Homo sapiens GN=RPS27A PE=1 SV=2	70	18296	1	Known & Predicted
Q8IX04	Ubiquitin-conjugating enzyme E2 variant 3 OS=Homo sapiens GN=UEVLD PE=1 SV=2	50	52516	1	Predicted
P00749	Urokinase-type plasminogen activator OS=Homo sapiens GN=PLAU PE=1 SV=2	141	49901	4	Predicted
P07911	Uromodulin OS=Homo sapiens GN=UMOD PE=1 SV=1	330	72451	8	Common with HA-
O00322	Uroplakin-1a OS=Homo sapiens GN=UPK1A PE=2 SV=1	44	29429	1	Predicted
Q9UK41	Vacuolar protein sorting-associated protein 28 homolog OS=Homo sapiens GN=VPS28 PE=1 SV=1	52	25694	1	None
Q9UN37	Vacuolar protein sorting-associated protein 4A OS=Homo sapiens GN=VPS4A PE=1 SV=1	63	49152	1	None
Q12907	Vesicular integral-membrane protein VIP36 OS=Homo sapiens GN=LMAN2 PE=1 SV=1	190	40545	5	Common with HA-
Q7Z5L0	Vitelline membrane outer layer protein 1 homolog OS=Homo sapiens GN=VMO1 PE=1 SV=1	178	22034	3	Common with HA-
P04004	Vitronectin OS=Homo sapiens GN=VTN PE=1 SV=1	66	55069	1	Predicted
P25311	Zinc-alpha-2-glycoprotein OS=Homo sapiens GN=AZGP1 PE=1 SV=2	54	34465	1	None

Supplementary Table 6.2: Proteins identified in HA- fraction. Uniprot accession, protein description, score, molecular weight (in Dalton), number of protein matches and prediction of palmitoylation or known status are indictaed in the table.

Accession	Description	Score	Mol wt.	prot_matches
P54802	Alpha-N-acetylglucosaminidase OS=Homo sapiens GN=NAGLU PE=1 SV=2	138	82670	2
P05090	Apolipoprotein D OS=Homo sapiens GN=APOD PE=1 SV=1	119	21547	3
Q9NZT1	Calmodulin-like protein 5 OS=Homo sapiens GN=CALML5 PE=1 SV=2	45	15883	1
P31944	Caspase-14 OS=Homo sapiens GN=CASP14 PE=1 SV=2	67	27947	1
Q8IWA5	Choline transporter-like protein 2 OS=Homo sapiens GN=SLC44A2 PE=1 SV=3	40	81610	1
P39059	Collagen alpha-1(XV) chain OS=Homo sapiens GN=COL15A1 PE=1 SV=2	63	142202	1
P01024	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2	45	188569	1
O60494	Cubilin OS=Homo sapiens GN=CUBN PE=1 SV=5	200	407262	3
P81605	Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2	68	11391	1
Q6E0U4	Dermokine OS=Homo sapiens GN=DMKN PE=1 SV=3	60	47282	1
Q08554	Desmocollin-1 OS=Homo sapiens GN=DSC1 PE=1 SV=2	74	101406	2
Q02413	Desmoglein-1 OS=Homo sapiens GN=DSG1 PE=1 SV=2	120	114702	3
P15924	Desmoplakin OS=Homo sapiens GN=DSP PE=1 SV=3	122	334021	3
Q01469	Fatty acid-binding protein, epidermal OS=Homo sapiens GN=FABP5 PE=1 SV=3	50	15497	1
Q5D862	Filaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1	78	249296	1
O75223	Gamma-glutamylcyclotransferase OS=Homo sapiens GN=GGCT PE=1 SV=1	53	21222	1
P04406	Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens GN=GAPDH PE=1 SV=3	68	36201	1
P00738	Haptoglobin OS=Homo sapiens GN=HP PE=1 SV=1	104	45861	2
Q86YZ3	Hornerin OS=Homo sapiens GN=HRNR PE=1 SV=2	162	283140	4
P01876	Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=2	132	38486	3
P01877	Ig alpha-2 chain C region OS=Homo sapiens GN=IGHA2 PE=1 SV=3	46	37301	1
P01857	Ig gamma-1 chain C region OS=Homo sapiens GN=IGHG1 PE=1 SV=1	146	36596	2
P01859	Ig gamma-2 chain C region OS=Homo sapiens GN=IGHG2 PE=1 SV=2	76	36505	1
P01614	Ig kappa chain V-II region Cum OS=Homo sapiens PE=1 SV=1	69	12782	1
P0CF74	Ig lambda-6 chain C region OS=Homo sapiens GN=IGLC6 PE=4 SV=1	57	11441	1
P01871	Ig mu chain C region OS=Homo sapiens GN=IGHM PE=1 SV=3	59	49960	1
P01591	Immunoglobulin J chain OS=Homo sapiens GN=IGJ PE=1 SV=4	102	18543	2
Q14624	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	104	103521	2
P14923	Junction plakoglobin OS=Homo sapiens GN=JUP PE=1 SV=3	118	82434	3
P13645	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6	731	59020	16
P13646	Keratin, type I cytoskeletal 13 OS=Homo sapiens GN=KRT13 PE=1 SV=4	188	49900	5
P02533	Keratin, type I cytoskeletal 14 OS=Homo sapiens GN=KRT14 PE=1 SV=4	557	51872	12
P19012	Keratin, type I cytoskeletal 15 OS=Homo sapiens GN=KRT15 PE=1 SV=3	166	49409	5

P08779	Keratin, type I cytoskeletal 16 OS=Homo sapiens GN=KRT16 PE=1 SV=4	514	51578	12
Q04695	Keratin, type I cytoskeletal 17 OS=Homo sapiens GN=KRT17 PE=1 SV=2	345	48361	9
P08727	Keratin, type I cytoskeletal 19 OS=Homo sapiens GN=KRT19 PE=1 SV=4	139	44079	4
P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3	732	62255	16
P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6	1202	66170	27
P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2	1020	65678	25
P12035	Keratin, type II cytoskeletal 3 OS=Homo sapiens GN=KRT3 PE=1 SV=3	144	64549	3
P19013	Keratin, type II cytoskeletal 4 OS=Homo sapiens GN=KRT4 PE=1 SV=4	143	57649	4
P13647	Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 PE=1 SV=3	615	62568	16
P02538	Keratin, type II cytoskeletal 6A OS=Homo sapiens GN=KRT6A PE=1 SV=3	546	60293	17
P04259	Keratin, type II cytoskeletal 6B OS=Homo sapiens GN=KRT6B PE=1 SV=5	472	60315	15
Q14CN4	Keratin, type II cytoskeletal 72 OS=Homo sapiens GN=KRT72 PE=1 SV=2	78	56470	2
Q7RTS7	Keratin, type II cytoskeletal 74 OS=Homo sapiens GN=KRT74 PE=1 SV=2	111	58229	3
Q8N1N4	Keratin, type II cytoskeletal 78 OS=Homo sapiens GN=KRT78 PE=2 SV=2	128	57629	4
P05787	Keratin, type II cytoskeletal 8 OS=Homo sapiens GN=KRT8 PE=1 SV=7	45	53671	1
P01042	Kininogen-1 OS=Homo sapiens GN=KNG1 PE=1 SV=2	49	72996	1
P98164	Low-density lipoprotein receptor-related protein 2 OS=Homo sapiens GN=LRP2 PE=1 SV=3	163	540376	4
O43451	Maltase-glucoamylase, intestinal OS=Homo sapiens GN=MGAM PE=1 SV=5	51	211031	1
O00187	Mannan-binding lectin serine protease 2 OS=Homo sapiens GN=MASP2 PE=1 SV=4	102	77193	1
O96009	Napsin-A OS=Homo sapiens GN=NAPSA PE=1 SV=1	53	45700	1
P59665	Neutrophil defensin 1 OS=Homo sapiens GN=DEFA1 PE=1 SV=1	97	10536	2
Q6UX06	Olfactomedin-4 OS=Homo sapiens GN=OLFM4 PE=1 SV=1	100	57529	2
P05154	Plasma serine protease inhibitor OS=Homo sapiens GN=SERPINA5 PE=1 SV=3	148	45760	2
P01833	Polymeric immunoglobulin receptor OS=Homo sapiens GN=PIGR PE=1 SV=4	140	84429	2
P01133	Pro-epidermal growth factor OS=Homo sapiens GN=EGF PE=1 SV=2	286	137613	8
P12273	Prolactin-inducible protein OS=Homo sapiens GN=PIP PE=1 SV=1	80	16847	1
P15309	Prostatic acid phosphatase OS=Homo sapiens GN=ACPP PE=1 SV=3	100	44880	2
P02760	Protein AMBP OS=Homo sapiens GN=AMBP PE=1 SV=1	128	39886	3
P31151	Protein S100-A7 OS=Homo sapiens GN=S100A7 PE=1 SV=4	52	11578	1
P05109	Protein S100-A8 OS=Homo sapiens GN=S100A8 PE=1 SV=1	95	10885	2
P06702	Protein S100-A9 OS=Homo sapiens GN=S100A9 PE=1 SV=1	134	13291	3
A6NMY6	Putative annexin A2-like protein OS=Homo sapiens GN=ANXA2P2 PE=5 SV=2	94	38806	2
A8MT79	Putative zinc-alpha-2-glycoprotein-like 1 OS=Homo sapiens PE=5 SV=2	46	23080	1
P04279	Semenogelin-1 OS=Homo sapiens GN=SEMG1 PE=1 SV=2	47	52157	1
P02787	Serotransferrin OS=Homo sapiens GN=TF PE=1 SV=3	145	79294	3
Q96P63	Serpin B12 OS=Homo sapiens GN=SERPINB12 PE=1 SV=1	49	46646	1
P29508	Serpin B3 OS=Homo sapiens GN=SERPINB3 PE=1 SV=2	58	44594	1
P02768	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2	277	71317	7

Q9H2G4	Testis-specific Y-encoded-like protein 2 OS=Homo sapiens GN=TSPYL2 PE=1 SV=1	46	79615	1
P10599	Thioredoxin OS=Homo sapiens GN=TXN PE=1 SV=3	78	12015	1
P07911	Uromodulin OS=Homo sapiens GN=UMOD PE=1 SV=1	467	72451	10
Q12907	Vesicular integral-membrane protein VIP36 OS=Homo sapiens GN=LMAN2 PE=1 SV=1	128	40545	2
Q7Z5L0	Vitelline membrane outer layer protein 1 homolog OS=Homo sapiens GN=VMO1 PE=1 SV=1	73	22034	1

CONCLUSIONS AND FUTURE WORK

This thesis presented an alternative method for removing contamination of high-abundance proteins while isolating urinary membrane vesicles (40-100nm). Although they have been called exosomes in the literature, multiple studies have emphasised that the differential centrifugation method enriches vesicles of various types and not only exosomes. The alternative method we have developed for removing contamination with high-abundance proteins while isolating urinary membrane vesicles involved a mild detergent (CHAPS) treatment. This method is also better at preserving the biological activity of protein constituents of urinary membrane vesicles, as demonstrated in case of DPP-IV and nephrilysin. This method would be superior to the methods published by previous workers (employing DTT) if an activity-based assay or biomarkers is to be developed. Building on this method proteomic analysis of the isolated urinary membrane vesicles was carried out. The methodology used was developed with minimum number of steps possible and 437 proteins in the high-speed pellet and 155 proteins in the low speed pellet were identified. Twenty % of the proteins identified in the high-speed pellet were previously not described as being part of exosomes which validates our method. This method is relatively easy to perform and less time and labor-consuming. Therefore, this method can be applied to large number of samples for qualitative and quantitative proteomic analysis of these vesicles. This can also be applied on clinical samples for quantitative studies for biomarker discovery.

Most of the studies in urine have been focussed on smaller vesicles (Exosomes and ‘exosome-like’ vesicles) and bigger vesicles like microvesicles and microparticles have been largely ignored. Microvesicles also have the potential to provide a different set of biomarkers for various diseases because they are mainly released upon cell stimulation. In a diseased state, abnormal cell stimulation can occur and microvesicles can provide a non-invasive way

of detecting and monitoring these changes. However, to the best of our knowledge, there are no methods available to specifically isolate microvesicles from urine. This is also supported by the fact that in the differential centrifugation method exosomes are present in low speed pellet. We have developed a lipid affinity-based method to isolate bigger membrane vesicles (mainly 200-750nm). No smaller vesicles were found in the isolated fraction using this method. It is for the first time that such a method has been developed. The fraction isolated using this method was also largely free of high-abundant proteins like albumin and Tamm-Horsfall protein. This method is very easy to perform and requires no sophisticated instrument or specialised training. It can be applied to a large number of samples with ease. It would open up an area of investigation to further characterise microvesicles from urine. Building on this method, similar methods to isolate microvesicles from other body fluid can also be developed in the future.

For further characterization of the membrane vesicles (high-speed pellet), glycoproteome of these vesicles was established using various glycobiological tools and techniques. One hundred and eight glycoproteins were identified unambiguously. The surface glycan profile of these vesicles was characterised and surface membrane glycoproteins were also identified. This study is a first attempt on characterising and identifying glycoprotein constituents of urinary membrane vesicles. This would greatly aid future studies on identifying glyco-biomarkers of renal and cardiovascular diseases. A plate-based FLLA assay was developed which can be expanded in the future to screen patient urine samples. If differences in surface glycosylation are found among healthy and disease samples a diagnostic test can be developed, building on this study, in the future.

Other PTMs, like ubiquitination and palmitoylation, were also profiled and a number of candidate modified proteins were identified. Ubiquitination and palmitoylation play a role in sorting proteins to membrane vesicles. Future studies can build upon this study to identify the

mechanism of sorting of proteins into membrane vesicles as well as validate the PTM status of these proteins. Both these modifications are dynamic and tightly controlled and differences in profile of these PTMs can likely be found among healthy and disease state which could identify new biomarkers or shed light upon molecular mechanism of disease progression.

APPENDIX-A

IDENTIFICATION OF UBIQUITIN-CONJUGATED PROTEINS IN URINARY EXOSOMES AND EXOSOME-LIKE VESICLES

A.1 Introduction

'Ubiquitin-like' proteins (Ubl) comprises of a family of compounds which are structurally and functionally similar to a significant extent (Welchman, Gordon & Mayer, 2005; Kerscher, Felberbaum & Hochstrasser, 2006). The C-terminus of this class of proteins can be ligated to lysine side chain amino group of the substrate proteins by the action of three enzymes namely activating enzyme (E1), conjugating enzyme (E2) and a ligase (E3) (Pickart, 2001; Weissman, 2001). The substrates can be singly or multiply monoubiquitinated implying the addition of a single ubiquitin moiety to one or several lysine at a time. Alternatively, they can be polyubiquitinated implying the addition of a chain of ubiquitin proteins to a single lysine residue in the substrate (Wilkinson, 1995). Unlike other Ubls, ubiquitination might utilise hundreds of E1, E2 and E3 enzymes and sometimes an E4 ubiquitin elongation factor may be needed (Koegl *et al.*, 1999; Hoppe, 2005; Welchman, Gordon & Mayer, 2005; Chiu, Sun & Chen, 2007; Jin *et al.*, 2007). Lys48 linked ubiquitin chains target the substrate proteins for degradation by proteasome system while Lys63 linked chains is implied in DNA damage repair, cellcular signalling, intracellular trafficking and ribosomal biogenesis. All these functions have been reviewed extensively in previous reports (Hochstrasser, 2004; Pickart & Eddins, 2004; Chen, 2005; Welchman, Gordon & Mayer, 2005; Kerscher, Felberbaum & Hochstrasser, 2006; Mukhopadhyay & Riezman, 2007).

