Investigation of lysosomal phosphoproteome changes in altered cholesterol metabolism in Niemann-Pick Disease Type C (NPC)

Dissertation

zur Erlangung des Doktorgrades (PhD)

der Medizinischen Fakultät

der Rheinischen Friedrich-Wilhelms-Universität

Bonn

von

Fatema Akter

aus

Brahmanbaria, Bangladesh

Bonn, 2019

Dedicated to

My father Md. Shafiqul Islam, who believed that only education can bring a change

Table of Contents

List of Abbreviation

A	Alanine	LE	Late endosome
BSA	Bovine serum albumin	LSD	Lysosomal Storage Disorder
CID	Collision-induced dissociation	m/z	Mass-to-charge ratio
CoIP	Co-immunoprecipitation	MBP	Maltose Binding Protein
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats	MEF	Mouse embryonic fibroblast cells
D	Aspartic Acid	MS	Mass spectrometry
DAPI	4',6-diamidin-2-phenylindol	4MU	4- methylumbelliferyl
DMSO	Dimethyl sulfoxide	NPC	Niemann-Pick Type C disease
E	Glutamic acid	PAGE	Polyacrylamide gel electrophoresis
ECL	Enhanced chemiluminescence	PCR	Polymerase chain reaction
ER	Endoplasmic reticulum	PTM	Post-translational modification
FCS	Fetal calf serum	S	Serine
FDR	False discovery rate	SDM	Site-directed mutagenesis
FeDex	Colloidal iron dextran particles	SDS	Sodium dodecyl sulfate
GO	Gene ontology	SILAC	Stable Isotope Labeling by Amino Acids in Cell Culture
HEK293	Human embryonic kidney cells	SPE	Solid phase extraction
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid	SPIONs	Superparamagnetic Iron Oxide Nanoparticles
CoIP	Immunoprecipitation	StageTips	Stop-and-go-extraction tips
KO	Knockout	Tris/HCl	Tris(hydroxymethyl)-aminomethan hydrochloride
LC-MS/MS	Liquid chromatography-tandem mass spectrometry	U18666A	(3β)-3-[2-(Diethylamino)ethoxy]androst-5-en-17-one hydrochloride

1. Introduction

1.1 Lysosomes

Lysosomes, membrane-limited cellular organelles, can degrade a vast array of cellular components including proteins, carbohydrates, nucleic acids, and lipids. This is why they are also referred to as "cellular waste bags" (de Duve 2005). Extracellular materials are transported to the lysosome either by endocytosis or phagocytosis whereas intracellular components are delivered by autophagy. More than 60 different hydrolytic enzymes in the lysosomal lumen degrade the substrates and the degradation products are transported out of the lysosome either via exocytosis (Saftig and Klumperman 2009) or via specific catabolite exporters (Ruivo et al. 2009). Apart from the soluble hydrolases, more than 50 different integral membrane proteins are also known to be present in lysosomes which are mainly transporters, ion channels, SNAREs and members of the vATPase complex (Xu and Ren 2015). Proton-importing v-type ATPase membrane proteins in lysosomes maintain the constant acidic pH (~4.5) in the lumen (Pu et al. 2016). This acidic pH which is optimal for the hydrolytic enzymes also ensures that in the case of lysosomal rupture, there is no degradation of other cellular components. In addition to the luminal and integral membrane proteins, there are several proteins and protein complexes which associate with the lysosomal surface and elicit lysosomal functions such as nutrient sensing by mTORC1 (Zoncu et al. 2011), or lysosomal positioning by the BORC complex (Pu et al. 2015). There are still many lysosomal proteins whose functions are unknown, which creates an active field of research in order to determine proteins involved in lysosomal biology.

1.2 Lysosomal Storage Disorders (LSDs)

Lysosomal storage disorders (LSDs) are a group of about 50 different genetic disorders caused by deficiencies or mutations in lysosomal proteins (Ballabio and Gieselmann 2009). In case any lysosomal enzymes are affected their substrates accumulate inside the lysosome resulting in progressive lysosomal dysfunction, ultimately leading to cellular dysfunction and cell death. The combined incidence of LSDs is estimated to be approximately 1 in 5,000 live births (Fuller et al. 2006). Most LSDs are present as infantile, juvenile or adult forms, which depending on the severity they present acute neurological or other symptoms or premature death (Platt et al. 2012). Lysosomal storage results in various pathogenic events such as inflammation, altered calcium homeostasis, oxidative stress, altered lipid trafficking,

reduced autophagy, endoplasmic reticulum stress and autoimmune responses (Vitner et al. 2010). Several diagnostic tests like enzymatic analysis, single-gene sequencing and most recently, next-generation sequencing (particularly whole-exome sequencing) are used for detecting LSDs (Platt et al. 2018). Even though no complete recovery is possible but for treating these disorders with enzyme replacement therapies (Zimran 2011), small-molecule therapies including substrate reduction (Aerts et al. 2006) and chaperone therapies (Hughes et al. 2017) are currently in use. However, despite the great progress in understanding the genetic, molecular and biochemical mechanisms of LSDs, almost nothing is known about how lysosomal accumulation causes dysfunction at the cellular level (Marques and Saftig 2019).

1.3 Niemann-Pick Disease Type C (NPC)

Niemann-Pick diseases are a heterogeneous group of autosomal recessive LSD which are classified as belonging to three different groups: type A, B, and C (Crocker and Farber 1958). Type A and B result from a deficiency of the lysosomal enzyme acid sphingomyelinase (Hollak et al. 2012); whereas type C occurs either by a mutation in NPC1 (in 95 % cases) or in NPC2 gene (in 5 % cases) (Vanier 2010). NPC is a neurovisceral atypical lysosomal lipid storage disorder with an incidence of 1 in 120,000 live births and it's severity ranges from early neonatal death to chronic neurodegeneration in adults (Vanier 2010). The exact molecular mechanism of NPC1 and NPC2 proteins is still unclear (Vanier 2010; Vance and Karten 2014) but it has been shown that an impaired egress of unesterified cholesterol from late endosomes/ lysosomes is the main feature of NPC disease (Vanier 2015). The NPC1 protein has 13 transmembrane domains and localizes to the lysosomal membrane (Davies and Ioannou 2000) whereas NPC2 is localized in the lysosomal lumen. These two proteins function in a coordinated fashion to facilitate the intracellular transport of cholesterol from the lysosome to the ER (Sleat et al. 2004). NPC2 protein binds to the acid lipase cleaved cholesterol esters from low-density lipoproteins (LDLs) and transfers them to the sterol-sensing domain (SSD) at the N-terminal part of membrane-bound NPC1, which then passes the cholesterol across the lysosomal membrane (Kwon et al. 2009) (Figure 1.1).
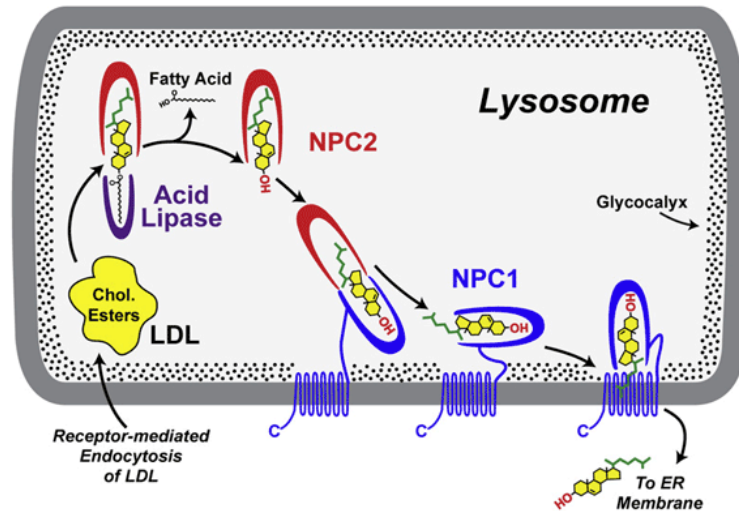
Figure 1.1: Pathway of cholesterol transportation by NPC1 and NPC2 proteins.
Adopted from Kwon et al. (Kwon et al. 2009)

In order to better understand NPC disease, several models have been developed. Studying fibroblasts from NPC patients was the first model (Liscum et al. 1989) followed by NPC mutant cell lines from different species like Chinese hamster (Cadigan et al. 1990), *Caenorhabditis elegans* (Sym et al. 2000), yeast (Malathi et al. 2004), *Drosophila melanogaster* (Huang et al. 2005) and knock out mouse models of NPC1 (Miyawaki et al. 1982) and NPC2 (Nielsen et al. 2011). Several chemical tools like progesterone, imipramine, and U18666A have been found to strongly inhibit the re-esterification of lysosomal cholesterol in the ER (Lange et al. 1997). When focusing on U18666A, which is an amphipathic steroid 3-β-[2-(diethylamino) ethoxy]androst-5-en-17-one, it is widely used in cell culture to mimic NPC disease phenotype (Liscum and Klansek 1998; Higgins et al. 1999; Lange et al. 2002; Karten et al. 2003; Pacheco et al. 2007). It has been suggested that U18666A blocks cholesterol egression from lysosomes, alters the transport and subcellular distribution of NPC protein (Higgins et al. 1999; Ko et al. 2001) and in doing so, U18666A may directly interact with the NPC protein to produce its effect (Liscum and Klansek 1998). U18666A treatment of normal steroidogenic cells showed a marked increase in NPC1 mRNA and the protein level (Watari et al. 2000). Moreover, studies on NPC-deficient CHO cells established that U18666A's effect on lipid accumulation is at least partially dependent on the presence of intact NPC protein. Transfection of cultured cells with mutated-nonfunctional NPC, accumulate cholesterol and gangliosides, which doesn't increase further after U18666A treatment. However, the capacity of U18666A to block lipid trafficking and raise cellular lipid content was restored by "knocking in" of the NPC protein (Sugimoto et al. 2001).

Intralysosomal cholesterol accumulation can be verified using fluorescence microscopy by filipin staining (Arthur et al. 2011). Filipin is a polyene antifungal antibiotic produced naturally by the bacteria *Streptomyces filipinensis*, and upon staining it forms a specific complex with unesterified cholesterols and emits blue fluorescence (Whitfield et al. 1955). In the current study, U18666A is used for creating the NPC phenotype in mouse embryonic fibroblasts (MEFs) and filipin is used for the staining of accumulated cholesterol.

1.4 Enrichment of lysosomes

For proteome research and the diagnosis of specific protein defects, it is necessary to purify organelles from cultured cells in order to acquire organelle-associated proteins. Additionally, mass spectrometry-based proteomics can identify and quantify thousands of proteins from a sample mixture and shed light on the understanding of many diseases. However, due to the sample complexity or possible contamination by other organelles, organelle-specific studies can become challenging sometimes. Therefore, first enriching the organelle followed by a mass spectrometry analysis would be the best approach for proper identification of novel proteins. There are various techniques which have been established and used for the enrichment of lysosomes. One of the most commonly used methods is subcellular fractionation which uses different density gradient materials like sucrose (de Duve et al. 1955), Metrizamide (Wattiaux et al. 1978), Percol (Schroeder et al. 2004), and Nycodenz (Chapel et al. 2013). Alternatively, by using an *in vivo* approach of injecting Triton WR-1339 into rats, hepatic lysosomes can be isolated due to their generated density shift (Wattiaux et al. 1963). Lysosomes can also be enriched by the immunoprecipitation technique where lysosomal proteins are epitope-tagged (Zoncu et al. 2011; Abu-Remaileh et al. 2017). Although these methods can provide a high yield of lysosomes, they also contain microsomal or mitochondrial contaminants. Specific targeting of lysosomal compartments by magnetic nanoparticles is an attractive approach for lysosome enrichment which can be used in a variety of cell lines. For this method, cells are supplemented with either colloidal iron dextran nanoparticles (FeDex) (Diettrich et al. 1998) or superparamagnetic iron oxide nanoparticles (SPIONs) (Walker and Lloyd-Evans 2015) and during this process, these nanoparticles are endocytosed in an unspecific manner. After a period of pulse and chase, they are delivered into the lysosomal compartment and by using a strong magnetic field lysosome enriched fractions are collected (Figure 1.2). SPIONs are

remaining in the fluid phase (supernatant) while running on the magnetic field and they are stable in the acidic pH environment of the lysosome, thereby allowing a rapid enrichment of intact lysosomes (Walker and Lloyd-Evans 2015).
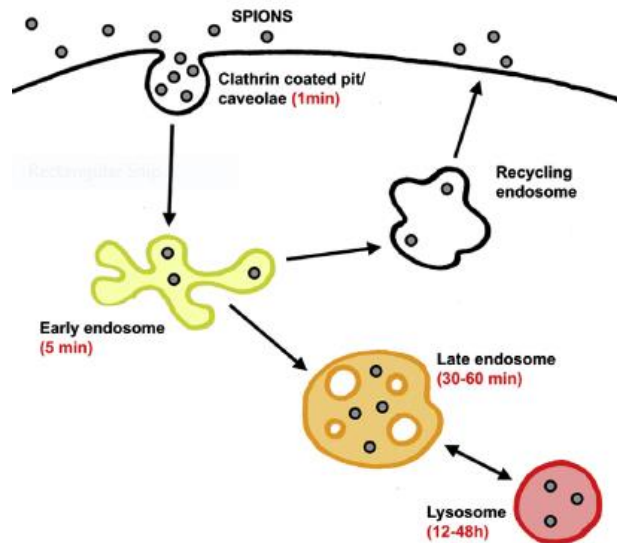
Figure 1.2: Trafficking of SPIONs through the endocytic system.
Schematic representation of the different stages of the endocytic system and the loading times of SPIONs into each individual compartment (red text in brackets). Adapted from (Walker and Lloyd-Evans 2015).