Another important function of monoubiquitination has been reported to be in the sorting of proteins at the multivesicular bodies (MVB) into the luminal vesicles which may later be released as exosomes (Katzmann, Odorizzi & Emr, 2002; Raiborg, Rusten & Stenmark, 2003; Babst, 2005). Mono-ubiquitinated membrane proteins are engaged with endosomal sorting machinery with the help of ubiquitin-interaction motifs (UIMs) found e.g. in the hepatocyte-growth-factor-regulated tyrosine-kinase substrate (Hrs or Vps27 in yeast) and in the signal transduction adapter molecule (STAM) (Clague, 2002). Hrs recruits endosomal

sorting complex required for transport-I (ESCRT-I) to the endosomes by interacting with ESCRT-I subunit tumour suppressor gene-101 (TSG101) (Pornillos *et al.*, 2003). This is followed by assembly of ESCRT-II and III complexes on the endosomes and subsequent concentration of cargo of ubiquitinated receptors (Katzmann, Babst & Emr, 2001; Babst *et al.*, 2002b). After the sorting, recruitment of Doa4 by ESCRT-III results in deubiquitination of sorted membrane proteins before they can be incorporated into MVB vesicles (Amerik *et al.*, 2000). Deubiquitination of cargo and sorting complex subunits before they can be incorporated into MVB vesicles implies that ubiquitinated proteins should not be present in exosomes. However, it has been reported that dendritic cell and human B-cell derived exosomes contain ubiquitinated proteins (Buschow *et al.*, 2005). It was also found in the same study that majority of these proteins were not integral membrane proteins. These cytoplasmic proteins were polyubiquinated although it could not be ruled out if monoubiquitinated proteins were present as well. In this paper it was suggested that polyubiquitinated proteins were involved in regulation of MVB protein sorting. However, the identity of these proteins as well as the relevance of their presence in exosomes remains unclear.

To identify the target proteins to be modified by 'ubiquitin-like' modifiers, the main approach has been the expression of tagged version of these modifiers in cell lines and subsequent enrichment of these modifier-conjugated proteins by affinity chromatography (Vertegaal *et al.*, 2006). This is followed by their identification or quantification by mass spectrometric (MS) analysis. However, this approach can not be applied to human body fluids or tissues. Therefore, we have performed multiple immuno-affinity chromatography on extracts of exosomes isolated from human urine. This enrichment of ubiquitin conjugated proteins was followed by identification of these proteins by MS.

A.2 Material and methods

A.2.1 Isolation of exosomes from urine

Membrane vesicle fraction of urine (P200, 000g) was isolated form urine of healthy human volunteers as described in Chapter 2.

A.2.2 Immunoaffinity chromatography (IAC)

Extract of exosomes was prepared by incubating 1mg of exosomal pellet (P200,000g) with 1% (w/v) beta-octyl glucoside (BOG) overnight at 4°C. The extract was diluted 1.5 times to reach .66% (w/v) BOG and 10mM phosphate (Buffer A, pH 7.2). Two different antibodies were used for IAC. One was clone FK-1 (Enzo life sciences, Famingdale, NY) which reocgnises only poly-ubiquitinated proteins while the other was anti-ubiquitin antibody (Dako, Carpentaria, CA). A negative control was taken by performing IAC using an unrelated rabbit IgG antibody. Antibodies were coupled to Dynabeads (M270, amine, Dynal, Inviteogen, Carlsbad, CA) as described for anti-albumin antibodies in methods section 2.2.11 in Chapter 2. Buffer A was used for binding of antibody-beads and proteins in extract. After overnight incubation at 4°C (on rotation) the beads were washed five times with phosphate buffered saline (PBS) and bound proteins were eluted by incubating beads with 100mM glycine (pH 2.3) for 2 hours at room temperature. The eluate was dialysed and the proteins were reduced, alkylated and trypsin digested (as described in Chapter 2) followed by their identification by LC-MS/MS.

A.2.3 SDS-PAGE and western blotting

SDS-PAGE and subsequent transfer of proteins to PVDF membrane was done as described in methods section of Chapter 2. Anti-ubiquitin antibody from Dako was rabbit IgG and it was used in dilutions of 1:1,000 and incubated with blots overnight at +4°C on rotation. The secondary antibody anti-rabbit IgG-IRDye 800 (Li-Cor biosciences) was used in dilutions of

1:5,000 for one hour at room temperature (RT). Anti-polyubiquitin antibody (Fk-1, Enzo life sciences) was an IgM. Therefore, it was biotinylated using Biotin N-hydroxysuccinimide ester (NHS-Biotin, Sigma, St Louis, MO). Briefly, freshly prepared NHS-biotin in dimethylsulfoxide (DMSO) was added to 100µg of antibody at final NHS-biotin concentration of 10mM. The solution was incubated at RT for 2 hours. Hundred mM Tris buffer (Final concentration) was added to the antibody solution to quench remaining NHS-biotin for 30 minutes. The resulting solution was directly incubated with blots at dilution of 1:500 overnight at +4°C on rotation. The next day, blot was developed with streaptavidin-IRDye 800 for 1 hour at RT and images were acquired with Odyssey imaging system (Li-Cor biosciences, Lincoln, NE).

A.2.4 LC-MS/MS identification of bound proteins

LC-MS/MS analysis was carried out as described in Chapter 5.

A.2.5 Bioinformatic analysis and gene ontology

Proteins identified in immuno-precipitations of both the antibodies were analysed by IPA (Ingenuity systems, Redwood City, CA), DAVID bioinformatics resources 6.7 (Huang, Sherman & Lempicki, 2009a; Huang, Sherman & Lempicki, 2009b) (NIH, USA) and Blast2Go software (Conesa *et al.*, 2005). They were classified into various categories based on biological processes they are involved in, their sub-cellular localisation and molecular functions.

A.3 Results

A.3.1 Immuno-affinity chomatography

In a previous report it was shown that exosomes contain mostly poly-ubiquitinated, soluble proteins and very few membrane proteins (Buschow *et al.*, 2005). Moreover, in our analysis of membrane vesicle proteome in Chapter 3, we have found polyubiquitin. Membrane vesicle fraction (P200,000g) as well as P18,000g and SN200,000 were probed with both the antibodies for presence of ubiquitinated proteins (Figure A.1).

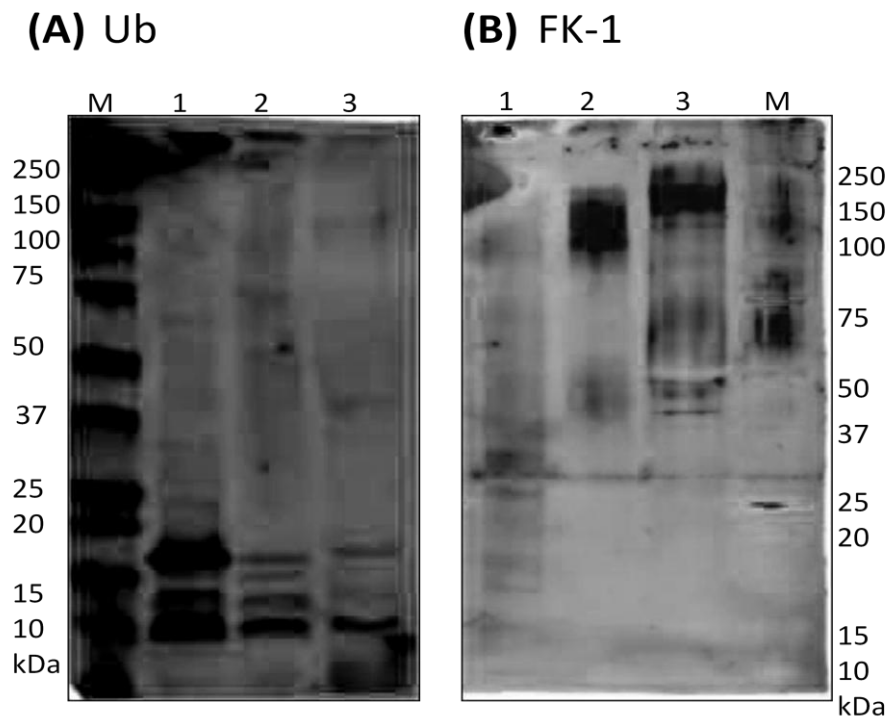

Figure A.1: This figure present western blotting of various fractions of human urine obtained by differential centrifugation method, as described in Chapter 2. Ub is the anti-ubiquitin antibody from Dako and FK-1 antibody is from Enzo Life sciences, which recognises only poly-ubiquitinated proteins. Molecular weight markers (in kDa) are shown in panel A and B at left and right to the picture, respectively. Lane 1: P18,000g, Lane2: P200,000g, Lane3: SN200,000g.

As can be seen in the Figure A.1 panel A, the ubiquitin antibody from Dako recognises multiple proteins (8-10 clearly visible as well as other faint bands) in various fractions. P200,000g presents a band at proximately 10kDa which could be free ubiquitin (Mol wt. approximately 8kDa). Then there are multiple proteins from 12-18kDa which should be ubiquitinated proteins. There are other high molecular weight bands present at very low intensity. In FK-1 blot of P200,000g (Lane2, Panel B) there are multiple bands from 37 to 180kDa. There is a smear of high intensity from approximately 120 to 180kDa. In SN200,000g, there is a big smear from 150 to 180kDa and multiple proteins below this

molecular weight up to 40kDa. There are multiple ubiquitinated proteins present in P200,000g as seen in Figure A.1. Therefore, we have employed immuno-affinity chromatography (on P200,000g) using two different antibodies. One of which is FK-1 which recognises only poly-ubiquitinated proteins. The second antibody (Anti-ubiquitin from Dako) was raised against cow erythrocyte ubiquitin and cross-reacts strongly with human ubiquitin. The resulting antibody has been previously shown to detect presence of ubiquitinated filamentous inclusions in human chronic, neurodegenerative disorders like Alzheimer's disease (Lowe *et al.,* 2001). This antibody most likely recognises both mono- and poly-ubiquitinated proteins as well as free ubiquitin (the band at ~50kDa is common to both antibodies).

A.3.2 Identification of enriched proteins by LC-MS/MS

Excluding keratins, 47 proteins were identified from FK-1 IAC (Supplementary Tables A.1 & A.2) and 47 from Ub IAC (Supplementary Tables A.3 & A.4). Seventy one unique proteins were identified in total while 23 proteins were shared among the elutions of these two antibodies (Figure A.2).

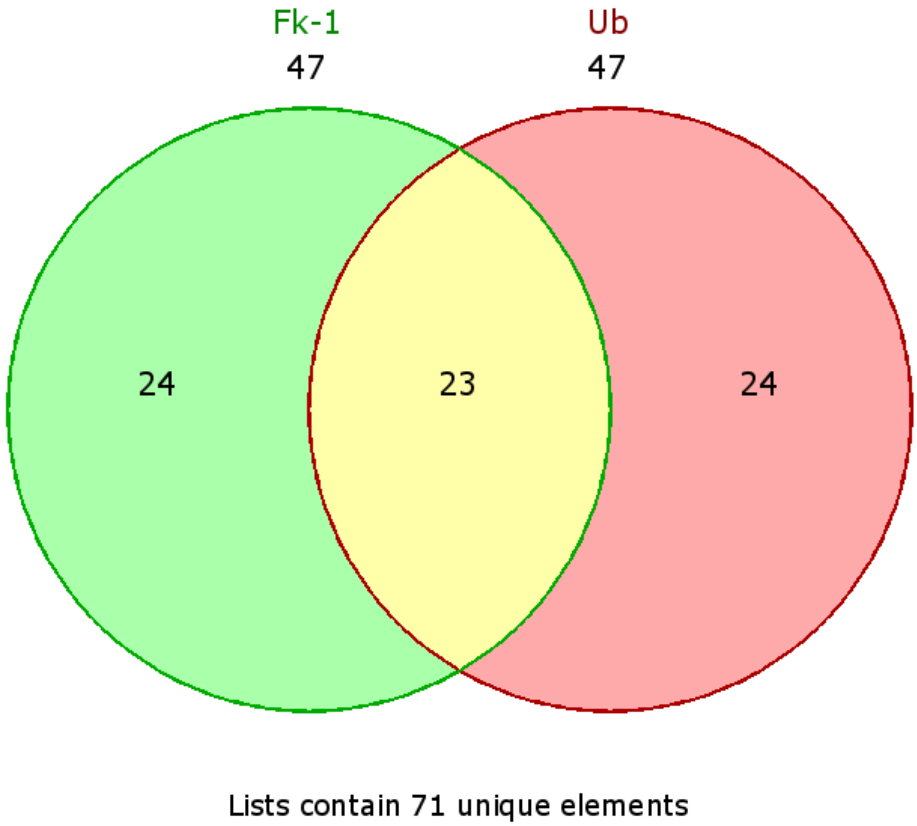

Figure A.2: Comparison between the proteins identified in Fk-1 and Ub immunoprecipitates.

An unrelated antibody (Rabbit IgG) immobilised to the dynabeads and performed the IAC with exosomal extract using the same protocol as the one used for anti-ubiquitin antibodies. As a control, all the proteins identified in this manner serve as our negative control (56 proteins, Supplementary Table A.5). In total 16 proteins from negative control were common with proteins identified in FK-1 and Ub IAC. The remaining 55 proteins are candidate proteins as being ubiquitinated proteins.

A.3.3 Comparison with previous studies

There have been multiple studies identifying the substrate for ubiquitin conjugation although our approach to identify ubiquitin-conjugated proteins in exosomes is the first one. We compare 71 non-redundant proteins identified in our study to all previous studies which have identified ubiquitin substrates. Out of 71 proteins, 39 proteins have been identified as being ubiquitinated in previous studies with many of them identified in multiple studies. These proteins are presented in Table A.1. Out of these 39 proteins, 14 proteins were also found in the negative control. After removing these 14 proteins from consideration, 25 potentially ubiquitinated proteins in urinary exosomes can be found in Table A.1.

Table A.1: Proteins previously known to be ubiquitinated in various studies. Uniprot accession, entrez gene name, sub-cellular localisation, type of proteins, presence of protein in Fk-1 and Ub dako antibody elutions and reference of the previous studies are given in the Table. Rows highlighted in light blue colour are the proteins which were also identified in negative control. Sub-cellular localisation and type of proteins is presented as annotated by IPA (Ingenuity systems).

Uniprot accession	Symbol	Entrez Gene Name	Location	Type (s)	Present in IP	References
A6NDJ8	RAB43L	RAB43 like protein			Ub Dako	(Lopitz-Otsoa *et al.*, 2012)
O43633	CHMP2A	charged multivesicular body protein 2A	Cytoplasm	Other	Ub Dako	(Wagner *et al.*, 2011), (Danielsen *et al.*, 2011)
P01834	IGKC	immunoglobulin kappa constant	Extracellular Space	Other	Both	(Danielsen *et al.*, 2011)
P01857	IGHG1	immunoglobulin heavy constant gamma 1 (G1m marker)	Extracellular Space	Other	Fk-1	hUbiqiotome
P01877	IGHA2	immunoglobulin heavy constant alpha 2 (A2m marker)	Extracellular Space	Other	Fk-1	(Wagner *et al.*, 2011), (Danielsen *et al.*, 2011)
P02788	LTF	lactotransferrin	Extracellular	Peptid	both	(Lopitz-

			Space	ase		Otsoa *et al.*, 2012)
P04279	SEMG1	semenogelin I	Extracellular Space	Other	Ub Dako	(Vanderwerf *et al.*, 2009), (Altun *et al.*, 2011)
P06576	ATP5B	ATP synthase, H+ transporting, mitochondrial F1 complex, beta polypeptide	Cytoplasm	Transp orter	Ub Dako	(Wagner *et al.*, 2011), (Danielsen *et al.*, 2011)
P08236	GUSB	glucuronidase, beta	Cytoplasm	Enzym e	Fk-1	(Wagner *et al.*, 2011)
P14923	JUP	junction plakoglobin	Plasma Membrane	Other	Fk-1	UbiProt, (Lopitz-Otsoa *et al.*, 2012)
P35573	AGL	amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase	Cytoplasm	Enzym e	Fk-1	hUbiqiotome
P53990	IST1	increased sodium tolerance 1 homolog (yeast)	Cytoplasm	other	Ub Dako	(Kim *et al.*, 2011), (Wagner *et al.*, 2011), (Danielsen *et al.*, 2011)
P59190	RAB15	RAB15, member RAS oncogene family	Cytoplasm	Enzym e	Ub Dako	(Lopitz-Otsoa *et al.*, 2012)
P60709	ACTB	actin, beta	Cytoplasm	Other	Fk-1	(Lopitz-Otsoa *et al.*, 2012), (Danielsen *et al.*, 2011)
P61626	LYZ	Lysozyme	Extracellular Space	Enzym e	Both	(Altun *et al.*, 2011), (Danielsen *et al.*, 2011)
Q13510	ASAH1	N-acylsphingosine amidohydrolase (acid ceramidase) 1	Cytoplasm	Enzym e	Fk-1	(Wagner *et al.*, 2011)
Q15678	PTPN14	protein tyrosine phosphatase, non-receptor type 14	Cytoplasm	Phosp hatase	Ub Dako	(Altun *et al.*, 2011)
Q7LBR1	CHMP1B	charged multivesicular body protein 1B	Plasma Membrane	Enzym e	Ub Dako	(Wagner *et al.*, 2011), (Danielsen *et al.*, 2011)
Q92820	GGH	gamma-glutamyl hydrolase (conjugase, folylpolygammaglutamyl hydrolase)	Cytoplasm	Peptid ase	Fk-1	(Wagner *et al.*, 2011), (Danielsen *et al.*, 2011)
Q9BY43	CHMP4A	charged	Cytoplasm	Other	Ub Dako	(Danielsen

		multivesicular body protein 4A				*et al.*, 2011)
Q9H082	RAB33B	RAB33B, member RAS oncogene family	Cytoplasm	Enzyme	Ub Dako	(Lopitz-Otsoa *et al.*, 2012), (Danielsen *et al.*, 2011)
Q9H444	CHMP4B	charged multivesicular body protein 4B	Cytoplasm	Other	Both	(Wagner *et al.*, 2011), (Danielsen *et al.*, 2011)
Q9NRW1	RAB6B	RAB6B, member RAS oncogene family	Cytoplasm	Enzyme	Ub Dako	(Lopitz-Otsoa *et al.*, 2012), (Danielsen *et al.*, 2011)
Q9NZZ3	CHMP5	charged multivesicular body protein 5	Cytoplasm	Other	Both	(Wagner *et al.*, 2011), (Danielsen *et al.*, 2011)
P10909	CLU	Clusterin	Cytoplasm	Chaperon	Fk-1	(Rizzi *et al.*, 2009)

These known ubiquitinated proteins contain multiple families and types of proteins. There are charged multivesicular body proteins (CHMP 1B, 2A, 4A, 4B and 5), RAS oncogene family members (RAB 6B, 15, 33B and RAB43 like), enzymes like acid ceramidase and gamma-glutamyl hydrolase and a tyrosine phosphatase PTPN14. CHMP proteins are part of endosomal sorting complex required for transport-III (ESCRT-III). A map of their interaction was drawn with ConsensusPathDB-human (Kamburov *et al.*, 2011) which is presented in Figure A.3.

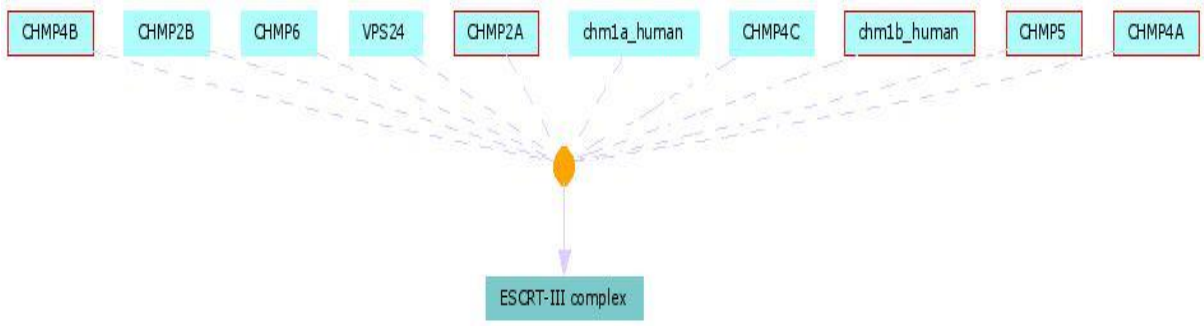

Figure A.3: Components of ESCRT-III complex. Boxes with red outline are the ones found in our study.

These proteins interact physically with each other to form ESCRT-III complex which is required for budding and scission of ILVs into the MVB endosomes (Wollert *et al.*, 2009). Their presence is expected as they are known to be secreted with exosomes although their status in exosomes as potentially ubiquitinated proteins is a new finding.

Proteins identified in our analysis were compared to previous studies published on proteomic analysis of urinary exosomes and exosome-like vesicles (Pisitkun, Shen & Knepper, 2004; Gonzales *et al.*, 2009; Wang *et al.*, 2011) (Knepper, Wang) as well as our dataset from Chpater 3.

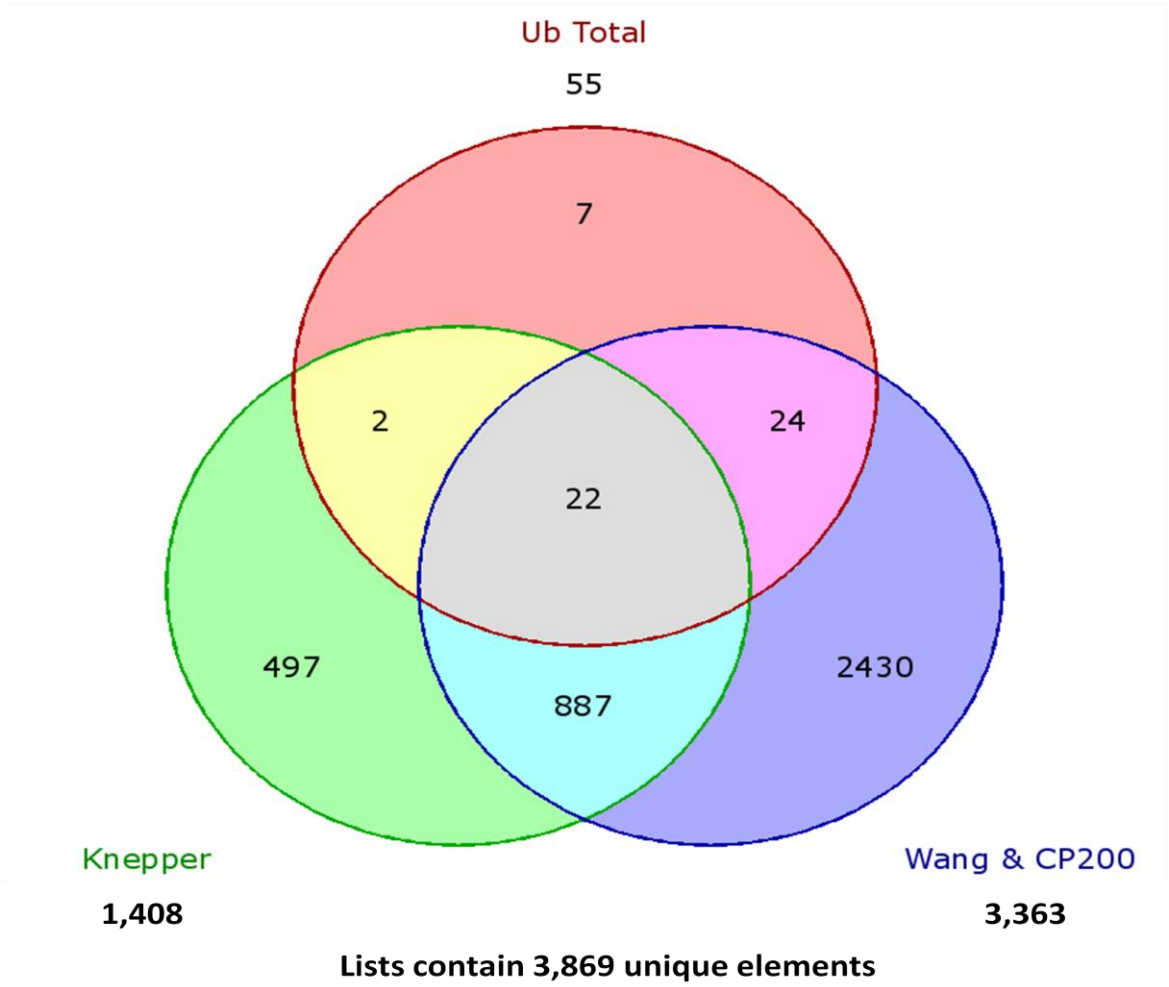

Figure A.4: Comparison of identified proteins with proteomic studies of urinary exosome and 'exosome-like' vesicle studies (Pisitkun, Shen & Knepper, 2004; Gonzales *et al.*, 2009; Wang *et al.*, 2011) (Knepper and Wang) and dataset from Chapter 3.

Seven proteins were found to be unique to our study. These proteins are identified for the first time as being part of urinary membrane vesicles. Two of these proteins are previously known to be associated with exosomes from other sources (Exocarta (Mathivanan *et al.*, 2012)). These proteins are indicated in the table A.2.

Table A.2: Proteins identified in urinary membrane vesicle pellet (P200,000g) for the first time.

Uniprot accesssion	Gene/protein Name	Known ubiquitinated or candidate	Found in Exocarta (Mathivanan *et al.*, 2012)	Mascot score	Protein matches
A6NDJ8	Putative Rab-43-like protein	Known	No	54	1
A6NMY6	Putative annexin A2-like protein	Candidate	Yes	71	2
P01596	Ig kappa chain V-I region CAR	Candidate	No	69	1
P01597	Ig kappa chain V-I region DEE	Candidate	No	45	1
P01614	Ig kappa chain V-II region Cum	Candidate	No	68	1
P01700	Ig lambda chain V-I region HA	Candidate	No	46	1
Q9NRW1	Ras-related protein Rab-6B	Known	Yes	54	1

Out of these 7 proteins presented in table A.2, two proteins putative Rab-43-like protein and RAS related protein Rab-6B are previously known to be ubquitinated (Table A.1). While 5 others are candidate ubiquitinated protein identified only in IAC. Other than 38 known ubiquitinated proteins, we also identified 30 candidate proteins which could be ubiquitinated. In this list of 30 proteins, we have excluded proteins which were also identified in our negative control. These 30 proteins are indicated in supplementary tables at the end of the text.

A.3.4 Bioinformatic analysis and Gene ontology

After removing the proteins also found in negative control, 55 proteins containing unique known and candidate ubiquitinated proteins were submitted to analysis by IPA (Ingnuity systems), DAVID bioinformatics resources (Huang, Sherman & Lempicki, 2009a; Huang, Sherman & Lempicki, 2009b) (NIH, USA) and Blast2Go software (Conesa *et al.*, 2005). These proteins were classified according to the cellular component to which they belong, their molecular functions and biological processes in which they are involved. Figure A.5 presents the sub-cellular localisation of these proteins and the biggest category is extra-

cellular space (44%) followed by cytoplasm (39%) and then plasma membrane (15%). Extracellular space proteins are over-represented in ubiquitin known and candidate identifications while plasma membrane proteins are under-represented.