1.5 Mass spectrometry-based quantitative proteomics

Mass spectrometry (MS)-based proteomics studies have become the method of choice for the analysis of complex protein samples in the last decades (Aebersold and Mann 2003). It has been successfully applied in diverse areas like determining protein composition of organelles, signaling pathways, protein-protein interaction, determining their substrates and large-scale mapping of protein phosphorylation in response to a stimulus (Baker 2010). Mass spectrometry analysis of proteins can be performed either by top-down or bottom-up approaches (Chait 2006). Top-down proteomics is used for whole-protein analysis whereas in bottom-up proteomics, after extracting proteins from different samples they are digested into peptides by sequence-specific enzymes e.g. trypsin. Peptides are then separated by reversed-phase chromatography and ionized by electrospray ionization before entering into the mass spectrometer (Figure 1.3) (Zhang et al. 2013).

A mass spectrometer consists of an ion source, a mass analyzer and a detector (Awad et al. 2015). Measurements in mass spectrometers are generally carried out at the gas phase on the ionized analytes. In order to volatize and ionize peptides, electrospray ionization (ESI) (Fenn et al. 1989) or matrix-assisted laser desorption/ ionization (MALDI) (Karas and Hillenkamp 1988) are the most commonly used

ionization methods. Mass analyzer measures mass (m) to charge (z) ratio (m/z) of the generated precursor ions and the detector detects m/z value and the intensity of each precursor ion (Awad et al. 2015). Using tandem mass spectrometry, a particular peptide precursor ion is further fragmented to product ions by collision with inert gas (e.g. nitrogen molecules or argon or helium atoms) which provides information for peptide sequencing (Steen and Mann 2004). Generated raw- data from mass spectrometers can be analyzed either manually or using the software in comparison to the existing databases, providing information regarding the identification and quantification of proteins (Schmidt et al. 2014).
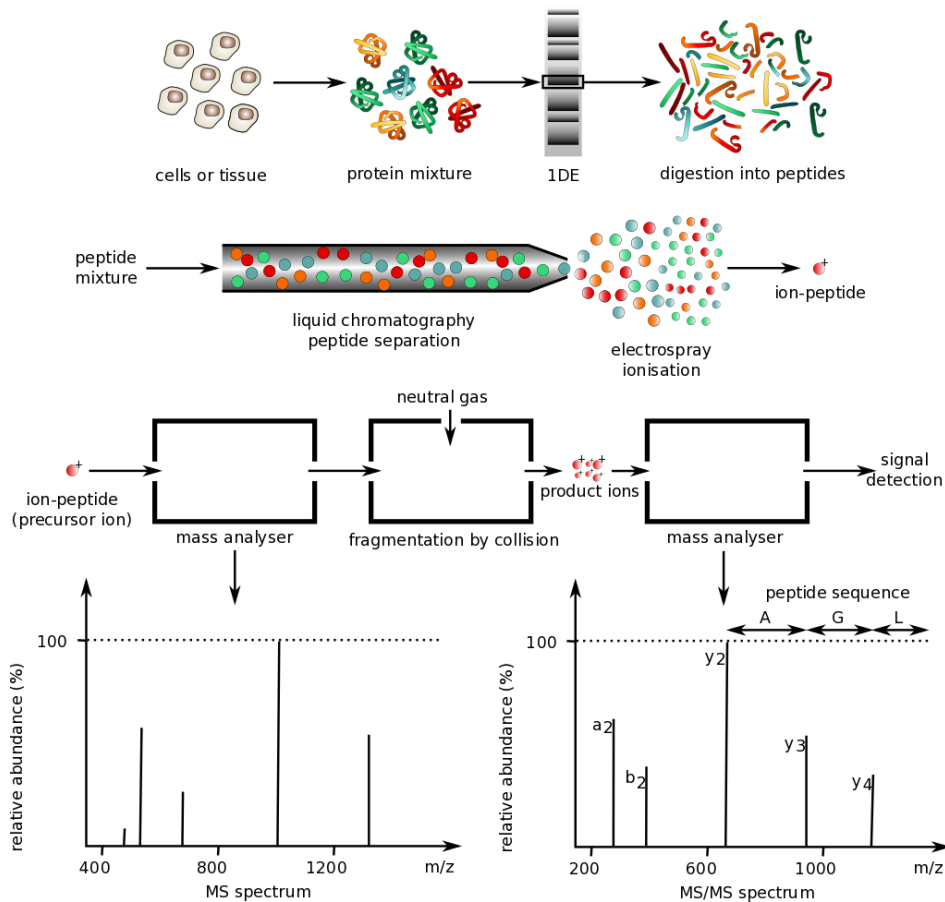
Figure 1.3: Schematics of bottom-up proteomics workflow.
Proteins are extracted from cells or tissues are either size-separated by 1D gel electrophoresis (1DE) or proceed in solution, digested into peptides by sequence-specific enzyme, separated by liquid chromatography (LC) and ionized in the gas phase using electrospray ionization (ESI). Mass analyzer and the detector detect the m/z ratio and relative abundance of ionized peptides providing MS spectrum. To obtain information about the peptide sequence or post-translational modification (PTM), precursor ions are fragmented by colliding the accelerated ionized peptides with neutral gas molecules followed by recording the m/z ratio of each fragment (MS/MS or MS2). Adopted from (Barillot et al. 2012)

Although, MS was initially considered as a qualitative method, with the recent development of instruments and software it has turned into a quantitative tool.

Quantitation can be performed either with label-free or label-based methods (Figure 1.4). In the label-free quantification method, proteins are quantitated both relatively and absolutely by utilizing the signal intensity and spectral counting of peptides (Cox et al. 2014). In contrast, label-based methods utilize stable isotopes incorporated into peptides and introduce an expectable mass difference with different experimental conditions. Several label-based methods are available where each has some advantages and disadvantages.
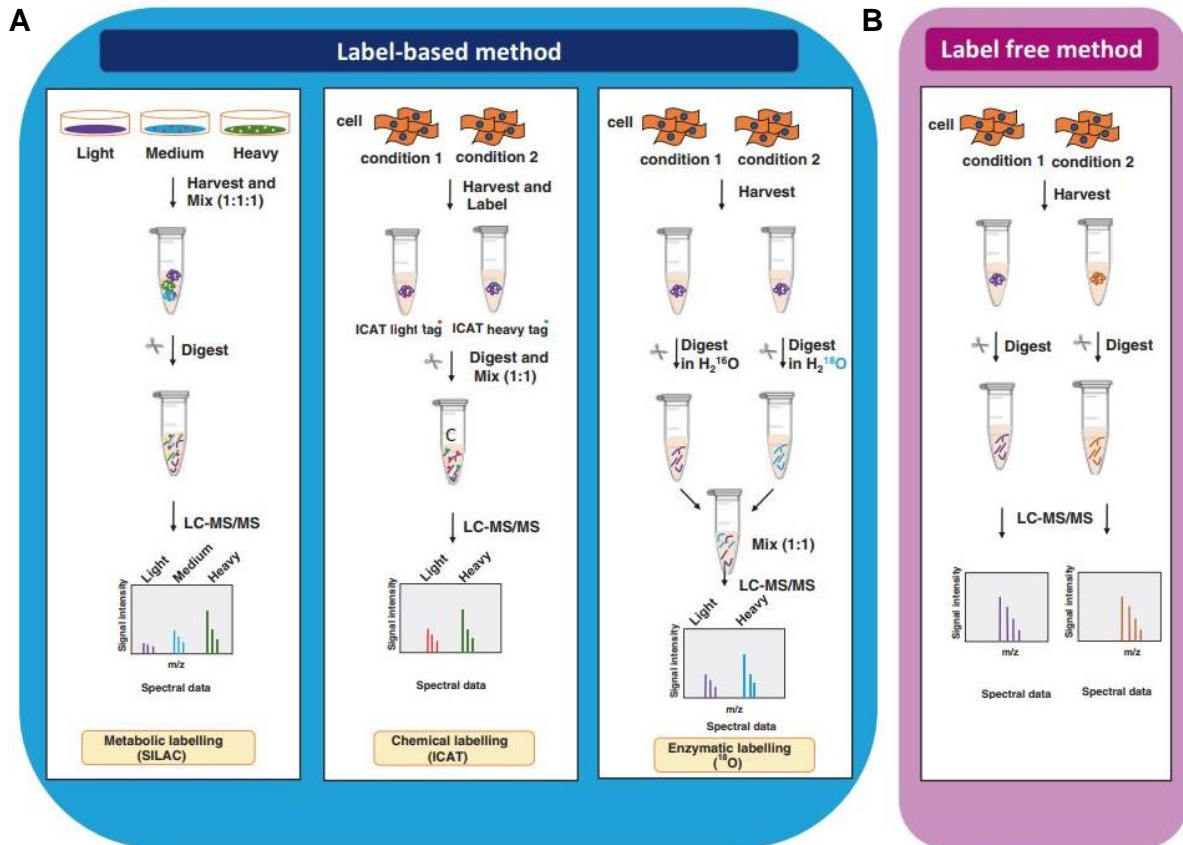
Figure 1.4: Common quantitative mass spectrometry workflows.
(A) In the label-based approach, peptides are labeled with stable isotopes at different stages depending on the type of technique applied and more than one biological samples are simultaneously profiled in a single MS run. After mixing the labeled samples in an equivalent ratio they are analyzed by LC-MS/MS. The quantitation is based on the comparison of the peak intensity ratio of the labeled peptide pairs. (B) In the label-free method, the samples are separately prepared and are subjected to individual LC-MS/MS analysis. Further quantification is based on comparing the counts of MS/MS spectra or peak intensity of the same peptide. Adapted from (Anand et al. 2017)

Proteins or peptides can be labeled metabolically *in vivo* by stable isotope labeling by amino acids in cell culture (SILAC) method (Ong et al. 2002) or chemically *in vitro* before or after digestion using for example dimethyl labeling (Boersema et al. 2009), tandem mass tags (TMT) (Thompson et al. 2003), or isobaric tags for relative and absolute quantification (iTRAQ) (Ross et al. 2004). In the SILAC method, stable isotopes (^{13}C, ^{15}N) of labeled amino acids (e.g. lysine or arginine is used when

trypsin functions as a protease during sample preparation) are incorporated metabolically into the peptides or protein sequences after several passages in the cell culture by feeding cells with supplemented SILAC media (Ong and Mann 2005). Samples can be mixed and used together for mass spectrometry analysis where ionization of both labeled samples are the same, generating a mass signal with a mass shift due to labeled stable isotope (Ong and Mann 2005).

1.6 Phosphorylation analysis by mass spectrometry

Phosphorylation is the most common protein post-translational modification and plays a pivotal role in numerous cellular processes like differentiation, proliferation, signal transduction, regulation of enzyme activity as well as protein-protein interaction (Steen et al. 2006). In eukaryotes, it is estimated that one-third of all proteins are phosphorylated and over 500 kinases are responsible for phosphorylation while approximately 150 phosphatases are responsible for dephosphorylation (Arrington et al. 2017). Despite the existence of many traditional approaches, it has been very challenging to understand the complexity of protein phosphorylation or phosphoprotein function in signaling networks (Mumby and Brekken 2005). The field of mass spectrometry-based phosphoproteomics has made it possible to study and discover many phosphoproteins and phosphorylation sites at a time. Prior to the phosphoproteome analysis by MS, enrichment of phosphopeptides is necessary due to the low abundance of phosphopeptides in a complex sample mixture. Several enrichment methods have been developed until now for phosphopeptides enrichment including immobilized metal ion affinity chromatography (IMAC) (Andersson and Porath 1986), metal oxide affinity chromatography (MOAC) (Larsen et al. 2005) and immunoaffinity chromatography (Rush et al. 2005). Besides these affinity methods, in order to separate negatively charged phosphopeptides from their non-phosphorylated counterparts different liquid chromatography (LC)- based fractionation approaches have been developed e.g. strong anion exchange chromatography (SAX) (Han et al. 2008), strong cation exchange chromatography (SCX) (Beausoleil et al. 2004), hydrophilic interaction liquid chromatography (HILIC) (McNulty and Annan 2008), and electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) (Alpert 2008). Recently, it has been observed that the enrichment of phosphopeptides prior to fractionation is more cost-effective and time-saving (Dehghani et al. 2018).

1.7 Phosphorylation studies on NPC

Several studies on Niemann-Pick Type C disease models have shown the regulation of the activity of different kinases and ultimately phosphorylation of various proteins. A study addressing several protein kinases in livers of NPC1 knock out mice showed that the expression level of protein kinase C (PKC)-alpha, PKC-zeta, proto-oncogene tyrosine-protein kinase src (pp60-src), and PKC-delta have increased significantly in NPC1 livers in comparison to wild-type organs (Garver et al. 1999). Another study on brains of NPC1 homozygous and heterozygous mice revealed that an abnormal cholesterol metabolism due to NPC1 mutation activates a mitogen-activated protein kinase (MAPK) which phosphorylates the tau protein leading to tauopathy in NPC (Sawamura et al. 2001). Another group investigated phosphorylated alpha-synuclein immunoreactivity in the brain of twelve NPC1 patients and suggested that the defect in intracellular cholesterol trafficking in NPC1 provokes aberrant phosphorylation of alpha-synuclein and tau protein (Saito et al. 2004). Research on human NPC1 fibroblasts showed hyperphosphorylation of vimentin by protein kinase c (PKC) in comparison to WT cells and moreover, activation of PKC can ameliorate NPC1 cholesterol transport block (Tamari et al. 2013). All these findings demonstrated significant alterations in phosphoproteome of NPC1 cells but no phosphoproteomics study has been performed on enriched lysosome fractions from NPC disease models yet.

1.8 The Mucopolysaccharidoses

The larger lysosomal storage disease (LSD) family comprises a group of seven rare genetic disorders, known as mucopolysaccharidoses (MPS) (Wijburg et al. 2013). As with most of the LSDs, each of the MPS is characterized by the deficit or mutation of a specific lysosomal enzyme that leads to the accumulation of undigested or partially digested macromolecules within lysosomes. In the MPSs, the accumulation of long unbranched polysaccharides, glycosaminoglycans (GAGs), produce progressive cellular damage that results in multisystemic disease. Mucopolysaccharidosis type III (MPS III or Sanfilippo syndrome) is the most common form of MPS, which is a neurodegenerative LSD whose symptoms are caused by the deficiency of enzymes involved exclusively in heparan sulfate degradation (Meikle et al. 1999). The primary characteristic of MPS III is the degeneration of the central nervous system, resulting in mental retardation and hyperactivity, typically during early childhood (de Ruijter et al. 2013). Mucopolysaccharidosis III is composed of four different subtypes: type A,

B, C, and D. All four subtypes are inherited in an autosomal recessive pattern. The incidence of the subtypes has a very uneven geographic distribution. Together, the reported incidence of all subtypes of MPS III varies between 0.28 and 4.1 per 100,000 live births, with types A and B being more common than types C and D (Valstar et al. 2008). To diagnose these diseases, a number of genetic and biochemical diagnostic methods have been developed and adopted by diagnostic laboratories. However, there is no effective therapy available for any form of MPS III, with treatment currently limited to clinical management of neurological symptoms. The availability of animal models for all forms of MPS III, whether spontaneous or generated via gene targeting, has contributed to improve the understanding of the MPS III subtypes and will deliver invaluable tools to appraise emerging therapies. Indeed, clinical trials to evaluate intrathecally-delivered enzyme replacement therapy in MPS IIIA patients, and gene therapy for MPS IIIA and MPS IIIB patients are planned or underway (Fedele 2015).

1.9 Heparan-α-glucosaminide N-acetyltransferase (HGSNAT)

HGSNAT is an integral lysosomal membrane protein which acetylates the non-reducing, terminal α-glucosamine residue of intralysosomal heparin or heparan sulfate. It uses cytosolic acetyl-CoA and converts it into the substrate for luminal α-N-acetylglucosaminidase (NAGLU) (Fan et al. 2006; Hrebícek et al. 2006). This reaction is crucial for the degradation of HS because there is no enzyme that can act on the unacetylated glucosamine molecule. Therefore, it catalyzes the only known synthetic reaction to occur in the lysosome. However, it can also acetylate the β-glucosamine residues of artificial substrates, which is the basis of HGSNAT activity assay coupled with β-hexosaminidase (Voznyi et al. 1993). Deficiency of HGSNAT results in an autosomal recessive disorder mucopolysaccharidosis IIIC (MPSIIIC, MIM:252930), also called Sanfilippo syndrome type C, which is characterized by the lysosomal storage of undegraded heparin and heparan sulfate (Kresse et al. 1978).

The mechanism by which HGSNAT uses cytosolic acetyl-CoA to transfer the acetyl group to the lysosomal luminal substrate is controversial (Fan et al. 2011). The most accepted model is a ping pong (double displacement) mechanism, where the enzyme gets acetylated by forming a covalently modified enzyme intermediate in the cytosol and then transfers the acetyl group to the intra-lysosomal heparin-α-glucosamine residue (Figure 1.5 (A)) (Bame and Rome 1986; Durand et al. 2010).
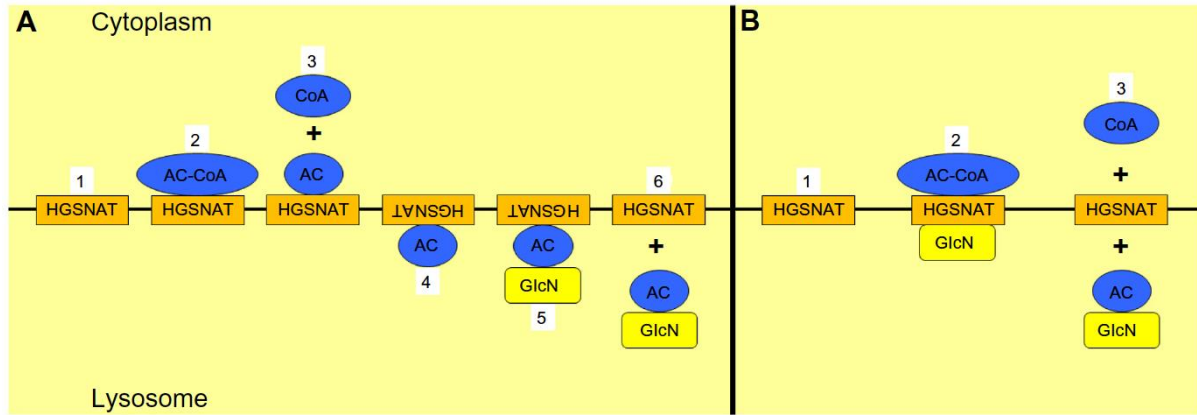
Figure 1.5: Proposed model for the mechanism of HGSNAT activity.
A. HGSNAT (1) acquires acetyl CoA from the cytoplasmic side of the lysosomal membrane (2) and gets acetylated at an active site histidine (3). A conformational change allows for the transfer of the acetyl group into the lysosome (4). Once heparan sulfate interacts with the active site, the terminal glucosamine residue of heparan sulfate (GlcN) acquires the acetyl group (5), thus forming N-acetylglucosamine (6); B. HGSNAT (1) catalyzes its reaction via a random ternary order complex (2), so that the process requires only one step and no direct acetylation of the enzyme as an intermediate (3). Adopted from (Fedele 2015)

Another alternative model proposes that the enzyme functions via a random-order ternary-complex mechanism where there is no specific order for binding the two substrates i.e., there is no release of an initial product (CoA), nor production of an acetylated enzyme intermediate before binding the second substrate (GlcNH$_2$) and generation of the final product (N-acetylglucosamine) (Figure 1.5 (B)) (Meikle et al. 1995; Fan et al. 2011).

HGSNAT primary sequence encodes a polypeptide of 635 amino acids with 11 transmembrane domains (TMDs) and a non-cleavable signal peptide with 42 N-terminal amino acids (Figure 1.6). The signal peptide integrates into cellular membranes as a single chain precursor of 70.5 kDa, which in turn post-translationally cleaved to a 27 kDa α-chain (Lys43 to Asn144) and a 44 kDa β-chain (Gly145 to Ile635) upon arrival in the lysosome (Fan et al. 2006, 2011; Hrebícek et al. 2006; Durand et al. 2010). The subunits are held together by disulfide bonds in between two cysteine residues (Cys123 and Cys434) in the lysosomal luminal loops of the enzyme (Durand et al. 2010).
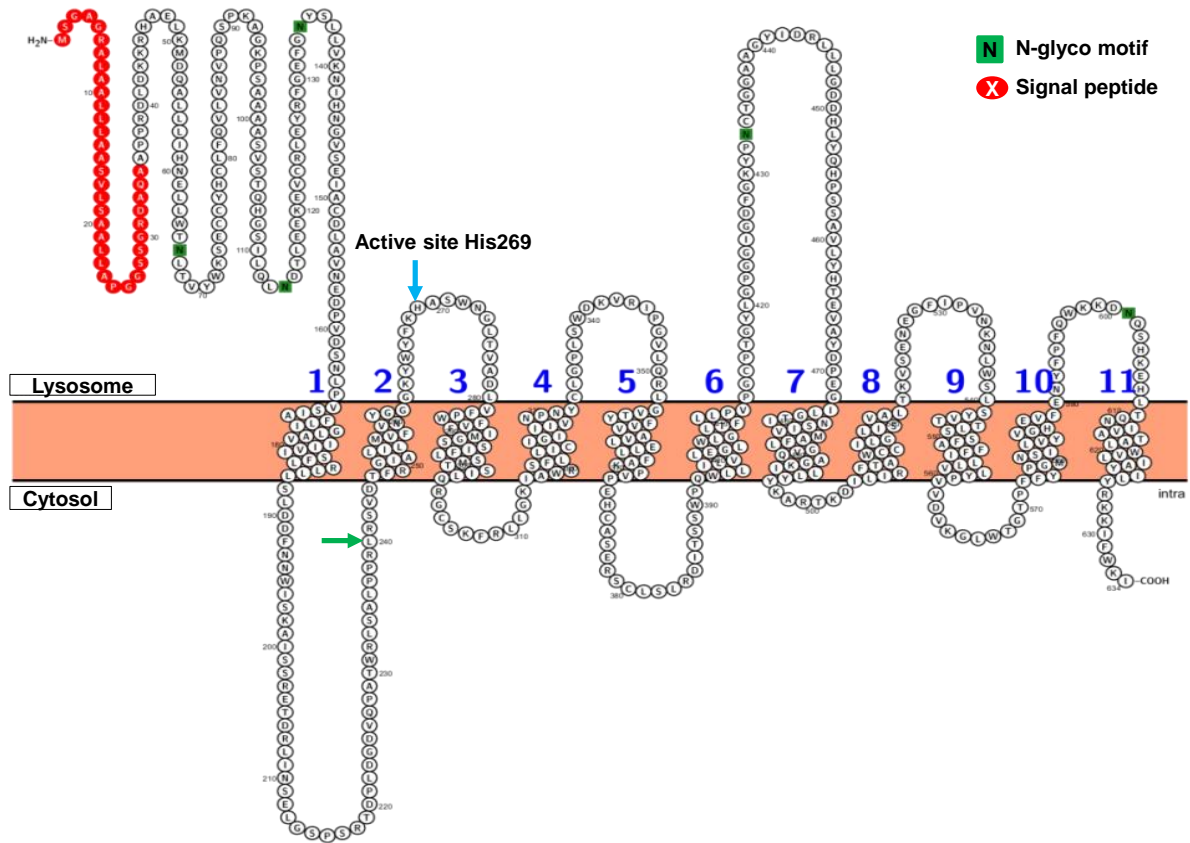
Figure 1.6: Topology of HGSNAT with its 11 transmembrane domains (TMDs).
Red circles showing the signal peptide; N: N-linked glycosylation motifs; Green arrow shows the last codon of the epitope for HGSNAT antibody, cyan arrow shows the active site histidine (His269) which transfers the acetyl group. Transmembrane segments predicted by the Protter tool (Omasits et al. 2014).

The nucleophile residue His269 is the active site of the protein which accesses the substrate acetyl-CoA in the cytoplasm and transfers it to the terminal glucosamine on heparan sulfate (Durand et al. 2010). There are also 5 consensus sites for Asn-linked glycosylation. The first cytosolic loop contains an intracellular localization signal in the form of a dileucine repeat-based lysosomal integral membrane-targeting motif [DE]XXXL[LI] (ETDRLI209). However, the role of putative tyrosine- and/or di-leucine repeat-based targeting motifs within a sequence towards the carboxy-terminal cytoplasmic tail (TALWVLIAYILYRKK630) remains controversial (Durand et al. 2010; Fan et al. 2011). The purification of the enzyme is very difficult because it is hydrophobic and the activity of HGSNAT is dependent on protein concentration, i.e the activity is completely lost if the purity of the enzyme increases (Ausseil et al. 2006). To date, 65 mutations in the HGSNAT gene have been identified out of which 28 are missense mutations. Expression studies with the mutations showed that these proteins are misfolded and retained in ER instead of lysosomal targeting having

negligible HGSNAT residual activity (Feldhammer et al. 2009a; Fedele and Hopwood 2010).

1.10 Heparan sulfate metabolism

Heparin/heparan sulfate glycosaminoglycan (HS-GAG) is a member of the glycosaminoglycan family consisting of a variably sulfated repeating disaccharide unit, glucuronic acid (GlcA) or iduronic acid (IdoA) linked to N-acetylglucosamine (GlcNAc). Two or three HS-GAG chains covalently attached to a core protein on the cell surface or in the extracellular matrix forming heparan sulfate proteoglycans (HSPGs)/ mucopolysaccharides (Sasisekharan and Venkataraman 2000; Esko et al. 2009). HSPGs can regulate many biological processes such as development, angiogenesis, blood coagulation, tissue repair and tumor metastasis (Stringer and Gallagher 1997; Tumova et al. 2000). Degradation of HS-GAG is required to maintain a natural turnover of GAGs. Defects in the degradative enzymes result in lysosomal storage diseases, where GAGs build up rather than being broken down and having pathological effects (Ballabio and Gieselmann 2009).
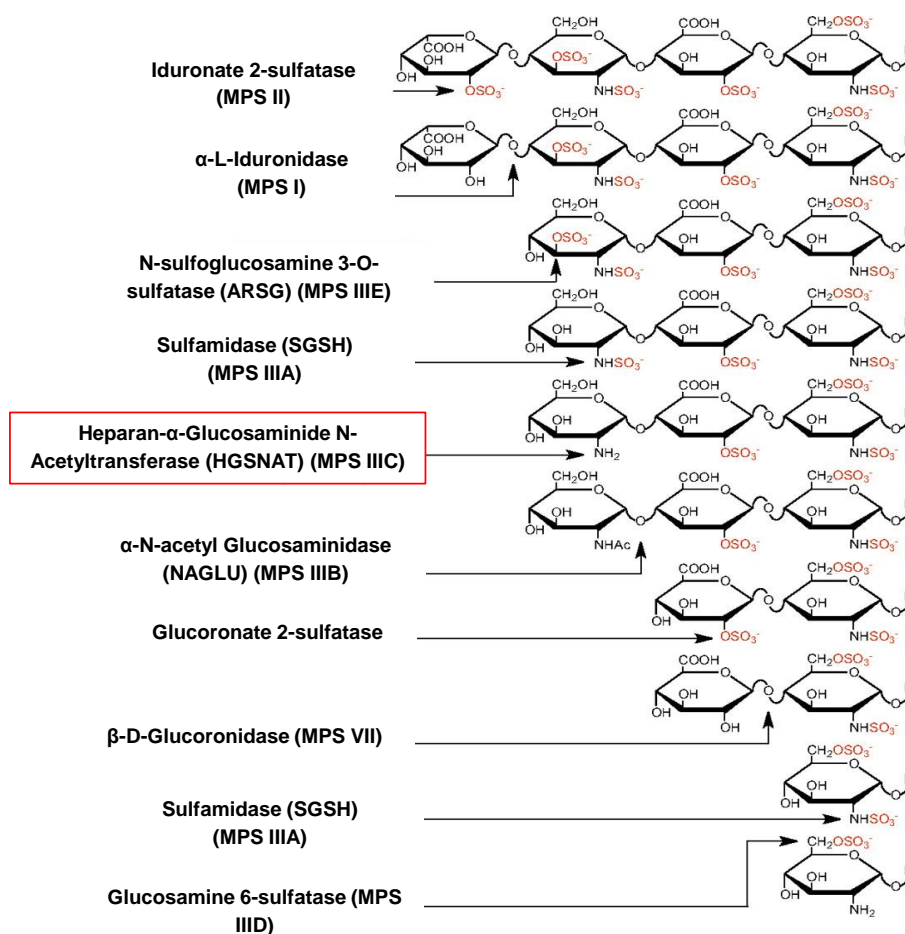
Figure 1.7: Heparan sulfate catabolism.
The scheme illustrates all nine different enzymes required for sequential catabolism of heparan sulfate (Adapted from (Kowalewski et al. 2012)).