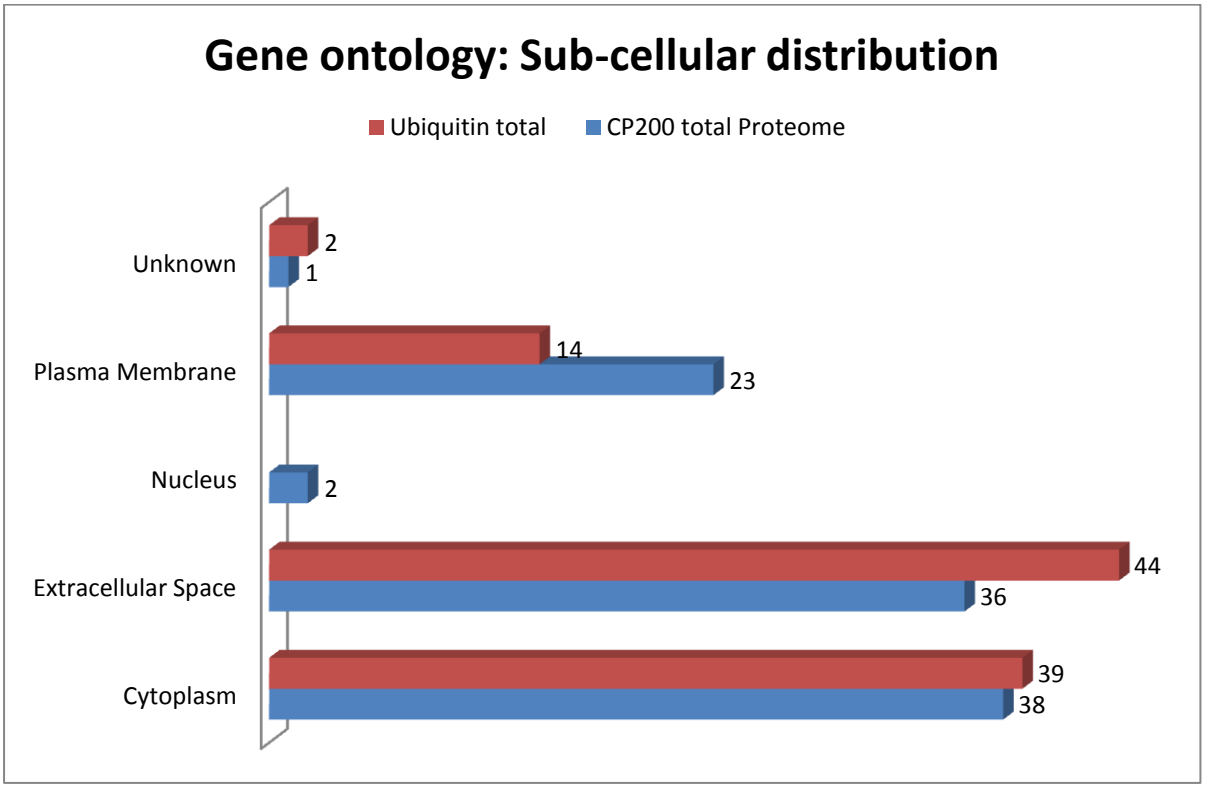

Figure A.5: Sub-cellular localisation of the proteins as annotated by IPA software (Ingenuity systems, USA).

Proteins present in extracellular space would be expected to endocytosed and reach endosome compartment where they could be target of ubiquitination. Cytoplasmic and plasma membrane associated proteins are already known targets of ubiquitination. However, plasma membrane proteins are under-represented compared to whole proteome of urinary membrane vesicles. It suggests that only a fraction of membrane proteins sorted to membrane vesicles are subjected to ubiquitination.

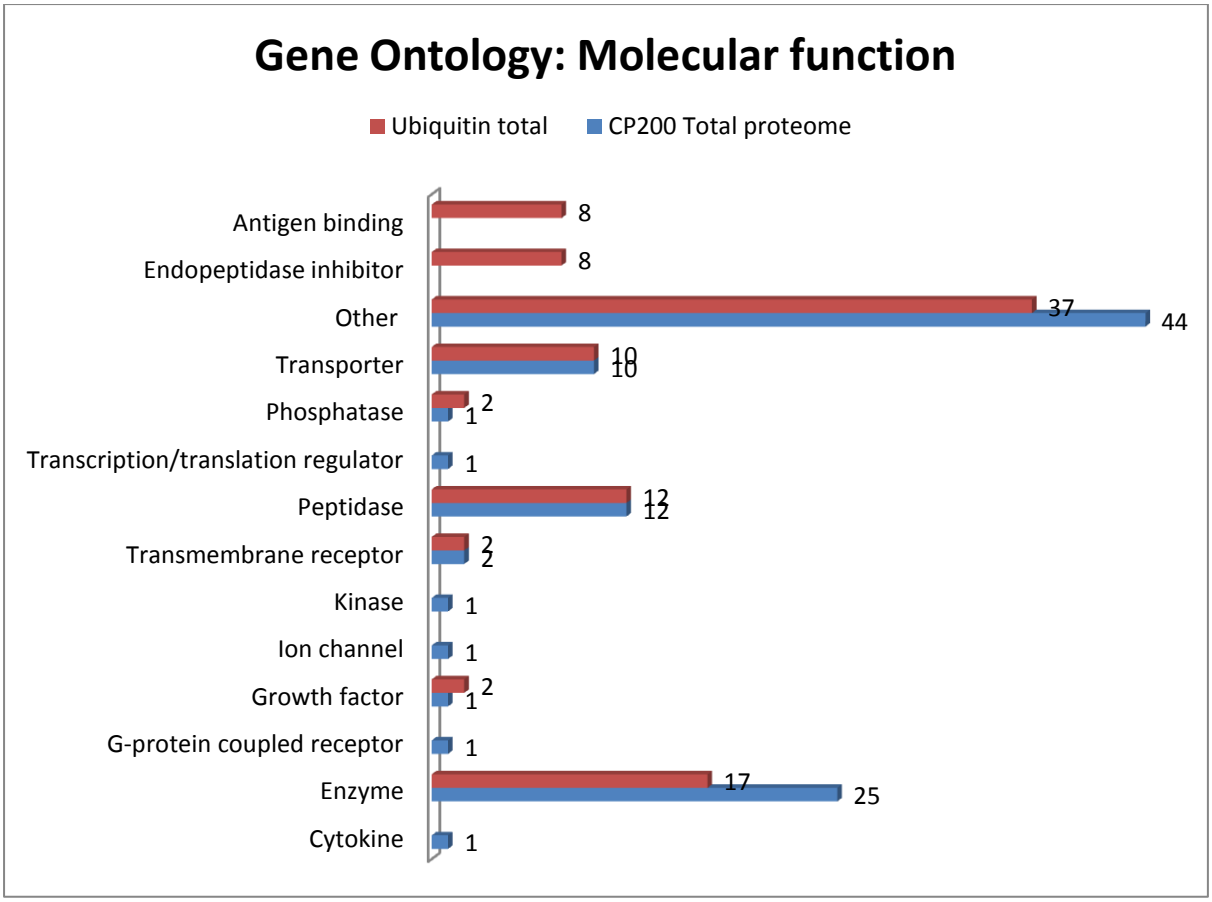

Figure A.6: Molecular functions of the proteins as annotated by IPA software (Ingenuity systems, USA).

When these proteins were classified according to molecular functions (Figure A.6), the biggest annotated category (38% were not annotated) was enzymes (17%) followed by proteases and protease inhibitors (13% and 8%, respectively) and then transporter proteins (10%). Proteins involved in vesicular transport (like RABs) and cargos sorting into ILVs (like CHMP proteins) are already known targets of ubiquitination. Compared to the total proteome (Chapter 3), peptidase and transporter were similarly represented while enzymes were under-represented in ubiquitinated protein identification. It suggests that among the enzymes sorted to membrane vesicles, only a subset is subject to ubiquitination. However, we have to also keep in mind that Chapter 3 is not the most exhaustive of datasets of membrane vesicles and many more proteins make up parts of these vesicles.

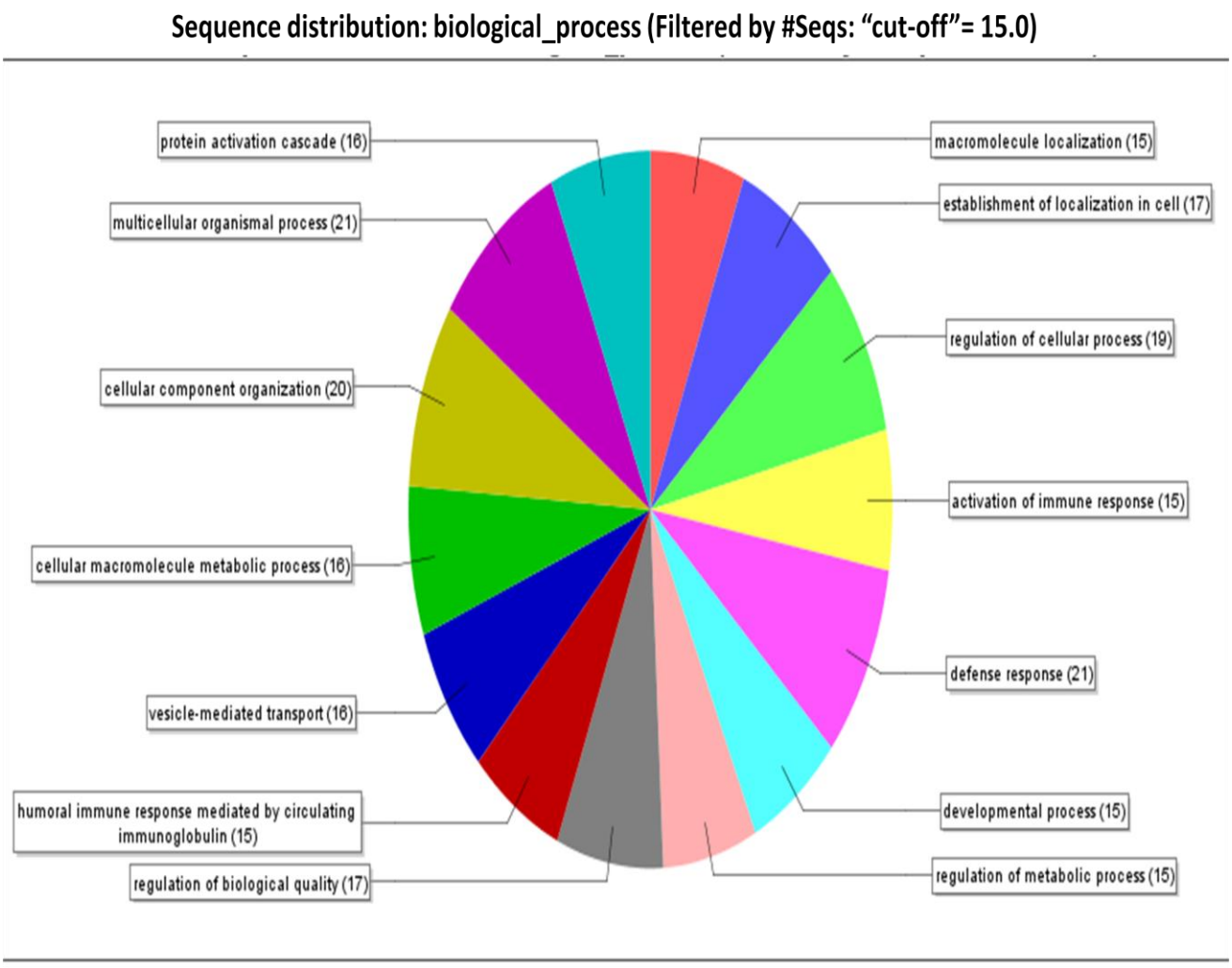

Figure A.7: Presents the annotation of the proteins according to the biological processes in which they are involved. This figure was generated by Blast2Go software and threshold for inclusion of a category put at minimum 15 sequences per category.

Biological processes, in which these proteins are involved, are shown in Figure A.7. Vesicle-mediated transport, cellular component organisation and macromolecule localisation are some of the major categories. Categories also enriched include those for metabolic process, regulation of cellular processes and immune and defense responses.

Proteins identified in our analysis (subtracting the negative control proteins) were subjected to biomarker filter analysis by IPA using cardiovascular and renal diseases, human species,

detection in blood/plasma and urine as various parameters. Biomarkers found in this analysis are presented in table A.3.

Table A.3: Biomarkers found in the proteins present in our identifications upon IPA biomarker filter analysis. Symbols, gene name, sub-cellular localisation, family of proteins, Uniprot Id, presence in plasma/urine, and type of bioarmker applications along with the one representative disease is indicated in the table. Cardiovascular disease and endocrine disorders were top disease in which these proteins have been sued as biomarkers.

Symbol	Entrez Gene Name	Location	Family	GenPept/UniProt/Swiss-Prot Accession	Plasma/Serum	Urine	Biomarker Application (s)	Diseases
APOE	Apolipoprotein E	Extracellular Space	Transporter	P02649	X	X	Diagnosis, Efficacy, Prognosis, Unspecified Application	Cardiovascular disease
CLU	Clusterin	Extracellular Space	Other	P10909	X	X	Diagnosis, Efficacy, Unspecified Application	Cardiovascular disease
EGF (includes EG:13645)	Epidermal growth factor	Extracellular Space	growth factor	P01133	X	X	Diagnosis, Efficacy, Prognosis, Response to Therapy, Unspecified Application	Endocrine system dosorders
F2	Coagulation factor II (thrombin)	Extracellular Space	Peptidase	P00734	X	X	Diagnosis	Cardiovascular disease
JUP	Junction plakoglobin	Plasma Membrane	Other	P14923	X	X	Diagnosis	Cardiovascular disease
KNG1	Kininogen 1	Extracellular Space	Other	P01042	x	X	Efficacy	Cardiovascular disease
LRP2	Low density lipoprotein receptor-	Plasma Membrane	Transporter	P98164	X	X	Diagnosis	Endocrine system

	related protein 2							dosorders
PLG	Plasminogen	Extracellular Space	Peptidase	P00747	X	X	Diagnosis	Cardiovascular disease
RETN	Resistin	Extracellular Space	Other	Q9HD89	X	X	Diagnosis, Disease Progression, Efficacy	Cardiovascular disease
SERPINA5	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 5	Extracellular Space	Other	P05154	X	X	Diagnosis	Cardiovascular disease
SLPI	Secretory leukocyte peptidase inhibitor	Cytoplasm	Other	P03973	X	X	Diagnosis, Prognosis	Cardiovascular disease

Upon further analysis of the protein list with IPA some proteins involved in glomerular injury were also identified which are presented in table A.4.

Table A.4: Presents the molecules involved in glomerular injury. Type of injury, name of molecules found and p-value of their enrichment is indicated in the table.

Category	Functions Annotation	p-Value	Molecules	# Molecules
Glomerular Injury	Glomerulosclerosis	6.07E-05	APOE,C3,CLU,IGHG1	4
Glomerular Injury	lipoprotein glomerulopathy	2.90E-03	APOE	1
Glomerular Injury	focal segmental glomerulosclerosis	4.54E-02	CLU	1

These proteins not only are good targets to verify their ubiquitination status in urinary exosomes but also to investigate if their status changes in animal model systems for glomerular injury. Upon annotation with DAVID, in our 30 new candidate proteins for

ubiquitination we found 3 proteins which are annotated as being part of cytoplasmic vesicles (Table A.5).

Table A.5: Proteins among the candidate ubiquitinated proteins annotated by DAVID as being part of cytoplasmic vesicles.

Serial Number	ID	Gene Name
1	789717	Epidermal growth factor (beta-urogastrone)
2	774646	Low density lipoprotein-related protein 2
3	774379	Polymeric immunoglobulin receptor

These proteins are strong candidates for ubiquitination as they end up in endosome compartments. Megalin, for example, is a multi-ligand receptor which acts with cubulin and is inetrnalised upon ligand binding. We searched for megalin ligands among our identifications and Figure A.8 shows the megalin interaction network that we found within our identifications.

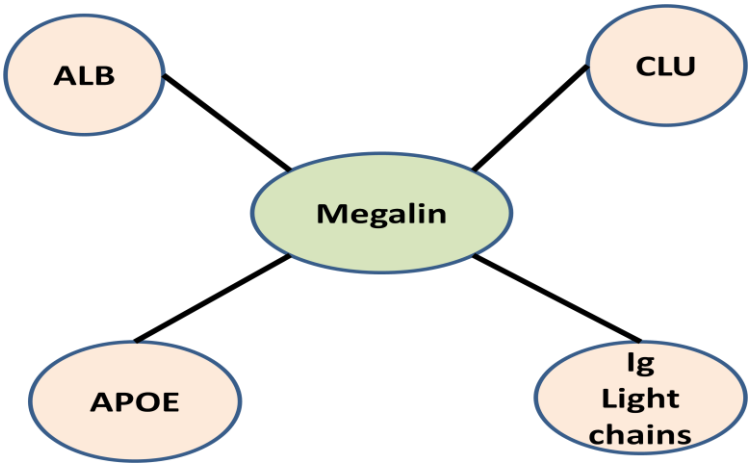

Figure A.8: Interacting proteins of Megalin that were found in our analysis. This figure was manually generated. Data from BioGRID[3.1] was used to search for Megalin interacting partners.

Clusterin, apolipoprotein-E, immunoglobulins (Ig) light chains and albumin are the megalin ligands identified in our study. While albumin and clusterin have been identified as ubiquitinated in a previous study (Lopitz-Otsoa *et al.,* 2012; Rizzi *et al.,* 2009), Ig light chains and apolipoprotein-E are novel candidates found in our study. However albumin was also found in negative control therefore left out of consideration. Immune and inflammation signalling proteins are also targets of ubiquitination (Chen, 2005) therefore we looked for such proteins in our list. We found a category of proteins in our list which were annotated as being involved in inflammatory response (Table A.6).

Table A.6: Proteins found in our identifications which are annotated by DAVID as being involved in inflammatory response.

ID	Gene Name
P02743	amyloid P component, serum
P10909	Clusterin
P00734	coagulation factor II (thrombin)
P01042	kininogen 1
P61626	lysozyme (renal amyloidosis)
O00187	mannan-binding lectin serine peptidase 2
P01024	similar to Complement C3 precursor; complement component 3; hypothetical protein LOC100133511

Analysis of proteins identified in our study, by IPA, identified two pathways enriched among the proteins identified. First one is a pathway which involves ubiquitin targets and ubiquitin specific protease 54 (USP54) binding partners. The pathway is shown in Figure A.9.

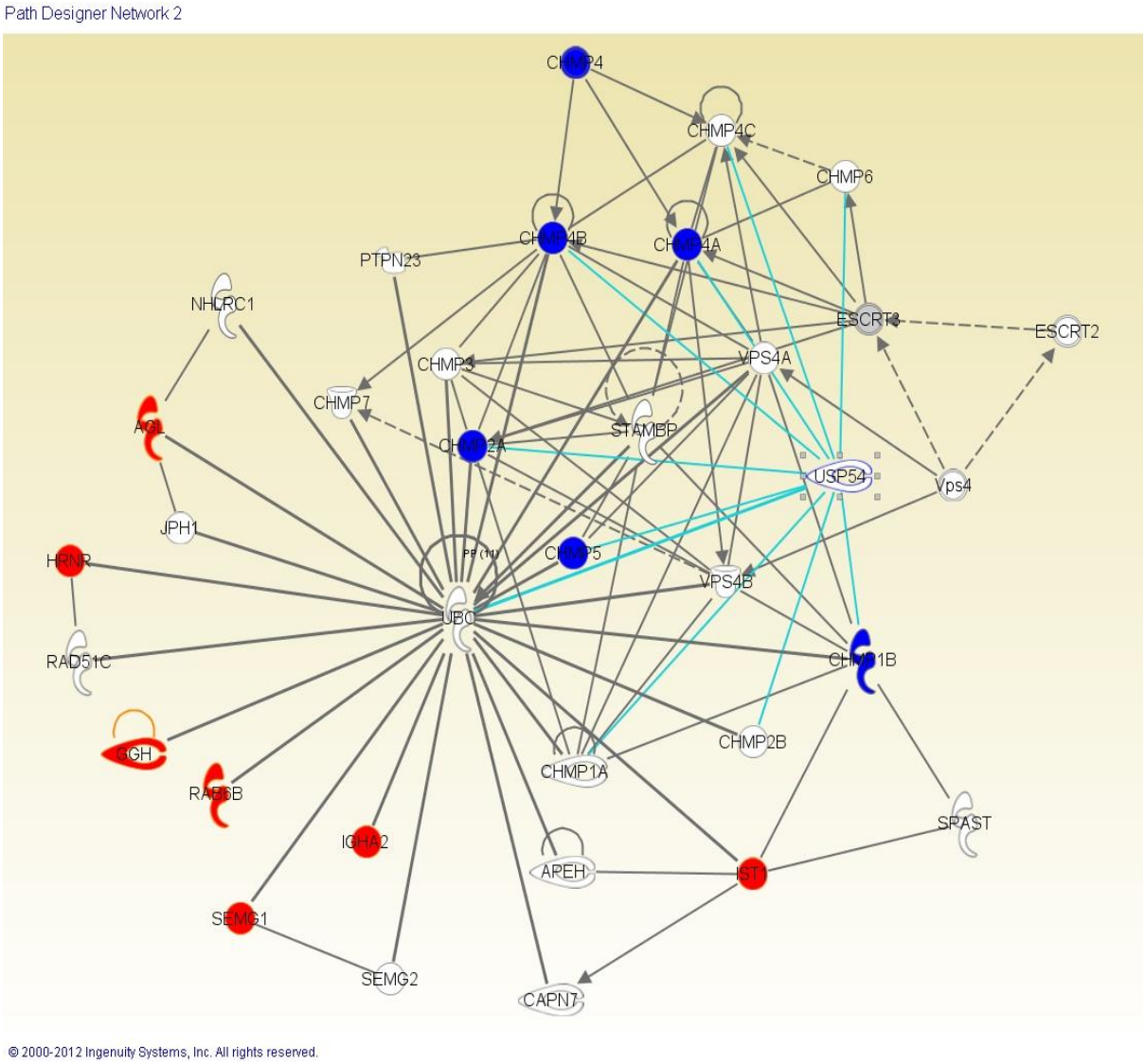

Figure A.9: A pathway found in our data by IPA. The picture was generated using Path Designer tool in the IPA. Proteins shown in red or blue are the ones which are present in our identifications.

This pathway was annotated by IPA as being involved in infectious diseases, cellular assembly and organization as well as DNA recombination, replication and repair. USP54, although inactive as a deubiquitinating protease (Quesada *et al.*, 2004), contains microtubule interacting and transport (MIT) domain which interacts with CHMP proteins (Rigden *et al.*, 2009). This domain is found in proteins involved in intracellular transport and this pathway fits well in context of exosomal secretion machinery.

Another pathway which was enriched in our identification (by IPA analysis) is related to nuclear factor kappa B (NFkB) activation. This pathway is shown in figure A.10.

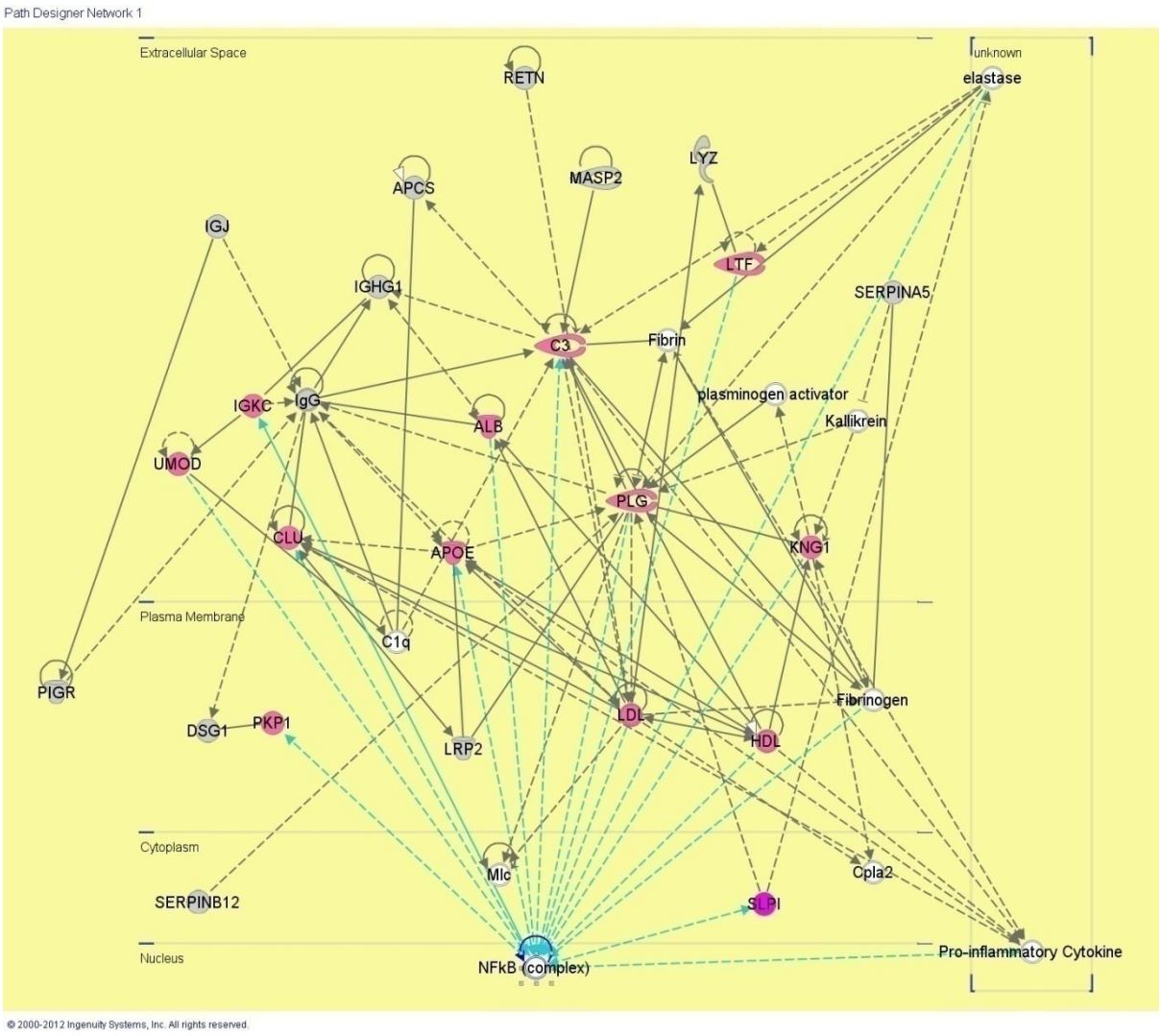

Figure A.10: A pathway found in our data by IPA. The picture was generated using Path Designer tool in the IPA. The proteins in pink or grey circles are the ones present in our identifications. Pink circles are the proteins which directly influence the nuclear localisation of NFkB.

The proteins shown in Figure A.10 are mostly either extracellular or part of plasma membrane and increase or decrease the nuclear localisation of NFkB. Previously it is known that NFkB activation in B-cells regulate HLA secretion with exosomes (Arita *et al.*, 2008).

Another study has also shown that exosome treatment of cells activate the NFkB pathway and lead to inflammation (Anand *et al.,* 2010).

A.4 Discussion

IAC was carried out on urinary exosomal extract and proteins were identified by LC-MS/MS. A number of proteins get enriched in the immunoprecipitations not only because of their specific affinity but also due to non-specific interaction with matrix stationary phase, protein-protein interactions and protein aggregation. To avoid the consideration of such proteins as ubiquitinated in our study an unrelated rabbit IgG was immobilised to the Dynabeads and IAC was carried out with same protocol as the anti-ubiquitin (Ub) antibody and all the proteins found in this way were considered as a negative control. We identified 71 proteins in total using two different antibodies, clone FK-1 which recognises only poly-ubiquitinated proteins and anti-ubiquitin (Dako) which is not so well characterised but may recognise both mono- and poly-ubiquitinated proteins. Thirty-nine previously known ubiquitinated proteins were identified but 14 of these proteins were present in negative control as well so they were not considered as potential Ub substrate. We prepared a membrane vesicle extract by incubating the exosomal pellet (P200,000g) with 1% (w/v) beta-octyl glucoside (BOG) overnight at 4°C. There are two reasons for that; one being that exosomes are enriched in 'lipid raft-like' domains and BOG can solubilise these rafts (Garner, Smith & Hooper, 2008) and second is that BOG has previously been shown to increase the specificity of immuno-precipitation (IP) (Zhang & Neubert, 2006). It was shown in the study that BOG, when present above its critical micelle concentration (.7%) in the IP buffer, increases both the sensitivity and selectivity of immuno-capture. This objective was achieved as 55% of the total proteins identified in our analysis are previously known to be ubiquitinated.

Among the proteins identified one interesting category was CHMP family proteins. We identified 5 of the 10 CHMP proteins known to constitute the ESCRT-III complex. This complex is required for ILV formation and scission (Wollert *et al.*, 2009). ESCRT-III interacts with the endosomal membrane (Whitley *et al.*, 2003) and recruits cargo to be

included in MVB pathway (Babst *et al.,* 2002a). ESCRT-III recruits Doa4 which deubiquitinates the cargo before it is included in MVB ILVs (Babst *et al.,* 2002a). However, this sequence of events is incomplete as according to this scenarios exosomes, which are derived from MVB later in the pathway, should be free of ubiquitinated proteins. But ubiquitinated proteins are present in exosomes as shown previously (Buschow *et al.,* 2005) and by our study. This means that ubiquitination plays as yet unknown role in MVB cargo sorting. The presence of ESCRT-III complex proteins in ubiquitinated form in exosomes also raises question about recycling of this complex as well as maintaining the free ubiquitin pool of the cells. Further studies will be required to answer these aspects of MVB sorting. It has been suggested that Ub dependent sorting of cell surface receptors into MVB might regulate their stability as well as turnover of misfolded proteins (Babst *et al.,* 2002a). Clusterin which was identified in our study is known to bind to unfolded proteins and chaperon-client complex is internalised and targeted for lysosomal degradation (Kounnas *et al.,* 1995). It is possible that subsequent to the ubiquitination of inetrnalised proteins a decision is taken to either degrade it in lysosome or secrete ILVs as exosome by fusing MVB with plasma membrane.