By endocytosis, cell-surface HSPGs reach endosomal compartment for their degradation, where heparanase cleaves off their side chains into oligosaccharides followed by their terminal degradation on lysosomes (Yogalingam and Hopwood 2001). Inside the lysosome, sequential action of several lysosomal enzymes e.g three glycosidases, one acetyltransferase, and three known sulfatases, degrade HS into glucuronic and L-iduronic acid (Figure 1.7) (Dorfman and Matalon 1976). Mutation or deficiency of any of these enzymes leads to lysosomal accumulation and urinary excretion of heparan sulfate resulting clinical onset of Sanfilippo syndrome (Andrade et al. 2015).

Moreover, after endocytosis, glypican-1 (Gpc-1), the most widely expressed HSPGs, and its heparan sulfate chains get deaminatively degraded by nitric oxides (NO) in the endosomal compartment (Fransson et al. 2004). This truncated Gpc-1 recycles back to the cell surface via the Golgi, where new HS chains are added (Edgren et al. 1997).

1.11 Aim of this study

The aim of this dissertation was to investigate the molecular events leading to the pathology of NPC disease by a mass spectrometry-based phosphoproteomics study. In order to make the aim of the dissertation achievable, the following specific objectives were settled:

I. Generation of NPC disease phenotype in a cell culture model by applying the inhibitor U18666A and confirmation of it by filipin staining.

II. Quantitative proteomics and phosphoproteomics analysis of lysosomes from SILAC labeled and U18666A treated MEFs.

III. Bioinformatics analysis, correlation of protein level and phosphorylation site data and selection of candidate protein.

IV. Follow up on identified candidate proteins using biochemistry and molecular biology methods:
   - Confirmation of protein abundance and localization using antibodies
   - Overexpression and knockdown studies
   - Detailed biochemical characterization of target proteins and their regulatory mechanisms

2. Materials and Methods

2.1 Materials

2.1.1 Chemicals/ Solutions/ Buffers

All mass spectrometry-related chemicals were HPLC or MS grade.

Chemicals / Solutions	Provider
$CaCl_2$, Ethanol, PMSF, KCl, NaF, Sodium orthovanadate	AppliChem GmbH (Darmstadt, Germany)
Clarity™ Western ECL substrate, Sodium dodecyl sulfate (SDS)	Bio-Rad Laboratories, Inc. (Hercules, CA, USA)
Acetic acid, ACN, FA, TFA	Biosolve (Valkenswaard, The Netherlands)
L-Arginine: HCL ($^{13}C_6$, 99%; $^{15}N_4$, 99%), L-Lysine: 2HCL ($^{13}C_6$, 99%; $^{15}N_2$, 99%)	Cambridge Isotope Laboratories, Inc. (Tewksbury, MA, USA)
Lysozyme, NaOH, TEMED	Carl Roth GmbH (Karlsruhe, Germany)
MycTrap agarose	ChromoTek GmbH (Planegg-Martinsried, Germany)
Roche cOmplete protease inhibitor cocktail	F. Hoffmann-La Roche AG (Basel, Switzerland)
Nonidet P-40, Triton X-100	Fluka Chemie Ag (Buchs, Switzerland)
Dulbecco's modified eagle's medium (DMEM): -L-glutamine, -pyruvate, DPBS, FCS, Penicillin, Streptomycin	Invitrogen Gibco (Paisley, United Kingdom)
Glycerol, HEPES (2,4 Hydroxyethyl piperazineethanesulfonic acid), Methanol, U18666A	Merck KGaA (Darmstadt, Germany)
Amylose resin	New England BioLabs (Ipswich, MA, USA)
2-Mercaptoethanol, Acrylamide, Ammonium persulfate (APS), BSA, Bromophenol blue, Citric acid monohydrate, DMSO, DTT, EDTA, Filipin, G418 solution, Glycolic acid, $MgCl_2$, Paraformaldehyde, Phosphoric acid, NaCl, Sodium pyrophosphate, Urea	Sigma-Aldrich (St. Louis, MO, USA)
DMEM for SILAC, Fluoromount-G™ with DAPI, IPTG (Isopropyl β-D-1- thiogalactopyranoside), TurboFect transfection reagent	Thermo Fisher Scientific (Waltham, MA, USA)

2.1.2 Equipment and consumables

Equipment and consumables	Provider
Empore C18 disks	3M Corporation (St. Paul, MN, USA)
Biometra T3 thermal cycler	Analytik Jena AG (Jena, Germany)
Branson Ultrasonics™ Sonifier S-250A with 2 mm tip; Ultrasonic water bath 2510	Branson Ultrasonics (Danbury, CT, USA)
Axiovert 200M microscope	Carl Zeiss AG (Oberkochen, Germany)
Axygen reaction tubes maximum recovery (1.5 mL)	Corning GmbH (Kaiserslautern, Germany)
ReproSil-Pur 120 C18 AQ	Dr. Maisch GmbH (Ammerbuch-Entringen, Germany)
Thermomixer Comfort	Eppendorf AG (Hamburg, Germany)
HisTrap^{TM} HP column (1 ml)	GE Healthcare Europe GmbH (Freiburg, Germany)
pH-Meter, Calimatic 761	Knick Elektronische Messgeräte GmbH & Co. (Berlin, Germany)
Vacuum centrifuge RVC 2-18	Martin Christ Gefriertrocknungsanlagen GmbH (Osterode am Harz, Germany)
QIAprep spin miniprep kit; QIAquick PCR purification kit	Qiagen Inc. (Hilden, Germany)
Sachtopore NP 5 µm 300 Å TiO_2 bulk material	Sachtleben Chemie GmbH (Duisburg, Germany)
Microplate reader GENios	Tecan (Männerdorf, Switzerland)
Liquid chromatograph EASY-nLC 1000; Mass spectrometer LTQ Orbitrap Velos/ Orbitrap Fusion Lumos; PureLink™ HiPure plasmid midiprep kit; PureLink™ Quick gel extraction kit; Centrifuge Labofuge 400; 37 °C Heraeus incubator	Thermo Fisher Scientific (Waltham, MA, USA)
FUSION SOLO 4M system	Vilber Lourmat (Eberhardzell, Germany)
PerfectBlue™ semi-Dry elektroblotter	VWR (Darmstadt, Germany)
Oasis HLB cartridge (10 mg/400 mg)	Waters Corporation (Milford, MA, USA)

2.1.3 Primers and oligonucleotides

All primers and oligonucleotides were synthesized by biomers.net GmbH (Ulm, Germany), except crRNA for CRISPR/Ca9 method by Dharmacon (GE Healthcare Dharmacon, Freiburg, Germany)

Primer internal name	Target gene	Sequence	Purpose	Tm* ºC
HGSNAT 211S-E,C-1	HGSNAT	ACTGATCGCCTCATCAATGAGG AGCTGGGATCTCCCAGC	SDM**	NA***
HGSNAT 211S-E,NC-2	HGSNAT	GCTGGGAGATCCCAGCTCCTCA TTGATGAGGCGATCAGT	SDM	NA
HGSNAT 211S-A,C-3	HGSNAT	ACTGATCGCCTCATCAATGCTG AGCTGGGATCTCCCAGC	SDM	NA
HGSNAT 211S-A,NC-4	HGSNAT	GCTGGGAGATCCCAGCTCAGC ATTGATGAGGCGATCAGT	SDM	NA
HGSNAT 215S-E,C-5	HGSNAT	ATCAATTCTGAGCTGGGAGAGC CCAGCAGGACAGACCCT	SDM	NA
HGSNAT 215S-E,NC-6	HGSNAT	AGGGTCTGTCCTGCTGGGCTCT CCCAGCTCAGAATTGAT	SDM	NA
HGSNAT 215S-A,C-7	HGSNAT	ATCAATTCTGAGCTGGGAGCTC CCAGCAGGACAGACCCT	SDM	NA
HGSNAT 215S-A,NC-8	HGSNAT	AGGGTCTGTCCTGCTGGGAGC TCCCAGCTCAGAATTGAT	SDM	NA
HGSNAT_FA_Fw-13	HGSNAT	AAAGAATTCATGAGCGGGGCG GG	Cloning	65
HGSNAT_FA_Re-14	HGSNAT	AAAAAGCTTGAGGCGGGGCGG CAG	Cloning	65
HGSNAT_Fwd_HindIII-20	HGSNAT	AAAAAGCTTATGAGCGGGGCG GG	Cloning	65
HGSNAT_Rev_NotI-21	HGSNAT	AAAGCGGCCGCTCACTTGTCGT CATGTCTTTGTAGTCCAGATCC TCTTCTGAGATGAGTTTTTGTTC GATTTTCCAAAAAATCTTCTTTC TATAGAG	Cloning	65
HGSNAT_FP_Fwd_EcoRI-22	HGSNAT	AAAGAATTCATGAGCGGGGCG GG	Cloning	65
HGSNAT_FP_Rev_HindIII-23	HGSNAT	AAAAAGCTTTCAGAGGCGGGG CGGCAG	Cloning	65
HGSNAT_SDM_2E_C_28	HGSNAT	ATCAATGAGGAGCTGGGAGAG CCCAGCAGGACAGACCCT	SDM	NA
HGSNAT_SDM_2E_NC_29	HGSNAT	AGGGTCTGTCCTGCTGGGCTCT CCCAGCTCCTCATTGAT	SDM	NA
HGSNAT_SDM_2A_C_30	HGSNAT	ATCAATGCTGAGCTGGGAGCTC CCAGCAGGACAGACCCT	SDM	NA

Primer internal name	Target gene	Sequence	Purpose	Tm* ºC
HGSNAT_SDM_2A_NC_31	HGSNAT	AGGGTCTGTCCTGCTGGGAGCTCCCAGCTCAGCATTGAT	SDM	NA
HGSNAT_EA_C_32	HGSNAT	ATCAATGAGGAGCTGGGAGCTCCCAGCAGGACAGACCCT	SDM	NA
HGSNAT_EA_NC_33	HGSNAT	AGGGTCTGTCCTGCTGGGAGCTCCCAGCTCCTCATTGAT	SDM	NA
HGSNAT_AE_C_34	HGSNAT	ATCAATGCTGAGCTGGGAGAGCCCAGCAGGACAGACCCT	SDM	NA
HGSNAT_AE-NC_35	HGSNAT	AGGGTCTGTCCTGCTGGGCTCTCCCAGCTCAGCATTGAT	SDM	NA
80_HGSNAT_DD_C	HGSNAT	GAAACTGATCGCCTCATCAATGATGAGCTGGGAGATCCCAGCAGGACAGACCCTCTC	SDM	NA
81_HGSNAT_DD_NC	HGSNAT	GAGAGGGTCTGTCCTGCTGGGATCTCCCAGCTCATCATTGATGAGGCGATCAGTTTC	SDM	NA
82_HGSNAT_rev_pTRE3G	HGSNAT	TTTACGCGTTCACTTGTCGTCATCGTCTTTGTAGTC	Cloning	65
85_Scarb2_rev_myc	Scarb2	AAAGCGGCCGCTCACAGATCCTCTTCTGAGATGAGTTTTTGTTCGGTTCGTATGAGGGGTGCTC	Cloning	64.6
crRNA_HGSNAT	HGSNAT	GTGATGTTCAGCCAGCAACG	CRISPR/Cas9***	NA
HGSNAT_fwd_crispr	HGSNAT	AGGAGAATCGCTTGACCCCAGGAGGTGGAGGTT	CRISPR/Cas9	64.43
HGSNAT_rev_crispr	HGSNAT	TCTTCCAGACTGGAGAGGCTGCAGAGTTTACTTTTACA	CRISPR/Cas9	65.03

*: Melting temperature; **: Site-directed mutagenesis; ***: Clustered regularly interspaced short palindromic repeat (CRISPR); ****: Not applicable. The annealing temperature for all site-directed mutagenesis PCRs was set at 60ºC.