Endocytosis of receptor ligand complexes leads them into endosomes and ubiquitination of endocytic cargo is a constitutive process (Katzmann, Odorizzi & Emr, 2002). In our candidate ubiquitinated proteins list at least 3 proteins EGF, megalin and polymeric immunoglobulin receptor were found which end up in endosomes upon endocytosis. These receptors need to be either recycled or degraded in the cell and Ub plays an important role in these two decisions. Decision of degradation further leads to another choice in which either MVB fuse with lysosomes or with plasma membrane releasing these proteins as part of exosomes. Therefore these 3 proteins are very strong candidate for ubiquitination and further studies will be needed to confirm this. Moreover, surface expression and ectodomain

shedding of EGF family member pro-amphiregulin is already known to be tightly controlled by ubiquitination (Fukuda *et al.,* 2012). There is no reason to believe that this can not be the case for EGF. Megalin is involved in internalisation of Ig light chains by proximal tubule (Li *et al.,* 2008) and can also lead to light chain induced nephrotoxicity in pathological conditions (Basnayake *et al.,* 2010). Ig light chains and megalin, both were also found in our study. IgA, IgM and their receptor, polymeric immunoglobulin receptor, all were identified. However, it remains to be established whether these interacting proteins were present as complexes and were enriched indirectly with only one or few of them being ubiquitinated or all of them are secreted as part of exosomes in ubiquitinated form.

NFkB is implicated in inductions and resolution of inflammations upon stimulation by a variety of molecules and ligands (Belen Sanz *et al.,* 2010). Both the classical and alternative pathways have been shown to be activated in kidney diseases and injury (Belen Sanz *et al.,* 2010). Moreover, NFkB activation is known to be upregulated in experimental renal disease (Guijarro & Egido, 2001). We have found a number of proteins in our identifications which affect the nuclear localisation of NFkB including Ig light chains as discussed previously. Exosomal localisation of NFkB effectors in ubiquitinated form raises questions whether exosomes serve as a dumping ground for this class of proteins? And whether their secretion with exosomes serves as a homeostatic mechanism which controls cellular levels of NFkB effectors? Further studies will be needed to establish the connection of exosomal proteins with NFkB pathway.

Another family of proteins known to be ubiquitinated and present in our analysis is RAB, RAS oncogene family members. RAB6B and RAB33B are involved in retrograde transport of cargo from early endosomes to trans-Golgi network (TGN) (Starr *et al.,* 2010) while RAB15 is involved in recycling of internalised receptors from recycling endosomes (Strick & Elferink, 2005). Receptors which are recycled might include polymeric immunoglobulin

receptor and megalin, as previously discussed. A picture emerges in which distinct steps in vesicle mediated transport from golgi to endosomes and vice versa, endosome to lysosome and endosome to plasma membrane is tightly regulated by ubiquitination of not only effector transport proteins but also of the cargo itself. While the exact reasons why only some of the ubiquitinated proteins and transporters end up in exosomes and not other components of the system remains to be answered in detail, this study provides the target molecules which can be studied further.

A.5 Reference

Altun, M., Kramer, H.B., Willems, L.I., McDermott, J.L., Leach, C.A., Goldenberg, S.J., Kumar, K.G.S., Konietzny, R., Fischer, R., Kogan, E., Mackeen, M.M., McGouran, J., Khoronenkova, S.V., Parsons, J.L., Dianov, G.L., Nicholson, B. & Kessler, B.M. (2011), "Activity-Based Chemical Proteomics Accelerates Inhibitor Development for Deubiquitylating Enzymes", *CHEMISTRY \& BIOLOGY*, vol. 18, no. 11, pp. 1401-1412.

Amerik, A., Nowak, J., Swaminathan, S. & Hochstrasser, M. (2000), "The Doa4 deubiquitinating enzyme is functionally linked to the vacuolar protein-sorting and endocytic pathways", *MOLECULAR BIOLOGY OF THE CELL*, vol. 11, no. 10, pp. 3365-3380.

Anand, P.K., Anand, E., Bleck, C.K.E., Anes, E. & Griffiths, G. (2010), "Exosomal Hsp70 Induces a Pro-Inflammatory Response to Foreign Particles Including Mycobacteria", *PLOS ONE*, vol. 5, no. 4.

Arita, S., Baba, E., Shibata, Y., Niiro, H., Shimoda, S., Isobe, T., Kusaba, H., Nakano, S. & Harada, M. (2008), "B cell activation regulates exosomal HLA production", *EUROPEAN JOURNAL OF IMMUNOLOGY*, vol. 38, no. 5, pp. 1423-1434.

Babst, M. (2005), "A protein's final ESCRT", *TRAFFIC*, vol. 6, no. 1, pp. 2-9.

Babst, M., Katzmann, D., Estepa-Sabal, E., Meerloo, T. & Emr, S. (2002 a), "ESCRT-III: An endosome-associated heterooligomeric protein complex required for MVB sorting", *DEVELOPMENTAL CELL*, vol. 3, no. 2, pp. 271-282.

Babst, M., Katzmann, D., Snyder, W., Wendland, B. & Emr, S. (2002 b), "Endosome-associated complex, ESCRT-II, recruits transport machinery for protein sorting at the multivesicular body", *DEVELOPMENTAL CELL*, vol. 3, no. 2, pp. 283-289.

Basnayake, K., Ying, W., Wang, P. & Sanders, P.W. (2010), "Immunoglobulin light chains activate tubular epithelial cells through redox signaling", *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY*, vol. 21, no. 7, pp. 1165-1173.

Belen Sanz, A., Dolores Sanchez-Nino, M., Mario Ramos, A., Antonio Moreno, J., Santamaria, B., Ruiz-Ortega, M., Egido, J. & Ortiz, A. (2010), "NF-kappa B in renal inflammation", *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY*, vol. 21, no. 8, pp. 1254-1262.

Buschow, S., Liefhebber, J., Wubbolts, R. & Stoorvogel, W. (2005), "Exosomes contain ubiquitinated proteins", *BLOOD CELLS MOLECULES AND DISEASES*, vol. 35, no. 3, pp. 398-403.

Chen, Z.J. (2005), "Ubiquitin signalling in the NF-kappa B pathway", *NATURE CELL BIOLOGY*, vol. 7, no. 8, pp. 758-765.

Chiu, Y., Sun, Q. & Chen, Z.J. (2007), "E1-L2 activates both ubiquitin and FAT10", *MOLECULAR CELL*, vol. 27, no. 6, pp. 1014-1023.

Clague, M. (2002), "Membrane transport: A coat for ubiquitin", *CURRENT BIOLOGY*, vol. 12, no. 15, pp. R529-R531.

Conesa, A., Gotz, S., Garcia-Gomez, J., Terol, J., Talon, M. & Robles, M. (2005), "Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research", *BIOINFORMATICS*, vol. 21, no. 18, pp. 3674-3676.

Danielsen, J.M.R., Sylvestersen, K.B., Bekker-Jensen, S., Szklarczyk, D., Poulsen, J.W., Horn, H., Jensen, L.J., Mailand, N. & Nielsen, M.L. (2011), "Mass Spectrometric Analysis of Lysine Ubiquitylation Reveals Promiscuity at Site Level", *MOLECULAR \& CELLULAR PROTEOMICS*, vol. 10, no. 3. pp. M110.

Fukuda, S., Nishida-Fukuda, H., Nakayama, H., Inoue, H. & Higashiyama, S. (2012), "Monoubiquitination of pro-amphiregulin regulates its endocytosis and ectodomain shedding", *BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS*, vol. 420, no. 2, pp. 315-320.

Garner, A.E., Smith, D.A. & Hooper, N.M. (2008), "Visualization of detergent solubilization of membranes: Implications for the isolation of rafts", *BIOPHYSICAL JOURNAL*, vol. 94, no. 4, pp. 1326-1340.

Gonzales, P.A., Pisitkun, T., Hoffert, J.D., Tchapyjnikov, D., Star, R.A., Kleta, R., Wang, N.S. & Knepper, M.A. (2009), "Large-Scale Proteomics and Phosphoproteomics of Urinary Exosomes", *JOURNAL OF THE AMERICAN SOCIETY OF NEPHROLOGY*, vol. 20, no. 2, pp. 363-379.

Guijarro, C. & Egido, J. (2001), "Transcription factor-kappa B (NF-kappa B) and renal disease", *KIDNEY INTERNATIONAL*, vol. 59, no. 2, pp. 415-424.

Hochstrasser, M. (2004), "Ubiquitin signalling: what's in a chain?", *NATURE CELL BIOLOGY*, vol. 6, no. 7, pp. 571-572.

Hoppe, T. (2005), "Multiubiquitylation by E4 enzymes: `one size' doesn't fit all", *TRENDS IN BIOCHEMICAL SCIENCES*, vol. 30, no. 4, pp. 183-187.

Huang, D.W., Sherman, B.T. & Lempicki, R.A. (2009 a), "Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists", *NUCLEIC ACIDS RESEARCH*, vol. 37, no. 1, pp. 1-13.

Huang, D.W., Sherman, B.T. & Lempicki, R.A. (2009 b), "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources", *NATURE PROTOCOLS*, vol. 4, no. 1, pp. 44-57.

Jin, J., Li, X., Gygi, S.P. & Harper, J.W. (2007), "Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging", *NATURE*, vol. 447, no. 7148, pp. 1135-U17.

Kamburov, A., Pentchev, K., Galicka, H., Wierling, C., Lehrach, H. & Herwig, R. (2011), "ConsensusPathDB: toward a more complete picture of cell biology", *NUCLEIC ACIDS RESEARCH*, vol. 39, no. 1, pp. D712-D717.

Katzmann, D., Babst, M. & Emr, S. (2001), "Ubiquitin-dependent sorting into the multivesicular body pathway requires the function of a conserved endosomal protein sorting complex, ESCRT-I", *CELL*, vol. 106, no. 2, pp. 145-155.

Katzmann, D., Odorizzi, G. & Emr, S. (2002), "Receptor downregulation and multivesicular-body sorting", *NATURE REVIEWS MOLECULAR CELL BIOLOGY*, vol. 3, no. 12, pp. 893-905.

Kerscher, O., Felberbaum, R. & Hochstrasser, M. (2006), "Modification of proteins by ubiquitin and ubiquitin-like proteins", *ANNUAL REVIEW OF CELL AND DEVELOPMENTAL BIOLOGY*, vol. 22, pp. 159-180.

Kim, W., Bennett, E.J., Huttlin, E.L., Guo, A., Li, J., Possemato, A., Sowa, M.E., Rad, R., Rush, J., Comb, M.J., Harper, J.W. & Gygi, S.P. (2011), "Systematic and Quantitative Assessment of the Ubiquitin-Modified Proteome", *MOLECULAR CELL*, vol. 44, no. 2, pp. 325-340.

Koegl, M., Hoppe, T., Schlenker, S., Ulrich, H., Mayer, T. & Jentsch, S. (1999), "A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly", *CELL*, vol. 96, no. 5, pp. 635-644.

Kounnas, M., Loukinova, E., Stefansson, S., Harmony, J., Brewer, B., Strickland, D. & Argraves, W. (1995), "Identification of glycoprotein-330 as an endocytic receptor for apolipoprotein-J/clusterin", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 270, no. 22, pp. 13070-13075.

Li, M., Balamuthusamy, S., Simon, E.E. & Batuman, V. (2008), "Silencing megalin and cubilin genes inhibits myeloma light chain endocytosis and ameliorates toxicity in human renal proximal tubule epithelial cells", *AMERICAN JOURNAL OF PHYSIOLOGY-RENAL PHYSIOLOGY*, vol. 295, no. 1, pp. F82-F90.

Lopitz-Otsoa, F., Rodriguez-Suarez, E., Aillet, F., Casado-Vela, J., Lang, V., Matthiesen, R., Elortza, F. & Rodriguez, M.S. (2012), "Integrative analysis of the ubiquitin proteome isolated using Tandem Ubiquitin Binding Entities (TUBEs)", *JOURNAL OF PROTEOMICS*, vol. 75, no. 10, pp. 2998-3014.

Lowe, J., Mayer, J., Landon, M. & Layfield, R. (2001), "Ubiquitin and the molecular pathology of neurodegenerative diseases", *NEUROPATHOLOGY AND GENETICS OF DEMENTIA*, eds. M. Tolnay & A. Probst, Swiss Soc Neuropathol, , pp. 169.

Mathivanan, S., Fahner, C.J., Reid, G.E. & Simpson, R.J. (2012), "ExoCarta 2012: database of exosomal proteins, RNA and lipids", *NUCLEIC ACIDS RESEARCH*, vol. 40, no. D1, pp. D1241-D1244.

Mukhopadhyay, D. & Riezman, H. (2007), "Proteasome-independent functions of ubiquitin in endocytosis and signaling", *SCIENCE*, vol. 315, no. 5809, pp. 201-205.

Pickart, C. (2001), "Mechanisms underlying ubiquitination", *ANNUAL REVIEW OF BIOCHEMISTRY*, vol. 70, pp. 503-533.

Pickart, C. & Eddins, M. (2004), "Ubiquitin: structures, functions, mechanisms", *BIOCHIMICA ET BIOPHYSICA ACTA-MOLECULAR CELL RESEARCH*, vol. 1695, no. 1-3, pp. 55-72.

Pisitkun, T., Shen, R. & Knepper, M. (2004), "Identification and proteomic profiling of exosomes in human urine", *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, vol. 101, no. 36, pp. 13368-13373.

Pornillos, O., Higginson, D., Stray, K., Fisher, R., Garrus, J., Payne, M., He, G., Wang, H., Morham, S. & Sundquist, W. (2003), "HIV Gag mimics the Tsg101-recruiting activity of the human Hrs protein", *JOURNAL OF CELL BIOLOGY*, vol. 162, no. 3, pp. 425-434.

Quesada, V., Diaz-Perales, A., Gutierrez-Fernandez, A., Garabaya, C., Cal, S. & Lopez-Otin, C. (2004), "Cloning and enzymatic analysis of 22 novel human ubiquitin-specific proteases",

BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 314, no. 1, pp. 54-62.

Raiborg, C., Rusten, T. & Stenmark, H. (2003), "Protein sorting into multivesicular endosomes", *CURRENT OPINION IN CELL BIOLOGY*, vol. 15, no. 4, pp. 446-455.

Rigden, D.J., Liu, H., Hayes, S.D., Urbe, S. & Clague, M.J. (2009), "Ab initio protein modelling reveals novel human MIT domains", *FEBS LETTERS*, vol. 583, no. 5, pp. 872-878.

Rizzi, F., Caccamo, A.E., Belloni, L. & Bettuzzi, S. (2009), "Clusterin Is a Short Half-Life, Poly-Ubiquitinated Protein, Which Controls the Fate of Prostate Cancer Cells", *JOURNAL OF CELLULAR PHYSIOLOGY*, vol. 219, no. 2, pp. 314-323.

Starr, T., Sun, Y., Wilkins, N. & Storrie, B. (2010), "Rab33b and Rab6 are Functionally Overlapping Regulators of Golgi Homeostasis and Trafficking", *TRAFFIC*, vol. 11, no. 5, pp. 626-636.