2.1.4 Antibodies

Target name	Manufacturer	Host species	Product No.	Application	Dilution
CTSD	Santa Cruz Biotechnology, Inc. (Dallas, Texas, USA)	Goat	sc6486	WB	1:1000
FLAG	Sigma-Aldrich (St. Louis, MO, USA)	Mouse	F1804	WB and ICC**	WB (1:5000) ICC (1:400)
HGSNAT	In-house/Pineda (Berlin, Germany)	Rabbit	-	WB	1:500
Myc	Abcam (Cambridge, United Kingdom)	Rabbit	ab9106	ICC	1:400

Target name	Manufacturer	Host species	Product No.	Application	Dilution
LAMP2	The Developmental Studies Hybridoma Bank (DSHB) (Iowa, IA, USA)	Rat	ABL-93	ICC	1:100
β-Actin	Sigma-Aldrich (St. Louis, MO, USA)	Mouse	A5316	WB	1:5000

*: Western blot, **: Immunocytochemistry

2.1.5 Enzymes

Enzyme	Provider
Conventional and/or Fast Digest restriction enzymes (HindIII, NotI, XbaI, BamHI, BglII, SalI, and MluI)	Thermo Fisher Scientific (Waltham, MA, USA)
Phusion high-fidelity DNA polymerase	Thermo Fisher Scientific (Waltham, MA, USA)
rLys-C	Promega (Madison, WI, USA)
Sequencing grade modified trypsin, porcine	Promega (Madison, WI, USA)
T4 DNA ligase	Thermo Fisher Scientific (Waltham, MA, USA)
Trypsin-EDTA solution 0.05%	Invitrogen Gibco (Paisley, United Kingdom)

2.1.6 Software

Software	Provider
AxioVision SE64 Rel.4.9.1	Carl Zeiss AG (Oberkochen, Germany)
FusionCapt Advance Solo 4 16.15	VilberLourmat (Eberhardzell, Germany)
GraphPad Prism 6.07	GraphPad Software Inc. (San Diego, CA, USA)
MaxQuant 1.5.2.8	Max Planck Institute of Biochemistry (Planegg, Germany)
Office Professional Plus 2010	Microsoft Corporation (Redmond, WA, USA)
Perseus 1.6.5.0	Max Planck Institute of Biochemistry (Planegg, Germany)
Thermo Xcalibur 2.2	Thermo Fisher Scientific (Waltham, MA, USA)

2.2 Methods

2.2.1 Eukaryotic cell line culture methods

2.2.1.1 Cell maintenance

Mouse Embryonic Fibroblasts (MEFs) and Human Embryonic Kidney (HEK293) cells were cultured in Dulbecco's Modified Eagle's medium (DMEM) containing 10 % fetal calf serum (FCS), 100 IU/ml Penicillin, 100 µg/ml Streptomycin and 2 mM L-Glutamine. In case of SILAC cells, other than SILAC DMEM with 10 % dialyzed SILAC FCS, light labeled natural amino acids (Lys0, Arg0) or medium (Lys4: 4,4,5,5,-D4; Arg6: $^{13}C_6$) or heavy labeled lysine and arginine (Lys8: $^{13}C_6 \bullet ^{15}N_2$; Arg10: $^{13}C_6 \bullet ^{15}N_4$) was provided. For labeling, cells were grown on an appropriate SILAC medium for a total of five cell doublings, where ~97 % of proteins got isotope incorporation (Ong and Mann 2006). Cells were passaged by using 0.05 % (w/v) trypsin solution. All cells were grown at 37 °C incubator 5 % CO_2. For long term storage, cells were kept in 10 % DMSO containing medium in a liquid nitrogen tank.

2.2.1.2 Coating of cell culture dishes

Before seeding HEK293 cells, culture dishes were coated with poly-L-Lysine (PLL). 5 ml of 100 µg/ml PLL solution was added to each 10 cm dish and incubated for at least 15 minutes in a humidified 37 °C incubator. After removing the solution, dishes were washed three times with 1x PBS. Coated dishes were either used immediately or stored at 4°C up to 2 weeks. Diluted PLL solutions were re-used up to five times.

2.2.1.3 Lysosome enrichment

Enrichment of lysosomes was performed as mentioned (Thelen et al. 2017). Briefly, $3x10^6$ MEF cells per10 cm dish were seeded with DMEM plus 2.5 % FCS and 48 hours later, 10 % magnetite medium was added to the cells. In the case of HEK293 cells, $5x10^6$ cells were seeded on PLL coated 10 cm dish directly in 10 % magnetite medium. After 24 hours of magnetite treatment, the medium was removed, cells were washed thrice with pre-warmed 1x PBS and supplemented with DMEM for a chase of 24-36 hours. Before harvesting, the medium was removed and cells were washed twice with ice-cold 1x PBS. Subsequently, 2 ml ice-cold Buffer A (250 mM Sucrose, 15 mM KCl, 1.5 mM MgAc, 10 mM HEPES pH 7.4, 1mM $CaCl_2$, 1 mM $MgCl_2$, 1 mM DTT, 1x Protease Inhibitor Cocktail (PIC), 1 mM NaF, 1 mM β-Glycerophosphate, 1 mM Sodium Orthovanadate, 10 mM Sodium Pyrophosphate) was added per plate, and scraped using a cell scraper. Then cells were homogenized 25 strokes using a 15 ml douncer and post-nuclear supernatant (PNS) was collected by centrifugation at

500 g for 10 minutes. After a second homogenization, both PNS were pooled and passed by gravity through a magnetite LS column placed in a MidiMACS Separator. The column was washed with 3 ml of Buffer A, removed immediately from the stand and lysosomes were eluted from the column with 2x 500 µl of Buffer A using a plunger. After pelleting intact lysosomes at 20,000 g for 30 min at 4 °C, the enrichment efficiency was evaluated by the β-hexosaminidase assay (2.2.4.1).

2.2.1.4 Cell harvest

Cells were harvested either by trypsinization or using a cell scraper. In either case, cells were washed with prewarmed PBS. For trypsinization, 1 ml 0.05 % (w/v) trypsin solution was added per 10 cm dish, incubated for 5 minutes at 37 °C and the reaction was stopped by adding 10 ml prewarmed medium. Subsequently, cells were transferred to a 15 ml tube, centrifuged for 5 minutes at 500 g at RT, the supernatant was removed, and pelleted cells were either cultured further or cryopreserved with 10 % DMSO. For harvesting cells using a cell scraper, after washing them with prewarmed PBS, culture dishes were placed on ice and were washed with ice-cold PBS. Subsequently, 1 ml of ice-cold PBS was added, and cells were mechanically dissociated from the culture dish using a cell scraper and transferred into a 1.5 ml tube and pelleted down at 500 g for 5 minutes at 4° C. Cell pellets were snap-frozen using liquid nitrogen and stored at −80 °C until further analysis.

2.2.1.5 Transient transfection of eukaryotic cells

TurboFect® transfection reagent was used in order to transfect eukaryotic cells according to the manufacturer's instructions with some modifications. Briefly, cells were seeded 24 hours prior to transfection to reach a confluency of 70-90 % at the time of transfection. For a 10 cm cell culture dish, 10 µg of purified plasmid DNA and 20 µl of TurboFect reagent were added to a tube containing 1 ml of serum-free medium and vortexed immediately. After 20 minutes of incubation at RT, the mixture was added dropwise to the cells and the cell culture dish was placed in a 37 °C incubator after mild shaking.

2.2.1.6 CRISPR/Cas9 gene editing

Prior to the start of CRISPR/Cas9 gene editing, puromycin concentration for HEK293 cells was determined which kills all untransfected cells. The day before transfection, HEK293 cells were seeded in a 6-well plate with complete media. For transfection, crRNA, tracrRNA and Cas9 plasmids were resuspended in nuclease-free 10 mM Tris pH 7.5 buffer with a final concentration of 10 µM, 10 µM and 100 ng/µl respectively

and mixed thoroughly on an orbital shaker at RT for 30 minutes. During transfection, crRNA, tracrRNA, Cas9 plasmid and Dharmafect Duo® solution (ordered from Dharmacon) were added to complete DMEM media at a final concentration of 50 nM, 50 nM, 2 ng/µl and 6 ng/µl respectively with a total volume of 500 µl. Subsequently, cells were transfected with this solution, while for control purposes, one well was supplied only with Cas9 plasmid and one remained untreated (negative control). 48 hours post-transfection, media was replaced with 2 µg/ml puromycin containing media in all wells except the negative control. 48 hours later transfected cells were split into 1:5 dilution and cultured for another 24 hours with puromycin supplementation. After limiting dilution selected single cells were cultured until 90 % confluency and split into three 96 well plates. Two plates containing cells were frozen with DMEM, 20 % FCS and 10 % DMSO media at -80 °C for later use, where cells in another plate were lysed with 30 µL lysis buffer (1 mM $CaCl_2$, 3 mM $MgCl_2$, 1 mM EDTA, 0.2 mg/mL Proteinase K, 1 % Triton X-100 and 10 mM TRIS-HCl) to isolate genomic DNA and performed 2 consecutive PCR for initial screening: one with primers binding to gene of interest and one with barcode primers binding to the forward primers having unique hangover for each clone. After confirmation from both PCR, all lysates were pooled in one Eppendorf tube and sent for sequencing.

2.2.1.7   Stable transfection using the piggyBac transposon system

HGSNAT KO HEK293 cells were converted to HGSNAT KO Tet-On cells by piggyBac transposon system (Li et al. 2013). Before stable transfection, the concentration of G418 antibiotic needed to kill all untransfected cells was determined. Cells were co-transfected with a rtTA gene containing piggyBac transposon plasmid and transposase enzyme-containing piggyBac helper plasmid using a transposon to helper plasmid ratio of 5:1, 2.5:1 and 1:1. TurboFect transfection reagent was used for transfecting the cells. 48 hours post-transfection, 300 µg/ml G418 antibiotics was added to the transfected cells and continued the culture for 2 weeks.

2.2.1.8   Microscopy related methods

2.2.1.8.1   Filipin staining

MEF cells were seeded in a 24-well plate at a density of 10,000 cells per 12 mm coverslip. Next day, cells were treated for 24 hours with either 3 µg/ml of U18666A or DMSO and subsequently, 24 hours post-treatment, cells were PBS washed, fixed for 10 minutes with 4 % paraformaldehyde (PFA), and washed thrice with PBS. After permeabilizing cells with 0.1 % Triton X-100 for 30 minutes, they were PBS washed

and stained for 2 hours with 125 µg/ml filipin in PBS/10% FCS in the dark at RT. Then after washing cells thrice with PBS, they were mounted using DAPI-fluoromount-G. Finally, stained cells were observed using an Axiovert 200M microscope equipped with an AxioCamMR3 using an objective with 20x magnification. Images were recorded with an exposure time of 350 milliseconds.

2.2.1.8.2 Immunocytochemistry staining and microscopy of cells

For immunocytochemistry studies, NIH3T3 cells were cultured on 12 mm coverslips, washed with PBS and fixed with 4% PFA for 10 minutes at RT. Then, coverslips were washed thrice with PBS, transferred to a wet chamber, permeabilized with 0.1% Triton X-100 in PBS for 10 min and blocked using 10 % FCS in PBS for 1 hour at RT. Subsequently, cells were incubated with respective primary antibody (see 2.1.4) for overnight at 4°C, washed thrice with PBS and incubated with secondary antibody for 1 hour in the dark at RT. Finally, the coverslips were washed twice with PBS, once with water and mounted using DAPI-fluoromount-G. Immunocytochemistry images were acquired using the Axiovert 200M equipped with an AxioCamMR3 with the following filter sets: filter set 38 (excitation BP 470/40, beam splitter FT 495, emission BP 525/50) for green fluorescence; filter set 43 HE (excitation BP 550/25 (HE), emission FT 570 (HE), beam splitter BP 605/70 (HE)) for red fluorescence; filter set 49 (DAPI) (excitation G 365, beam splitter FT 395, emission BP 445/50).

2.2.2 Molecular cloning-related methods

2.2.2.1 Polymerase chain reaction (PCR)

PCR reaction was carried on using Phusion High-Fidelity (HF) DNA Polymerase enzyme. For one PCR reaction following components were added to a PCR tube: 10 µl of 5x Phusion HF buffer, 5 µl of 2 mM dNTPs, 0.5 µM each of the forward and reverse primers, 20 ng of template DNA, 5 µl of DMSO, 0.5 µl of the Phusion HF DNA Polymerase (1U), and ddH$_2$O to a total volume of 50 µl. After vortexing and brief centrifugation PCR tube was placed in a thermocycler with the following PCR program: 30 seconds at 98 °C for initial denaturation and 35 cycles of 10 seconds at 98 °C for denaturation, 30 seconds at 'X' °C for annealing, and 30 seconds/kb at 72 °C for extension following with 10 minutes final extension at 72 °C. The primer annealing temperatures ('X') were adjusted to 3 °C above their calculated melting temperatures (Tm) (see 2.1.3)

2.2.2.2 Site-directed mutagenesis (SDM)

By using PCR, candidate phosphosites were mutated to either glutamic acid (E) or aspartic acid (D) and alanine (A). Mutation points were placed in the middle of the primers and the melting temperature was designed to be more than 80 °C. For one mutagenesis reaction following components were added to a PCR tube: 10 µl of 5x Phusion HF buffer, 5 µl of 10 mM dNTPs, 20 pmol each of the forward and reverse primers, 50 ng of the template plasmid, 5 µl of DMSO, 0.5 µl of the Phusion HF DNA Polymerase (1U), and ddH$_2$O to a total volume of 50 µl. Then the reaction tube was placed in a preheated thermocycler, with the following program: 1 cycle at 95 °C for 1 minute, 16 cycles of 50 seconds at 95 °C for denaturation, 50 seconds at 60 °C for annealing, and 30 seconds/kb at 72 °C for extension, with a final 10 minutes extension at 72 °C. Next, 1 µl of DpnI was added directly to the PCR products and incubated at 37 °C for 4 hour to digest the parental DNA template.