Strick, D. & Elferink, L. (2005), "Rab15 effector protein: A novel protein for receptor recycling from the endocytic recycling compartment", *MOLECULAR BIOLOGY OF THE CELL*, vol. 16, no. 12, pp. 5699-5709.

Vanderwerf, S.M., Svahn, J., Olson, S., Rathbun, R.K., Harrington, C., Yates, J., Keeble, W., Anderson, D.C., Anur, P., Pereira, N.F., Pilonetto, D.V., Pasquini, R. & Bagby, G.C. (2009), "TLR8-dependent TNF-alpha overexpression in Fanconi anemia group C cells", *BLOOD*, vol. 114, no. 26, pp. 5290-5298.

Vertegaal, A.C.O., Andersen, J.S., Ogg, S.C., Hay, R.T., Mann, M. & Lamond, A.I. (2006), "Distinct and overlapping sets of SUMO-1 and SUMO-2 target proteins revealed by quantitative proteomics", *MOLECULAR \& CELLULAR PROTEOMICS*, vol. 5, no. 12, pp. 2298-2310.

Wagner, S.A., Beli, P., Weinert, B.T., Nielsen, M.L., Cox, J., Mann, M. & Choudhary, C. (2011), "A proteome-wide, quantitative survey of *in vivo* ubiquitylation sites reveals widespread regulatory roles", *MOLECULAR \& CELLULAR PROTEOMICS*, vol. 10, no. 10.

Wang, Z., Hill, S., Luther, J.M., Hachey, D.L. & Schey, K.L. (2011), "Proteomic analysis of urine exosomes by multidimensional protein identification technology (MudPIT)", *PROTEOMICS*, vol. 12, no. 2, pp. 329-338.

Weissman, A. (2001), "Themes and variations on ubiquitylation", *NATURE REVIEWS MOLECULAR CELL BIOLOGY*, vol. 2, no. 3, pp. 169-178.

Welchman, R., Gordon, C. & Mayer, R. (2005), "Ubiquitin and 'ubiquitin-like' proteins as multifunctional signals", *NATURE REVIEWS MOLECULAR CELL BIOLOGY*, vol. 6, no. 8, pp. 599-609.

Whitley, P., Reaves, B., Hashimoto, M., Riley, A., Potter, B. & Holman, G. (2003), "Identification of mammalian Vps24p as an effector of phosphatidylinositol 3,5-bisphosphate-dependent endosome compartmentalization", *JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 278, no. 40, pp. 38786-38795.

Wilkinson, K. (1995), "Roles of ubiquitinylation in proteolysis and cellular-regulation", *ANNUAL REVIEW OF NUTRITION*, vol. 15, pp. 161-189.

Wollert, T., Wunder, C., Lippincott-Schwartz, J. & Hurley, J.H. (2009), "Membrane scission by the ESCRT-III complex", *NATURE*, vol. 458, no. 7235, pp. 172-U2.

Zhang, G. & Neubert, T. (2006), "Use of detergents to increase selectivity of immunoprecipitation of tyrosine phosphorylated peptides prior to identification by MALDI quadrupole-TOF MS", *PROTEOMICS*, vol. 6, no. 2, pp. 571-578.

Supplementary Table A.1: Proteins identified in Fk-1 IAC sample 1. Uniprot accession, protein description, score, molecular weight (in Dalton) and number of protein matches are indicated in the table.

Accession	Description	Score	Mol Wt.	prot_matches
Q13510	Acid ceramidase OS=Homo sapiens GN=ASAH1 PE=1 SV=5	60	45095	1
P05090	Apolipoprotein D OS=Homo sapiens GN=APOD PE=1 SV=1	78	21552	3
P02649	Apolipoprotein E OS=Homo sapiens GN=APOE PE=1 SV=1	50	36248	1
P08236	Beta-glucuronidase OS=Homo sapiens GN=GUSB PE=1 SV=2	41	75033	1
P31944	Caspase-14 OS=Homo sapiens GN=CASP14 PE=1 SV=2	82	27952	2
Q9NZZ3	Charged multivesicular body protein 5 OS=Homo sapiens GN=CHMP5 PE=1 SV=1	42	24613	1
P10909	Clusterin OS=Homo sapiens GN=CLU PE=1 SV=1	45	53041	1
P01024	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2	57	188596	1
P01040	Cystatin-A OS=Homo sapiens GN=CSTA PE=1 SV=1	41	11000	1
P81605	Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2	108	11393	4
Q02413	Desmoglein-1 OS=Homo sapiens GN=DSG1 PE=1 SV=2	138	114720	5
P15924	Desmoplakin OS=Homo sapiens GN=DSP PE=1 SV=3	258	334063	9
Q92820	Gamma-glutamyl hydrolase OS=Homo sapiens GN=GGH PE=1 SV=2	46	36347	1
P35573	Glycogen debranching enzyme OS=Homo sapiens GN=AGL PE=1 SV=3	46	176856	1
Q86YZ3	Hornerin OS=Homo sapiens GN=HRNR PE=1 SV=2	133	283156	3
P01877	Ig alpha-2 chain C region OS=Homo sapiens GN=IGHA2 PE=1 SV=3	128	37315	4
P01857	Ig gamma-1 chain C region OS=Homo sapiens GN=IGHG1 PE=1 SV=1	132	36605	4
P01766	Ig heavy chain V-III region BRO OS=Homo sapiens PE=1 SV=1	41	13334	1
P01762	Ig heavy chain V-III region TRO OS=Homo sapiens PE=1 SV=1	41	13580	1
P01764	Ig heavy chain V-III region VH26 OS=Homo sapiens PE=1 SV=1	40	12748	1
P01834	Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 SV=1	139	11776	3
P01596	Ig kappa chain V-I region CAR OS=Homo sapiens PE=1 SV=1	69	11812	1
P01597	Ig kappa chain V-I region DEE OS=Homo sapiens PE=1 SV=1	148	11770	2
P01614	Ig kappa chain V-II region Cum OS=Homo sapiens PE=1 SV=1	68	12784	1
P01616	Ig kappa chain V-II region MIL OS=Homo sapiens PE=1 SV=1	70	12164	1
P01620	Ig kappa chain V-III region SIE OS=Homo sapiens PE=1 SV=1	71	11884	2
P01700	Ig lambda chain V-I region HA OS=Homo sapiens PE=1 SV=1	46	12005	1
P0CG05	Ig lambda-2 chain C regions OS=Homo sapiens GN=IGLC2 PE=1 SV=1	51	11461	1
P01871	Ig mu chain C region OS=Homo sapiens GN=IGHM PE=1 SV=3	41	49972	1
P01591	Immunoglobulin J chain OS=Homo sapiens GN=IGJ PE=1 SV=4	62	18551	1
P14923	Junction plakoglobin OS=Homo sapiens GN=JUP PE=1 SV=3	84	82447	2
Q14525	Keratin, type I cuticular Ha3-II OS=Homo sapiens GN=KRT33B PE=2 SV=3	87	47345	3
P13645	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6	1187	59024	31
P13646	Keratin, type I cytoskeletal 13 OS=Homo sapiens GN=KRT13 PE=1 SV=4	191	49905	7
P02533	Keratin, type I cytoskeletal 14 OS=Homo sapiens GN=KRT14 PE=1 SV=4	619	51877	19
P19012	Keratin, type I cytoskeletal 15 OS=Homo sapiens GN=KRT15 PE=1	212	49413	8

	SV=3			
P08779	Keratin, type I cytoskeletal 16 OS=Homo sapiens GN=KRT16 PE=1 SV=4	569	51584	18
Q04695	Keratin, type I cytoskeletal 17 OS=Homo sapiens GN=KRT17 PE=1 SV=2	289	48366	9
P08727	Keratin, type I cytoskeletal 19 OS=Homo sapiens GN=KRT19 PE=1 SV=4	161	44079	6
P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3	1611	62259	42
Q9NSB4	Keratin, type II cuticular Hb2 OS=Homo sapiens GN=KRT82 PE=1 SV=3	49	58008	1
P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6	2188	66173	61
Q7Z794	Keratin, type II cytoskeletal 1b OS=Homo sapiens GN=KRT77 PE=1 SV=3	481	62154	12
P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2	788	65683	22
P12035	Keratin, type II cytoskeletal 3 OS=Homo sapiens GN=KRT3 PE=1 SV=3	113	64552	3
P19013	Keratin, type II cytoskeletal 4 OS=Homo sapiens GN=KRT4 PE=1 SV=4	207	57656	6
P13647	Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 PE=1 SV=3	425	62572	14
P02538	Keratin, type II cytoskeletal 6A OS=Homo sapiens GN=KRT6A PE=1 SV=3	454	60298	15
P04259	Keratin, type II cytoskeletal 6B OS=Homo sapiens GN=KRT6B PE=1 SV=5	516	60320	16
Q3SY84	Keratin, type II cytoskeletal 71 OS=Homo sapiens GN=KRT71 PE=1 SV=3	113	57778	3
Q14CN4	Keratin, type II cytoskeletal 72 OS=Homo sapiens GN=KRT72 PE=1 SV=2	97	56481	3
P05787	Keratin, type II cytoskeletal 8 OS=Homo sapiens GN=KRT8 PE=1 SV=7	107	53671	3
Q6KB66	Keratin, type II cytoskeletal 80 OS=Homo sapiens GN=KRT80 PE=1 SV=2	93	51016	2
P02788	Lactotransferrin OS=Homo sapiens GN=LTF PE=1 SV=6	125	80046	3
P98164	Low-density lipoprotein receptor-related protein 2 OS=Homo sapiens GN=LRP2 PE=1 SV=3	67	540699	1
P61626	Lysozyme C OS=Homo sapiens GN=LYZ PE=1 SV=1	42	16990	1
Q13835	Plakophilin-1 OS=Homo sapiens GN=PKP1 PE=1 SV=2	58	84142	1
P01833	Polymeric immunoglobulin receptor OS=Homo sapiens GN=PIGR PE=1 SV=4	78	84450	1
P01133	Pro-epidermal growth factor OS=Homo sapiens GN=EGF PE=1 SV=2	46	137677	1
P05109	Protein S100-A8 OS=Homo sapiens GN=S100A8 PE=1 SV=1	42	10886	1
P06702	Protein S100-A9 OS=Homo sapiens GN=S100A9 PE=1 SV=1	77	13292	2
A6NMY6	Putative annexin A2-like protein OS=Homo sapiens GN=ANXA2P2 PE=5 SV=2	71	38809	2
P02768	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2	106	71352	3
P02743	Serum amyloid P-component OS=Homo sapiens GN=APCS PE=1 SV=2	86	25487	2
Q9P2P6	StAR-related lipid transfer protein 9 OS=Homo sapiens GN=STARD9 PE=1 SV=3	48	521940	1

Supplementary Table A.2: Proteins identified in Fk-1 IAC sample 2. Uniprot accession, protein description, score, molecular weight (in Dalton) and number of protein matches are indicated in the table.

Accession	Description	Score	Mol Wt.	prot_matches
P60709	Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1	40	42052	1
Q9H444	Charged multivesicular body protein 4b OS=Homo sapiens GN=CHMP4B PE=1 SV=1	83	24935	3
P81605	Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2	77	11391	2
P01876	Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=2	85	38486	3
P01834	Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 SV=1	134	11773	4
P01591	Immunoglobulin J chain OS=Homo sapiens GN=IGJ PE=1 SV=4	129	18543	6
P13645	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6	1469	59020	70
P13646	Keratin, type I cytoskeletal 13 OS=Homo sapiens GN=KRT13 PE=1 SV=4	204	49900	8
P02533	Keratin, type I cytoskeletal 14 OS=Homo sapiens GN=KRT14 PE=1 SV=4	1038	51872	45
P08779	Keratin, type I cytoskeletal 16 OS=Homo sapiens GN=KRT16 PE=1 SV=4	660	51578	36
Q04695	Keratin, type I cytoskeletal 17 OS=Homo sapiens GN=KRT17 PE=1 SV=2	259	48361	10
P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3	1057	62255	35
P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6	1528	66170	75
Q7Z794	Keratin, type II cytoskeletal 1b OS=Homo sapiens GN=KRT77 PE=1 SV=3	528	62149	27
P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2	751	65678	36
P13647	Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 PE=1 SV=3	97	62568	4
P02538	Keratin, type II cytoskeletal 6A OS=Homo sapiens GN=KRT6A PE=1 SV=3	368	60293	19
P04259	Keratin, type II cytoskeletal 6B OS=Homo sapiens GN=KRT6B PE=1 SV=5	298	60315	12
P48668	Keratin, type II cytoskeletal 6C OS=Homo sapiens GN=KRT6C PE=1 SV=3	302	60273	8
Q3SY84	Keratin, type II cytoskeletal 71 OS=Homo sapiens GN=KRT71 PE=1 SV=3	107	57769	11
P59665	Neutrophil defensin 1 OS=Homo sapiens GN=DEFA1 PE=1 SV=1	52	10536	1
P01133	Pro-epidermal growth factor OS=Homo sapiens GN=EGF PE=1 SV=2	42	137613	1
P07911	Uromodulin OS=Homo sapiens GN=UMOD PE=1 SV=1	758	72451	25

Supplementary Table A.3: Proteins identified in Ub IAC sample 1. Uniprot accession, protein description, score, molecular weight (in Dalton) and number of protein matches are indicated in the table.