2.2.2.3 PCR product purification

PCR products were purified using QIAquick PCR Purification Kit. Binding buffer (PB buffer) was added to the PCR products at a 3:1 ratio (PB buffer: PCR product), the mixture was transferred to a fresh spin column, centrifuged at 13,000 g for 1 minute, flow-through (FT) was collected and reloaded onto the column twice. Then captured DNA was washed using wash buffer (PE buffer), FT was collected, reloaded onto the column, 2nd FT was discarded and centrifuged for 1 minute at 13,000 g. After changing the collection tube 35 µl of ddH$_2$O was added, incubated 5 minutes at RT and finally dissolved DNA was collected by centrifugation at 13,000 g for 1 minute.

2.2.2.4 Restriction digestion and gel extraction

1 µg of DNA were incubated with 10 units of appropriate restriction endonuclease(s) in the corresponding buffer at 37 °C for 1 hour. During double digestion, according to the manufacturer's recommendations, a reaction buffer compatible to both enzymes was chosen. Afterward, to confirm the size of digested products or to purify the DNA fragment from the gel, digested DNA was separated using agarose gel electrophoresis. Then nucleotides were extracted from the gel cut using the PureLink® Quick Gel Extraction Kit. The excised gel was dissolved in gel solubilization buffer (1:3), incubated 15 minutes at 55 °C, transferred to a column and centrifuged for 1 minute at 13,000 g. Subsequently, 500 µl washing buffer was added, centrifuged twice for 1 minute at 12,000 g, exchanged the collection tube and finally, nucleotides were eluted with 35 µl ddH$_2$O.

2.2.2.5 Ligation, bacterial transformation, and culture

To ligate the insert DNA with the vector, 20 ng of the linear vector were added with a 5-fold excess of the insert DNA. Then 1 unit of T4 DNA ligase was added to the mixture and incubated overnight at 16 °C. Transformation of bacterial was carried out using heat shock method, where purified or ligated plasmid was added to the chemically competent *E. coli* XL-1 Blue cells, and incubated for 30 minutes on ice. Then cells were heated for 45 seconds at 42 °C and shocked for 2 minutes on ice. Subsequently to each sample, 900 µl of prewarmed lysogeny broth (LB) medium (without antibiotic) was added, shaked for 1 hour with 250 rpm at 37°C in the orbital shaker. Then, 100 µl of the suspension was spread on LB-agarose plates containing the vector-specific antibiotics and incubated overnight at 37 °C. Next day, after picking single clones, they were inoculated into 5 ml LB-medium containing the respective antibiotic and incubated overnight at 37 °C in an orbital shaker at 250 rpm. Then, the plasmid DNA was isolated using QIAprep spin miniprep kit according to the manufacturer's protocol and after sequencing verification, respective plasmid DNA was purified using the PureLink HiPure plasmid midiprep kit according to the manufacturer's protocol.

2.2.3 Protein-related methods

2.2.3.1 Cell lysis

Cell lysates or enriched lysosomes were lysed using the lysis buffer consisting of 0.25 % Triton X-100 and Buffer A (see 2.2.1.3) including 1x protease inhibitor cocktail. The lysis buffer was added to the sample with 10:1 ratio (lysis buffer: cell pellet (v/v)) and the pellet was dissolved by pipetting. Afterward, using ultrasonic water bath lysates were sonicated for 10 minutes. Then they were freeze-thawed 6 times using liquid N_2 and water bath.

2.2.3.2 Determination of protein concentration

The protein concentration was measured based on the Lowry assay (Lowry et al. 1951) using the Bio-Rad DC assay kit. In brief, 5 µl of diluted samples were pipetted in triplicates in a 96 well plate and calibrated with a BSA serial dilution standard. Then reagent A' was prepared by adding 20 µl reagent S to 1 ml of reagent A followed by adding 25 µl of reagent A' and 200 µl of reagent B subsequently to the samples and BSA standard. After a 15 minutes incubation at RT, the absorbance of the samples was measured at 750 nm using TECAN microplate reader. Finally, the

total protein concentration of each sample was calculated based on the regression equation of the BSA standard curve.

2.2.3.3 SDS-PAGE and Western blotting

10 % SDS-PAGE gels were prepared with a combination of running gels (375 mM Tris/HCl, pH 8.8; 0.1 % (w/v) SDS; 10 % acrylamide; 0.1 % APS and 0.1 % TEMED) and stacking gels (75 mM Tris/HCl, pH 6.8; 0.05 % (w/v) SDS; 2.5 % acrylamide; 0.05 % APS and 0.1 % TEMED) using BioRad Mini-PROTEAN Tetra Cell Casting Stand. The gels were placed into the electrophoretic chamber with 1x SDS running buffer (190 mM Glycine, 0.1 % (w/v) SDS and 25 mM Tris/HCl). 10-50 µg of the protein samples were incubated for 10 minutes at 40 °C with a final concentration of 1 x Laemmlli buffer (1 % (v/v) β-mercaptoethanol, 2 % (w/v) SDS, 10 % (v/v) Glycerol, 4 % (w/v) Bromophenol blue, 60 mM Tris/HCl) (Laemmli 1970). Proteins were size separated at 120 V for ~1.5 hours and transferred to PVDF membrane (activated with methanol) for 1 hour with 200 mA using the Perfect Blue Semi-Dry Electro Blotter. After blocking the membranes for 1 hour with 5 % non-fat dry milk in TBS-T (10 mM Tris, 150 mM NaCl and 0.05 % Tween-20) they were washed thrice with TBS-T and incubated overnight at 4 °C with the primary antibody. The appropriate secondary antibody was applied for 1 hour after washing the membranes thrice with TBS-T. The protein expression signals were detected using the enhanced chemiluminescence (ECL) kit, visualized with the FUSION SOLO 4M System and illustrated and/or analyzed by the FusionCapt Advance software.

2.2.3.4 Recombinant protein expression and cell lysis

E. coli XL1 Blue cells containing the plasmid of interest were grown overnight in an orbital shaker with 250 rpm at 37 °C in LB-medium containing 100 µg/ml Ampicillin until mid-log phase, indicated by an OD595 of 0.6 to 1.0. Protein was expressed by adding 0.1 mM IPTG (final conc.), incubated for 4 hours at 37 °C and centrifuged at 4,000 rpm for 15 minutes. Then the bacterial pellet was resuspended using the lysis buffer [1x PBS, 200 µg/ml lysozyme, 1x protease inhibitor cocktail and 1 mM phenylmethylsulfonyl fluoride (PMSF)], incubated 30 minutes on ice and lysed by 2 times freezing and thawing followed by sonication using Branson sonifier 250 with a setting of 30 seconds, duty cycle: '70' and output control: '7'. Following centrifugation of the lysate at 20,000 rpm for 30 minutes, the supernatant containing protein extracts was sterile filtered using 0.45 µm filter and transferred into a new plastic tube.

2.2.3.5 Affinity purification and dialysis of MBP-HGSNAT fusion protein

The MBP-HGSNAT fusion protein was purified by using amylose resin beads followed by Ni^{2+}-NTA IMAC column. Briefly, the amylose column was made with 6 ml amylose resin and equilibrated with five column volumes of 1x PBS. Filtered protein extract was passed over the column with gravity flow followed by washing the column with 10 ml 1x PBS. Bound proteins were eluted from amylose column with 50 mM maltose in PBS. Eluate was pressed through a syringe coupled to the HisTrap HP^{TM} column followed by washing the column with 10 ml 1x PBS. Finally, MBP-HGSNAT fusion protein was eluted from the column using 150 mM imidazole in PBS in several fractions each with 500 μl volume. After combining all fractions proteins were dialyzed using SERVAPOR dialysis tubing (MWCO: 12,000-14,000 Da, RC, diameter: 16 mm) against 1x PBS at 4 °C overnight. Next day, volume was reduced by using Amicon Ultra-15 Centrifuge filter unit with centrifuging at 4,000 g, 15 minutes, 4 °C. Finally, protein concentration was measured.

2.2.3.6 Purification of HGSNAT anti-serum

A rabbit was immunized intradermally with 300-500 μg affinity-purified MBP-HGSNAT fusion protein (FP) and boosted every 20 days by subcutaneous injections. Serum was collected consecutively 60, 90 and 120 days post-injection and antibodies targeting HGSNAT were purified from the serum using HGSNAT protein immobilized in AminoLink Plus resin column. To that extent, AminoLink Plus resin slurry was equilibrated thrice with 5 ml coupling buffer (0.1 M sodium phosphate, 0.15 M sodium chloride, pH 7.2) and the concentrated MBP-HGSNAT FP was added to the equilibrated resin beads. Cyanoborohydride solution (5 M $NaCNBH_3$ in 1 M NaOH) was added to the reaction slurry in a fume hood to a final concentration of 50 mM and incubated overnight at 4 °C with constant rotation. The supernatant containing unbound proteins was collected by centrifugation at 500 g for 1 minute. Protein assay was performed with the supernatant to observe coupling efficiency. After washing the resin with coupling buffer, remaining active sites of AminoLink plus resin was blocked using quenching buffer (1 M Tris/HCl, pH 7.4) and cyanoborohydride solution. After collecting the supernatant beads were washed thrice with coupling buffer. Then the column was used for purification of HGSNAT antibody. First, the column was equilibrated with 1x PBS, then the diluted anti-serum (1:5 with 1x PBS) was passed over the column by gravity flow and flow-through was collected in fractions. After washing the column thrice with binding buffer, bound antibody was eluted with elution

buffer (0.2 M glycine/HCl, pH 2.5) and the eluate was immediately neutralized with neutralization buffer (1 M Tris/HCl, pH 9). Finally, all eluate fractions were pooled, the buffer was exchanged to 1x PBS using Amicon Ultra-15 centrifugal units (30 kDa cutoff) and protein content was measured by DC protein assay (see 2.2.3.2)

2.2.4 Enzymatic assays

2.2.4.1 β-hexosaminidase assay

To perform β-hexosaminidase assay (Wendeler and Sandhoff 2009), 25 µl samples from each fraction were loaded onto 4 wells of a 96 well plate, 8 µl 10% Triton X-100 was added to two of them followed by addition of 50 µl β-hexosaminidase substrate solution (100 mM sodium citrate, 0.2 % BSA, 10 mM p-Nitrophenyl-N-acetyl-2-β-D-Glucosaminide pH 4.6) to all samples. After incubating 15 minutes at 37 °C, 200 µl stop solution (0.4 M Glycine/NaOH pH 10.4) was added to each sample. Finally, absorbance was measured at 405 nm using a Tecan microplate reader. Total β-hexosaminidase activity (mU) was calculated using Beer-Lambert Law.

2.2.4.2 HGSNAT activity assay

HGSNAT activity assay was performed as reported earlier (Voznyi et al. 1993) with some modifications. In the case of HGSNAT overexpression, β-N-Acetyl hexosaminidase purified from *Talaromyces flavus* (Slámová et al. 2012) was added to the reaction mixture. To measure the activity, respective amount of enriched lysosomes (100 ng to 1 µg from overexpressed cells and 10 µg from wild type cells), 4 mM acetyl CoA, 1 mM fluorogenic substrate 4-methylumbelliferyl β-D-glucosaminide (4MU-βGlcNH$_2$) in McIlvain's buffer (100 mM sodium citrate, 200 mM sodium phosphate, pH 5.7), 1 unit β-N-Acetyl hexosaminidase and 0.2 % BSA to adjust the volume to 30 µl were mixed and incubated at 30 °C from 30 minutes (overexpressed samples) to 17 hours (WT samples). The reaction was stopped by adding 200 µl of 0.5 M Na$_2$CO$_3$/NaHCO$_3$ pH 10.7. Fluorescence of 4-methylumbellifereyl (4MU) was measured by Tecan microplate reader using 360 nm excitation and 460 nm emission wavelengths. Activities were calculated using methylumbelliferyl standard solution and expressed as mU/mg.

2.2.5 Mass spectrometry-based techniques

2.2.5.1 Chloroform-methanol precipitation

For protein precipitation, 1 ml ice-cold chloroform-methanol (2:1, v/v) was added to each sample and incubated on ice for 1 hour. Samples were centrifuged at 14,000 rpm for 30 minutes at 4 °C. All liquids were removed from the tube without disturbing

the middle protein layer. Proteins were washed by adding 1 ml methanol per tube, centrifuged at 14,000 rpm for 30 minutes at 4 °C. Finally, supernatants were discarded and protein pellets were air-dried.

2.2.5.2 In-solution digestion

Precipitated protein samples were resuspended in 8 M urea in 0.1 M TEAB, incubated at RT with 800 rpm for 45 minutes, reduced by adding 5 mM dithiothreitol (DTT) at 56 °C, 800 rpm for 25 minutes and alkylated with 20 mM acrylamide for 30 minutes in the dark. After quenching the reaction with 5 mM DTT at RT, 800 rpm for 15 minutes, urea concentration was reduced to 4 M, rLys-C was added at an enzyme to protein ratio of 1:100 and incubated overnight at 37 °C. Next day, urea concentration was further reduced to 1.6 M, trypsin was added at an enzyme-to-substrate ratio of 1:100 and incubated 10 hours at 37 °C. Digested peptides were then acidified to 1 % acetic acid (final concentration) and desalted.

2.2.5.3 In-gel digestion

For in-gel digestion, proteins were denatured and reduced by 1x Laemmli buffer for 10 minutes at 95 °C, alkylated by 20 mM acrylamide for 30 minutes at RT and size separated in a 10 % SDS gel. Proteins were stained overnight using Coomassie brilliant blue and de-stained using ddH_2O. Gel bands were excised into 1 mm cubes, washed several times with 30 % ACN, 0.07 M ammonium bicarbonate (NH_4HCO_3) each time for 30 minutes at 25 °C, 800 rpm to remove all Coomassie stains, dehydrated using 100 % ACN for 15 minutes at 800 rpm, 25 °C and dried by vacuum centrifugation. 1 µg diluted trypsin (1:1 with 0.1 M NH_4HCO_3) was added to each sample, thereby added more 0.1 M NH_4HCO_3 until all gel pieces were merged and incubated overnight at 37 °C. Next day, the supernatants containing peptides were collected to a new tube and remaining peptides from gel pieces were extracted by sequential addition of 50 % ACN, 0.1% TFA; 0.1 M NH_4HCO_3 and ACN. Finally, extracted peptides were dried using vacuum centrifugation.

2.2.5.4 Peptide purification by solid-phase extraction (SPE)

For samples containing a larger number of peptides (>10 µg), 10 mg Oasis HLB cartridges were applied. The cartridges were equilibrated using four times 1 ml 70 % ACN, 0.5 % AcOH followed by four times 1 ml 0.5 % AcOH. Acidified samples were loaded three times to the cartridge followed by washing the cartridge 10 times with 1 ml of 0.5 % AcOH. Peptides were eluted by 300 µl of 30 % ACN, 0.5 % AcOH, 300 µl

50 % ACN, 0.5 % AcOH and 500 µl 70 % ACN, 0.5 % AcOH and after combining all eluate fractions, they were dried using a vacuum centrifuge.

2.2.5.5 Enrichment of phosphopeptides

Phosphopeptide enrichment was performed as mentioned elsewhere (Villén and Gygi 2008). Briefly, dried peptides were resuspended in 500 µl 80 % ACN, 5 % TFA, 1 M glycolic acid and added to TiO_2 beads at a ratio of 1:6 (peptides to beads, w/w) followed by incubation at RT, 1200 rpm for 15 minutes. TiO_2 beads were then centrifuged down at 13,000 g for 1 minute and the supernatant was discarded. TiO_2 beads were transferred to new tubes using 1 ml of 80 % ACN, 1 % TFA, centrifuged, and the supernatant discarded. The beads were washed with 1 ml 20 % ACN, 0.1 % TFA, the supernatant was discarded after the centrifugation and dried by vacuum centrifugation. Phosphopeptides were eluted from TiO_2 beads by adding 200 µl 1% NH_4OH and incubation at RT, 1200 rpm for 15 minutes. Beads were pelleted by centrifugation and supernatants were transferred to new tubes and acidified using 10 µl formic acid (FA). Phosphopeptides were dried using a vacuum centrifuge and re-solubilized in 500 µl 70 % ACN, 0.1 % TFA. TiO_2 beads were again added to each sample at a ratio of 1:6 (peptides to beads) and then the beads were washed with 50 % ACN, 0.1 % TFA. Finally, phosphopeptides were eluted by incubation with 200 µl 1 % NH_4OH at RT, 1200 rpm for 15 minutes, acidified using 10 µl of FA.

2.2.5.6 StageTip desalting of peptides

Samples containing less than 10 µg of peptides were purified using StageTips (Rappsilber et al. 2003). Briefly, 4-6 discs of C18 Empore extraction discs were placed in 200 µl micropipette tips and equilibrated by sequential application of 20 µl of MeOH, 80 % ACN, 0.5 % AcOH and 0.5 % AcOH with centrifugation at RT, 2500 x g for 1 minute. In the case of phosphopeptides, StageTips were also primed with 20 µl of 50 mM citrate prior to loading the samples (Winter et al. 2009). Dried peptides were resuspended in 20 µl of 5 % FA, 5 % ACN solution, sonicated briefly in an ultrasonic water bath and then 80 µl of water was added to dilute ACN amount. Resuspended peptides were applied to the StageTips, washed using 100 µl of 0.5 % AcOH and peptides were eluted using 2x 20 µL 80 % ACN, 0.5 % AcOH. In between each step, StageTips were centrifuged at RT, 2,500 x g for 1 min. Eluted peptides were dried using a vacuum centrifuge.

2.2.5.7 SILAC CoIP using Anti-HGSNAT antibody

For co-immunoprecipitation (CoIP), SPIONs treated cells were dounced in Buffer A (see 2.2.1.3). After douncing 25 strokes, post-nuclear supernatant (PNS) was collected in a new falcon by centrifuging at 500 g for 10 minutes at 4 °C. Douncing was repeated with 2 ml Buffer A per sample and PNS were pooled. The protein content of the PNS sample was measured by lysing an aliquot of PNS using 1 % NP-40, 1 x PIC. The same amount of PNS from heavy and light cells was mixed at this stage and then split into two half where one part was used for lysosome enrichment using magnetic stand (see 2.2.1.3) and another part as PNS sample. After lysosome enrichment, both the eluate and PNS were lysed with 1 % NP-40, 1x PIC, 1 mM $MgCl_2$ and 1 mM $CaCl_2$. Then 50 µl of affinity-purified rabbit anti-HGSNAT antibody (1.5 µg/µl) was added to each sample (lysosome and PNS) and incubated overnight with mild shaking at 4 °C. Next, 70 µl protein A beads (crosslinked in 6 % beaded agarose with a capacity of 12-19 mg human IgG per ml resin) was washed thrice with ice-cold Buffer A and then added to each sample. After incubating for 3 hours at 4 °C with mild shaking, beads were centrifuged at 100 g for 2 minutes, unbound proteins were collected from the supernatant and beads were washed thrice with 2 ml Buffer A. Finally, protein complexes were eluted from the beads using 100 µl 2x Laemmli buffer by heating at 40 °C for 10 minutes and centrifuging at 2,000 g for 2 minutes. The eluates of both samples were alkylated using 20 mM acrylamide at RT for 30 minutes in dark. Samples were then loaded into 10 % SDS-PAGE gel and subjected to in-gel digestion (2.2.5.3) and desalted (2.2.5.6).

2.2.5.8 SILAC CoIP for Myc-tagged proteins

Myc-Trap A beads were used for precipitating Myc-tagged HGSNAT interacting proteins. $1x10^7$ overexpressed cells were resuspended in 1 ml pre-warmed paraformaldehyde (PFA) solution (0.5 % PFA in PBS), incubated at RT for 15 minutes with constant agitation and centrifuged at 2,000 g for 5 minutes. The cross-linking reaction was stopped by adding 1 ml of ice-cold stop solution (1.25 M glycine in PBS) per $1x10^7$ cells. Afterward, cell pellets were lysed with 300 µl of lysis buffer (40 mM HEPES pH 7.4, 1 % Triton X-100, 10 mM β-glycerophosphate, 10 mM sodium pyrophosphate, 2.5 mM $MgCl_2$, 1 mM PMSF and 2x of EDTA-free protease inhibitor cocktail) and then placed on ice for 30 minutes with extensive pipetting every 10 min. After centrifuging for 10 minutes at 20,000 g, 4°C, supernatants were transferred to new tubes and protein concentration was measured. 3 mg lysate was

diluted to 500 µl using dilution buffer (10 mM Tris/HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 1 mM PMSF and 2x of protease inhibitor cocktail), mixed the diluted lysate with prewashed beads and incubated at 4 °C with rotation for 3 hours. After washing beads thrice with the dilution buffer, immunocomplexes were dissociated from the beads using 100 µl of 2x Laemmli buffer. Following incubation for 10 min at 40° C, beads were centrifuged at 2500 g for 2 minutes and the supernatant was transferred into the new tube as an eluate. At this stage eluate from heavy, medium and light conditions were mixed, alkylated using 20 mM acrylamide for 30 minutes at RT in dark and 50 µl of the fresh eluate was loaded to 10 % SDS-PAGE gels, run until 1 cm in the separating gel and subjected to in-gel digestion.

2.2.6 Mass spectrometry analysis of samples

2.2.6.1 LC-MS/MS data acquisition

LC-MS/MS analyses of peptides were carried out by nano-UHPLC-MS/MS using a Dionex Ultimate 3000 system coupled to an Orbitrap Fusion Lumos mass spectrometer. Analytical columns were produced in house as follows: using a P-2000 laser puller, spray tips were generated with fused silica capillaries with an outer diameter of 360 µm, an inner diameter of 100 µm and packed with 1.9 µm Reprosil AQ C18 reverse-phase particles. Peptides were resuspended in 5 % ACN, 5 % FA and loaded directly on the analytical column at a flow rate of 600 nl/min 100 % solvent A (0.1 % FA in water). Before loading phosphopeptides, columns were equilibrated thrice with 50 mM citrate (Winter et al. 2009). Separation was performed at a flow rate of 300 nl/min with 90- or 240-min linear gradients from 5 % - 35 % solvent B (95 % ACN, 0.1 % FA). Survey spectra were acquired in the Orbitrap mass analyzer with a mass range of m/z 375-1575 at a resolution of 60,000. MS/MS fragmentation was performed on the 10 most abundant ions using higher-energy collision dissociation (HCD) followed by data acquisition in the Orbitrap at a resolution of 30,000. The precursor isolation width was set to 1.6 m/z in the quadrupole with automatic gain control (AGC) of 4×10^5 and 5×10^5 for MS1 and MS2 scans, respectively. Singly charged ions and such with unassigned charge states were excluded from MS/MS fragmentation. Dynamic exclusion was enabled with a repeat count set to 1 and repeat duration to 45 sec for 90 min gradient and 120 sec for 240 min gradient for reducing repeated fragmentation reactions of the same peptides.

2.2.6.2 Mass spectrometry data analysis

2.2.6.2.1 Protein identification and quantification using MaxQuant

The generated mass spectrometry raw files were processed in MaxQuant (version 1.5.2.8) with the Andromeda search engine (Cox and Mann 2008). The database used was either UniProt human (release 2015_04, 68485 entries) or UniProt mouse (release 2015_05, 53297 entries). The enzyme specificity was set to trypsin with maximum two missed cleavages allowed and quantification was performed using default parameters for 2 or 3 states SILAC [light (K0, R0), medium (K4, R6) and heavy (K8, R10)]. The fixed modification was set to propionamide at cysteine residues. Variable modifications were as follows: acetylation at protein N-termini, oxidation at methionine, carbamylation at lysine and peptide N-termini (in the case of urea-based lysis buffer), as well as phosphorylation at serine, threonine, and tyrosine.

2.2.7 Bioinformatics and statistical analysis

MaxQuant data were further analyzed by Perseus toolbox, version 1.6.5.0 (Tyanova et al. 2016) and Microsoft Excel (2010). In order to identify enriched gene ontology (GO), PANTHER (Protein Analysis Through Evolutionary Relationships (http://www.pantherdb.org)) (Ashburner et al. 2000) was used. For network analysis, STRING (https://string-db.org/) was used (Szklarczyk et al. 2019). Statistical processing and visualization of data were performed with GraphPad Prism, version 6.07 (www.graphpad.com).

3. Results

3.1 Proteomics and phosphoproteomics study on NPC

In order to investigate phosphoproteome changes due to altered cholesterol metabolism, U18666A was used to induce the phenotype of Niemann Pick Disease Type C (NPC) in mouse embryonic fibroblasts (MEFs). Lysosomes were enriched from these cells and changes in protein phosphorylation were detected by mass spectrometry-based proteomics and phosphoproteomics (Figure 3.1).
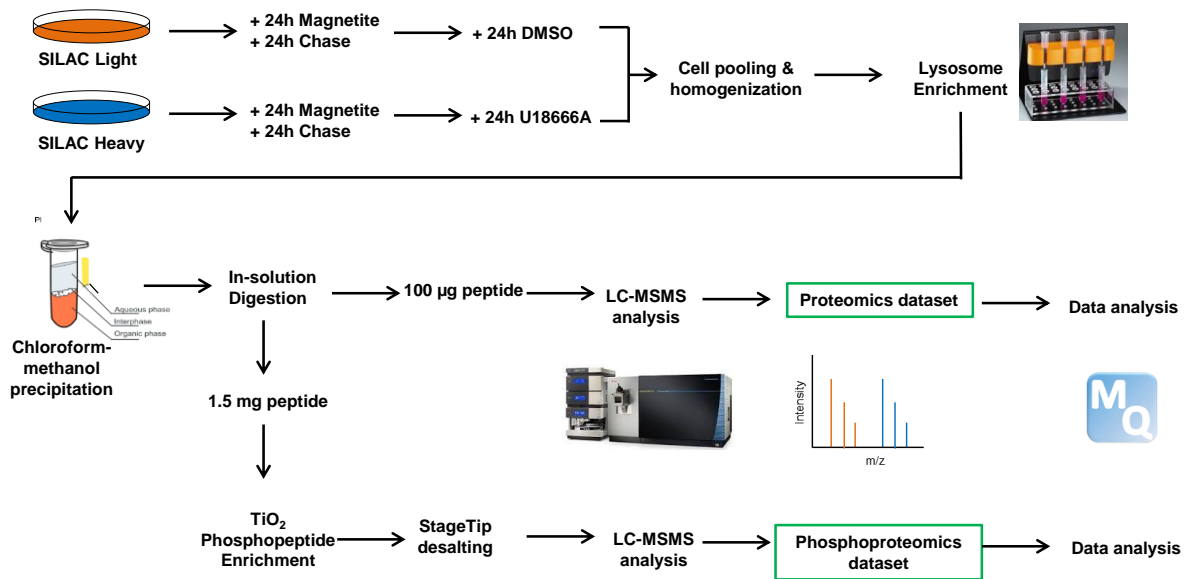
Figure 3.1: Schematic workflow of proteomic and phosphoproteomics studies.
Lysosomes were enriched from U18666A and DMSO treated SILAC heavy and light labeled MEF cells. After chloroform-methanol precipitation, proteins were digested in solution with rLysC and trypsin. 100 µg peptides were analyzed by LC-MS/MS for proteomic study. 1.5 mg peptides were enriched for phosphopeptide using TiO_2, desalted using StageTips and analyzed by LC-MS/MS. All proteomic and phosphoproteomic data were analyzed by MaxQuant.