Accession	Description	Score	Mol Wt.	prot_matches
P03973	Antileukoproteinase OS=Homo sapiens GN=SLPI PE=1 SV=2	59	15228	2
P05090	Apolipoprotein D OS=Homo sapiens GN=APOD PE=1 SV=1	96	21547	2
P06576	ATP synthase subunit beta, mitochondrial OS=Homo sapiens GN=ATP5B PE=1 SV=3	80	56525	1
P31944	Caspase-14 OS=Homo sapiens GN=CASP14 PE=1 SV=2	77	27947	1
P01040	Cystatin-A OS=Homo sapiens GN=CSTA PE=1 SV=1	52	11000	1
Q02413	Desmoglein-1 OS=Homo sapiens GN=DSG1 PE=1 SV=2	62	114702	1
P15924	Desmoplakin OS=Homo sapiens GN=DSP PE=1 SV=3	90	334021	3
Q01469	Fatty acid-binding protein, epidermal OS=Homo sapiens GN=FABP5 PE=1 SV=3	43	15497	1
Q5D862	Filaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1	56	249296	1
Q86YZ3	Hornerin OS=Homo sapiens GN=HRNR PE=1 SV=2	42	283140	1
P01876	Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=2	117	38486	4
P01834	Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 SV=1	88	11773	1
P01593	Ig kappa chain V-I region AG OS=Homo sapiens PE=1 SV=1	45	12099	1
P01597	Ig kappa chain V-I region DEE OS=Homo sapiens PE=1 SV=1	45	11768	1
P01620	Ig kappa chain V-III region SIE OS=Homo sapiens PE=1 SV=1	66	11882	1
Q15323	Keratin, type I cuticular Ha1 OS=Homo sapiens GN=KRT31 PE=2 SV=3	139	48633	6
Q14525	Keratin, type I cuticular Ha3-II OS=Homo sapiens GN=KRT33B PE=2 SV=3	157	47325	6
P13645	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6	809	59020	27
P13646	Keratin, type I cytoskeletal 13 OS=Homo sapiens GN=KRT13 PE=1 SV=4	213	49900	7
P02533	Keratin, type I cytoskeletal 14 OS=Homo sapiens GN=KRT14 PE=1 SV=4	526	51872	19
P19012	Keratin, type I cytoskeletal 15 OS=Homo sapiens GN=KRT15 PE=1 SV=3	205	49409	8
P08779	Keratin, type I cytoskeletal 16 OS=Homo sapiens GN=KRT16 PE=1 SV=4	393	51578	14
Q04695	Keratin, type I cytoskeletal 17 OS=Homo sapiens GN=KRT17 PE=1 SV=2	312	48361	12
P08727	Keratin, type I cytoskeletal 19 OS=Homo sapiens GN=KRT19 PE=1 SV=4	194	44079	10
P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3	1295	62255	36
P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6	1758	66170	46
Q7Z794	Keratin, type II cytoskeletal 1b OS=Homo sapiens GN=KRT77 PE=1 SV=3	370	62149	10
P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2	639	65678	21
P19013	Keratin, type II cytoskeletal 4 OS=Homo sapiens GN=KRT4 PE=1 SV=4	102	57649	2
P13647	Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 PE=1 SV=3	327	62568	10
P02538	Keratin, type II cytoskeletal 6A OS=Homo sapiens GN=KRT6A PE=1 SV=3	255	60293	9
P04259	Keratin, type II cytoskeletal 6B OS=Homo sapiens GN=KRT6B PE=1 SV=5	312	60315	11
P48668	Keratin, type II cytoskeletal 6C OS=Homo sapiens GN=KRT6C PE=1 SV=3	242	60273	9
P08729	Keratin, type II cytoskeletal 7 OS=Homo sapiens GN=KRT7 PE=1 SV=5	89	51411	2
Q3SY84	Keratin, type II cytoskeletal 71 OS=Homo sapiens GN=KRT71 PE=1 SV=3	99	57769	2
A6NCN2	Keratin-81-like protein OS=Homo sapiens PE=2 SV=3	52	54972	1
P01042	Kininogen-1 OS=Homo sapiens GN=KNG1 PE=1 SV=2	59	72996	1

P61626	Lysozyme C OS=Homo sapiens GN=LYZ PE=1 SV=1	41	16982	1
O00187	Mannan-binding lectin serine protease 2 OS=Homo sapiens GN=MASP2 PE=1 SV=4	63	77193	1
P05154	Plasma serine protease inhibitor OS=Homo sapiens GN=SERPINA5 PE=1 SV=3	94	45760	2
P00747	Plasminogen OS=Homo sapiens GN=PLG PE=1 SV=2	44	93247	1
P01133	Pro-epidermal growth factor OS=Homo sapiens GN=EGF PE=1 SV=2	100	137613	2
P05109	Protein S100-A8 OS=Homo sapiens GN=S100A8 PE=1 SV=1	60	10885	1
P06702	Protein S100-A9 OS=Homo sapiens GN=S100A9 PE=1 SV=1	88	13291	1
P00734	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	48	71475	1
A6NDJ8	Putative Rab-43-like protein ENSP00000330714 OS=Homo sapiens PE=5 SV=3	54	20425	1
P59190	Ras-related protein Rab-15 OS=Homo sapiens GN=RAB15 PE=1 SV=1	54	24660	1
Q9H082	Ras-related protein Rab-33B OS=Homo sapiens GN=RAB33B PE=1 SV=1	54	26043	1
Q9NRW1	Ras-related protein Rab-6B OS=Homo sapiens GN=RAB6B PE=1 SV=1	54	23561	1
Q9HD89	Resistin OS=Homo sapiens GN=RETN PE=2 SV=1	105	12096	3
Q96P63	Serpin B12 OS=Homo sapiens GN=SERPINB12 PE=1 SV=1	48	46646	1
P29508	Serpin B3 OS=Homo sapiens GN=SERPINB3 PE=1 SV=2	47	44594	1
P02768	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2	45	71317	1
P07911	Uromodulin OS=Homo sapiens GN=UMOD PE=1 SV=1	335	72451	10

Supplementary Table A.4: Proteins identified in Ub IAC sample 2. Uniprot accession, protein description, score, molecular weight (in Dalton) and number of protein matches are indicated in the table.

Accession	Description	Score	Mol Wt.	prot_matches
P05090	Apolipoprotein D OS=Homo sapiens GN=APOD PE=1 SV=1	42	21547	1
Q7LBR1	Charged multivesicular body protein 1b OS=Homo sapiens GN=CHMP1B PE=1 SV=1	50	22152	2
O43633	Charged multivesicular body protein 2a OS=Homo sapiens GN=CHMP2A PE=1 SV=1	41	25088	1
Q9BY43	Charged multivesicular body protein 4a OS=Homo sapiens GN=CHMP4A PE=1 SV=3	56	25083	1
Q9H444	Charged multivesicular body protein 4b OS=Homo sapiens GN=CHMP4B PE=1 SV=1	93	24935	3
Q9NZZ3	Charged multivesicular body protein 5 OS=Homo sapiens GN=CHMP5 PE=1 SV=1	116	24612	4
P81605	Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2	149	11391	6
P01876	Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=2	121	38486	6
P01834	Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 SV=1	233	11773	5
P01614	Ig kappa chain V-II region Cum OS=Homo sapiens PE=1 SV=1	48	12782	1
P01591	Immunoglobulin J chain OS=Homo sapiens GN=IGJ PE=1 SV=4	166	18543	9
P53990	IST1 homolog OS=Homo sapiens GN=IST1 PE=1 SV=1	131	39897	8
P13645	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6	1349	59020	59
P13646	Keratin, type I cytoskeletal 13 OS=Homo sapiens GN=KRT13 PE=1 SV=4	176	49900	8
P02533	Keratin, type I cytoskeletal 14 OS=Homo sapiens GN=KRT14 PE=1 SV=4	627	51872	30
P08779	Keratin, type I cytoskeletal 16 OS=Homo sapiens GN=KRT16 PE=1 SV=4	590	51578	31

Q04695	Keratin, type I cytoskeletal 17 OS=Homo sapiens GN=KRT17 PE=1 SV=2	80	48361	4
P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3	1509	62255	50
P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6	1569	66170	68
Q7Z794	Keratin, type II cytoskeletal 1b OS=Homo sapiens GN=KRT77 PE=1 SV=3	593	62149	26
P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2	836	65678	35
P13647	Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 PE=1 SV=3	129	62568	6
P02538	Keratin, type II cytoskeletal 6A OS=Homo sapiens GN=KRT6A PE=1 SV=3	155	60293	12
P04259	Keratin, type II cytoskeletal 6B OS=Homo sapiens GN=KRT6B PE=1 SV=5	341	60315	14
P48668	Keratin, type II cytoskeletal 6C OS=Homo sapiens GN=KRT6C PE=1 SV=3	137	60273	5
Q7RTS7	Keratin, type II cytoskeletal 74 OS=Homo sapiens GN=KRT74 PE=1 SV=2	66	58229	7
P02788	Lactotransferrin OS=Homo sapiens GN=LTF PE=1 SV=6	92	80014	3
P59665	Neutrophil defensin 1 OS=Homo sapiens GN=DEFA1 PE=1 SV=1	78	10536	3
P05154	Plasma serine protease inhibitor OS=Homo sapiens GN=SERPINA5 PE=1 SV=3	53	45760	1
Q9HD89	Resistin OS=Homo sapiens GN=RETN PE=2 SV=1	47	12096	1
P04279	Semenogelin-1 OS=Homo sapiens GN=SEMG1 PE=1 SV=2	74	52157	2
Q15678	Tyrosine-protein phosphatase non-receptor type 14 OS=Homo sapiens GN=PTPN14 PE=1 SV=2	40	136031	1
P07911	Uromodulin OS=Homo sapiens GN=UMOD PE=1 SV=1	992	72451	40

Supplementary Table A.5: Proteins identified in rabbit IgG IAC (negative control). Uniprot accession, protein description, score, molecular weight (in Dalton) and number of protein matches are indicated in the table.

Accession	Description	Score	Mol Wt.	prot_matches
P04745	Alpha-amylase 1 OS=Homo sapiens GN=AMY1A PE=1 SV=2	78	58415	1
P05090	Apolipoprotein D OS=Homo sapiens GN=APOD PE=1 SV=1	84	21547	3
P31944	Caspase-14 OS=Homo sapiens GN=CASP14 PE=1 SV=2	117	27947	3
P01040	Cystatin-A OS=Homo sapiens GN=CSTA PE=1 SV=1	132	11000	3
P01036	Cystatin-S OS=Homo sapiens GN=CST4 PE=1 SV=3	44	16489	1
P01037	Cystatin-SN OS=Homo sapiens GN=CST1 PE=1 SV=3	44	16605	1
P81605	Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2	146	11391	4
Q08554	Desmocollin-1 OS=Homo sapiens GN=DSC1 PE=1 SV=2	68	101406	1
Q02413	Desmoglein-1 OS=Homo sapiens GN=DSG1 PE=1 SV=2	146	114702	4
P15924	Desmoplakin OS=Homo sapiens GN=DSP PE=1 SV=3	128	334021	3
P14625	Endoplasmin OS=Homo sapiens GN=HSP90B1 PE=1 SV=1	61	92696	1
Q01469	Fatty acid-binding protein, epidermal OS=Homo sapiens GN=FABP5 PE=1 SV=3	40	15497	1
Q5D862	Filaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1	84	249296	1
Q86YZ3	Hornerin OS=Homo sapiens GN=HRNR PE=1 SV=2	150	283140	4
P01876	Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=2	41	38486	1

Q15323	Keratin, type I cuticular Ha1 OS=Homo sapiens GN=KRT31 PE=2 SV=3	307	48633	9
Q14525	Keratin, type I cuticular Ha3-II OS=Homo sapiens GN=KRT33B PE=2 SV=3	214	47325	7
O76011	Keratin, type I cuticular Ha4 OS=Homo sapiens GN=KRT34 PE=2 SV=2	187	50818	6
P13645	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6	826	59020	17
P13646	Keratin, type I cytoskeletal 13 OS=Homo sapiens GN=KRT13 PE=1 SV=4	323	49900	10
P02533	Keratin, type I cytoskeletal 14 OS=Homo sapiens GN=KRT14 PE=1 SV=4	585	51872	14
P19012	Keratin, type I cytoskeletal 15 OS=Homo sapiens GN=KRT15 PE=1 SV=3	268	49409	8
P08779	Keratin, type I cytoskeletal 16 OS=Homo sapiens GN=KRT16 PE=1 SV=4	544	51578	14
Q04695	Keratin, type I cytoskeletal 17 OS=Homo sapiens GN=KRT17 PE=1 SV=2	375	48361	9
P08727	Keratin, type I cytoskeletal 19 OS=Homo sapiens GN=KRT19 PE=1 SV=4	149	44079	4
P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3	713	62255	17
P78386	Keratin, type II cuticular Hb5 OS=Homo sapiens GN=KRT85 PE=1 SV=1	237	57306	9
O43790	Keratin, type II cuticular Hb6 OS=Homo sapiens GN=KRT86 PE=1 SV=1	375	55120	14
P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6	1151	66170	30
P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2	1033	65678	27
Q01546	Keratin, type II cytoskeletal 2 oral OS=Homo sapiens GN=KRT76 PE=1 SV=2	163	66370	6
P12035	Keratin, type II cytoskeletal 3 OS=Homo sapiens GN=KRT3 PE=1 SV=3	185	64549	5
P19013	Keratin, type II cytoskeletal 4 OS=Homo sapiens GN=KRT4 PE=1 SV=4	136	57649	5
P13647	Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 PE=1 SV=3	612	62568	18
P02538	Keratin, type II cytoskeletal 6A OS=Homo sapiens GN=KRT6A PE=1 SV=3	715	60293	20
P04259	Keratin, type II cytoskeletal 6B OS=Homo sapiens GN=KRT6B PE=1 SV=5	647	60315	19
P48668	Keratin, type II cytoskeletal 6C OS=Homo sapiens GN=KRT6C PE=1 SV=3	643	60273	18
Q14CN4	Keratin, type II cytoskeletal 72 OS=Homo sapiens GN=KRT72 PE=1 SV=2	127	56470	4
Q86Y46	Keratin, type II cytoskeletal 73 OS=Homo sapiens GN=KRT73 PE=1 SV=1	131	59457	5
O95678	Keratin, type II cytoskeletal 75 OS=Homo sapiens GN=KRT75 PE=1 SV=2	236	59809	8
A6NCN2	Keratin-81-like protein OS=Homo sapiens PE=2 SV=3	321	54972	12
Q52LG2	Keratin-associated protein 13-2 OS=Homo sapiens GN=KRTAP13-2 PE=1 SV=1	47	19912	1
P07195	L-lactate dehydrogenase B chain OS=Homo sapiens GN=LDHB PE=1 SV=2	66	36900	1
P59665	Neutrophil defensin 1 OS=Homo sapiens GN=DEFA1 PE=1 SV=1	46	10536	1
P12273	Prolactin-inducible protein OS=Homo sapiens GN=PIP PE=1 SV=1	43	16847	1
P31151	Protein S100-A7 OS=Homo sapiens GN=S100A7 PE=1 SV=4	45	11578	1
P05109	Protein S100-A8 OS=Homo sapiens GN=S100A8 PE=1 SV=1	115	10885	3
P06702	Protein S100-A9 OS=Homo sapiens GN=S100A9 PE=1 SV=1	108	13291	2
P14618	Pyruvate kinase isozymes M1/M2 OS=Homo sapiens GN=PKM2 PE=1 SV=4	113	58470	1
Q96P63	Serpin B12 OS=Homo sapiens GN=SERPINB12 PE=1 SV=1	46	46646	1
P29508	Serpin B3 OS=Homo sapiens GN=SERPINB3 PE=1 SV=2	42	44594	1
P02768	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2	71	71317	2

Q9H2G4	Testis-specific Y-encoded-like protein 2 OS=Homo sapiens GN=TSPYL2 PE=1 SV=1	43	79615	1
P10599	Thioredoxin OS=Homo sapiens GN=TXN PE=1 SV=3	66	12015	1
Q9BQE3	Tubulin alpha-1C chain OS=Homo sapiens GN=TUBA1C PE=1 SV=1	51	50548	1
P62979	Ubiquitin-40S ribosomal protein S27a OS=Homo sapiens GN=RPS27A PE=1 SV=2	71	18296	1