3.1.1 Induction of NPC by U18666A in MEFs

NPC phenotype was induced by treating SILAC-labelled MEF cells with U18666A, which is an inhibitor of intracellular cholesterol trafficking (Cenedella 2009). In due course, intra-lysosomal cholesterol accumulation was confirmed by filipin staining, which can bind to and fluorescently stain unesterified cholesterol (Arthur et al. 2011). MEF cells were cultured on coverslips in a 24 well plate and treated with either 3 µg/ml U18666A or DMSO (control) for 24 hours. Subsequently, 24 hours post-treatment, cells were fixed and stained with or without filipin. Figure 3.2 shows blue-stained accumulated cholesterol inside the lysosome and confirms the induction of NPC.
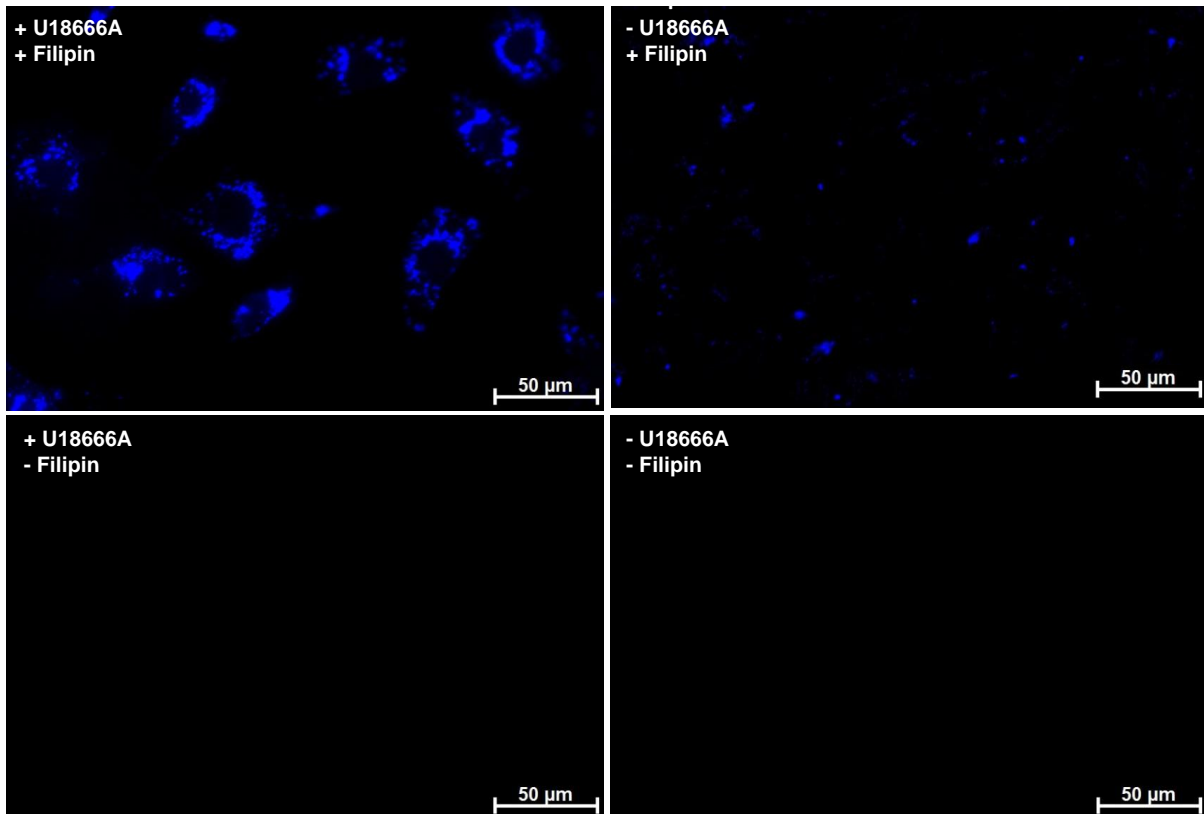
Figure 3.2: Filipin staining of unesterified cholesterol in MEFs.
MEFs were cultured on coverslips and treated with either U18666A or DMSO for 24 hours. After fixing and staining with or without filipin, imaging was performed with the Axiovert 200 M microscope, using an objective with 20x magnification. Images were recorded with an exposure time of 350 ms. Cholesterol accumulation is indicated by blue staining.

3.1.2 Enrichment of lysosomes from MEFs

To reduce sample complexity for mass spectrometry analysis it was decided to enrich lysosomes by SPIONs technique (Walker and Lloyd-Evans 2015). Therefore, SILAC light (Lys0, Arg0) or heavy (Lys8, Arg10) labeled MEF cells were cultured in 6x 10-cm dishes and 3 days later magnetite media was added on each plate. After 24 hours of pulse, the media were removed and cells were washed thrice with PBS. New media were added per plate with U18666A on heavy labeled cells and DMSO on light labeled cells. 24 h post-treatment cells were harvested, pooled, dounce homogenized and lysosome enrichment was performed using magnetic chromatography (see details in 2.2.1.3). To evaluate the enrichment efficiency and measure the number of intact lysosomes in the individual fractions, β-hexosaminidase enzyme assay (Wendeler and Sandhoff 2009) was performed.
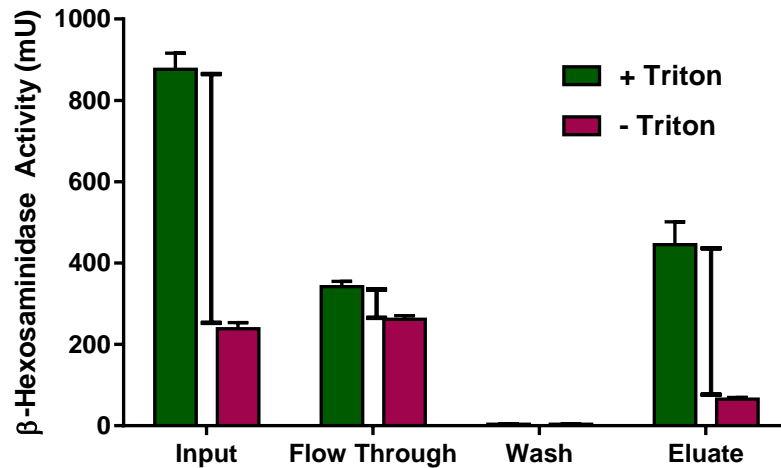
Figure 3.3: β-hexosaminidase assay showing the efficiency of lysosome enrichment from MEFs.
The total activity of β-hexosaminidase was determined in each fraction by enzymatic activity assay with (green bar) and without (pink bar) Triton X-100. Lines next to green bars indicate the number of intact lysosomes. Total enzymatic activity is shown in milliunits (mU). n=3, shown are mean values ± SEM.

The difference between with and without Triton X-100 shows the activity that is present in intact lysosomes and is therefore available for purification. The β-hexosaminidase assay shows, 73 % of the lysosomes were intact in the input (shown by capped line) from which 60 % were recovered in the eluate (whole green bar) (Figure 3.3). The eluate fraction from the magnetic column contained a total of 85 % intact lysosomes (shown by capped line) which were further used for mass spectrometry sample preparation.

3.1.3 Proteomic analysis of lysosomes from U18666A treated MEF cells

The lysosomal eluate was precipitated using chloroform-methanol precipitation method and lysed using urea lysis buffer. After reduction and alkylation precipitated proteins were digested overnight with rLys-C followed by trypsin digestion for 8 hours. Digested peptides were then desalted using Oasis HLB cartridges. Subsequently, 100 µg of the peptide was separated from the Oasis eluate and consequently, 1 µg was analyzed on an Orbitrap Fusion Lumos mass spectrometer. The raw data were further processed by MaxQuant (Figure 3.1). From 3 independent biological replicates, 3,592 proteins and 38,726 peptides were detected.

Output data from MaxQuant were further analyzed using the computational tool 'Perseus' (Tyanova et al. 2016). Correlation between three biological replicates was determined by Pearson's correlation coefficient (r) (Puth et al. 2014), where scatter plot of Figure 3.4 (A) shows 'r' as 0.821, 0.82 and 0.789 between replicate 1 and 2, 1 and 3, and 2 and 3, respectively. These relatively high positive Pearson's correlation

coefficients indicate high reproducibility of replicates. Venn diagram (Figure 3.4 (B)) shows, 78 % proteins were common in all three biological replicates and only 5.9 % proteins were detected in one replicate.
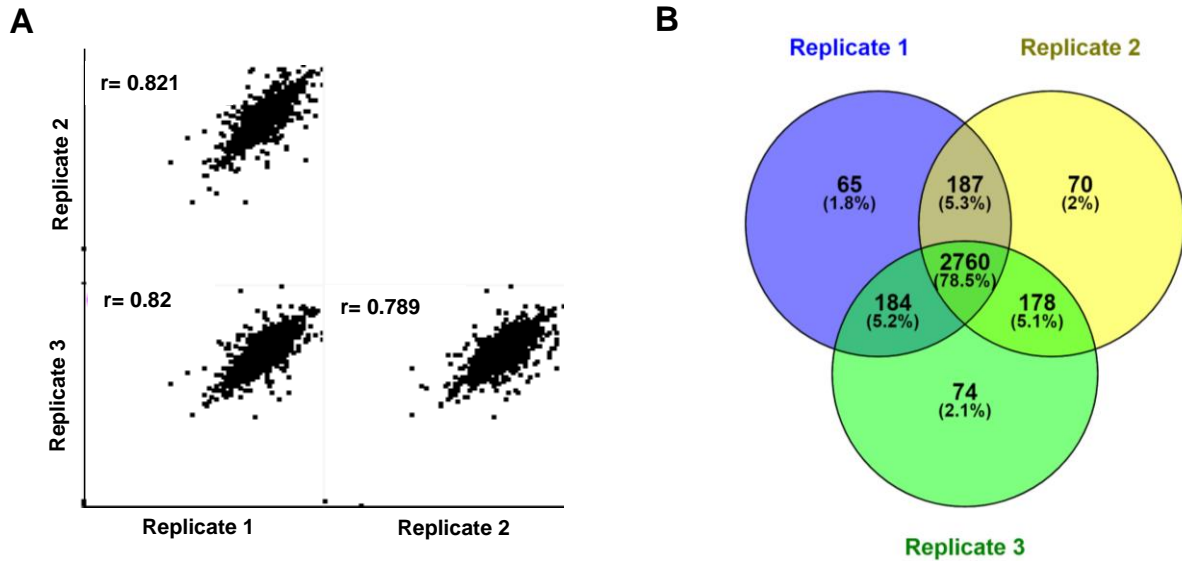
Figure 3.4: Total proteins from 3 biological replicates of proteomics dataset.
A. Proteomics data was analyzed by MaxQuant and 'ProteinGroups' output of MaxQuant was further assessed by Perseus. Correlation between the three biological replicates was determined by Pearson's correlation coefficient (r). B. Venn diagram showing overlap between the total proteins found in three biological replicates. Venn diagrams were created online using open-access tool Venny (Oliveros 2007).

Next, in order to evaluate the distribution of data in each replicate, normalized SILAC H/L ratios obtained from MaxQuant were log (base 2) converted and distribution was assessed using Perseus software. In Figure 3.5, all replicates are showing a normal distribution of SILAC H/L ratios towards '0'.
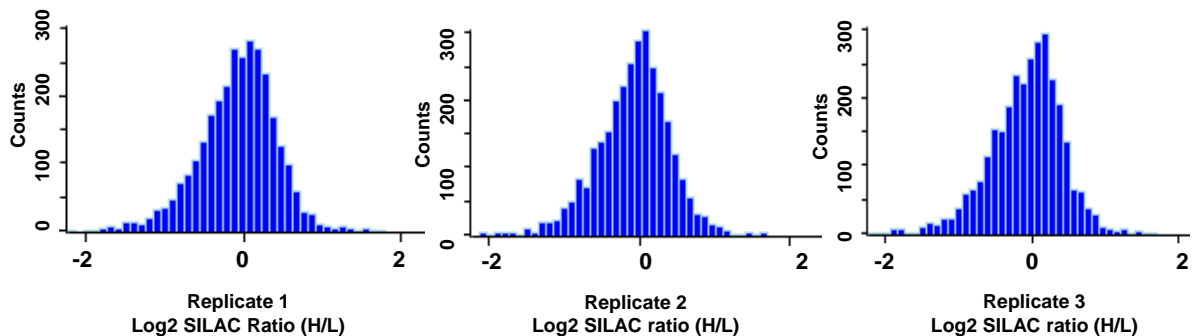
Figure 3.5: Distribution of SILAC H/L ratios of proteins.
Normalized ratios acquired from the MaxQuant were converted to log2 values and from these values' distribution of SILAC ratios for each replicate was assessed using Perseus. Y-axis (counts) shows the number of counts in each bin in the histogram.

The protein lists from the MaxQuant analyses were evaluated further for lysosomal and lysosomal-associated proteins. Therefore, two known lists of lysosomal proteins were used where one list contains in-house curated lysosomal proteins (189

proteins) and another one contains Panther gene ontology annotated lysosomal proteins (411 proteins) (see Appendix 6.3). In the following, these lists will be called as the curated list and Panther list. From this analysis, 142 lysosomal proteins were detected based on the curated list and 210 lysosomal proteins were found based on the Panther list. Next, the reproducibility of lysosome enrichment was compared between each replicate. Figure 3.6 (A) shows, according to curated list 89 % of the lysosomal proteins were common in all 3 replicates, whereas according to Panther list 83 % lysosomal proteins were found common in 3 biological replicates. Only 4.2 % and 4.6 % unique lysosomal proteins were found in single replicates based on the curated and Panther list respectively.
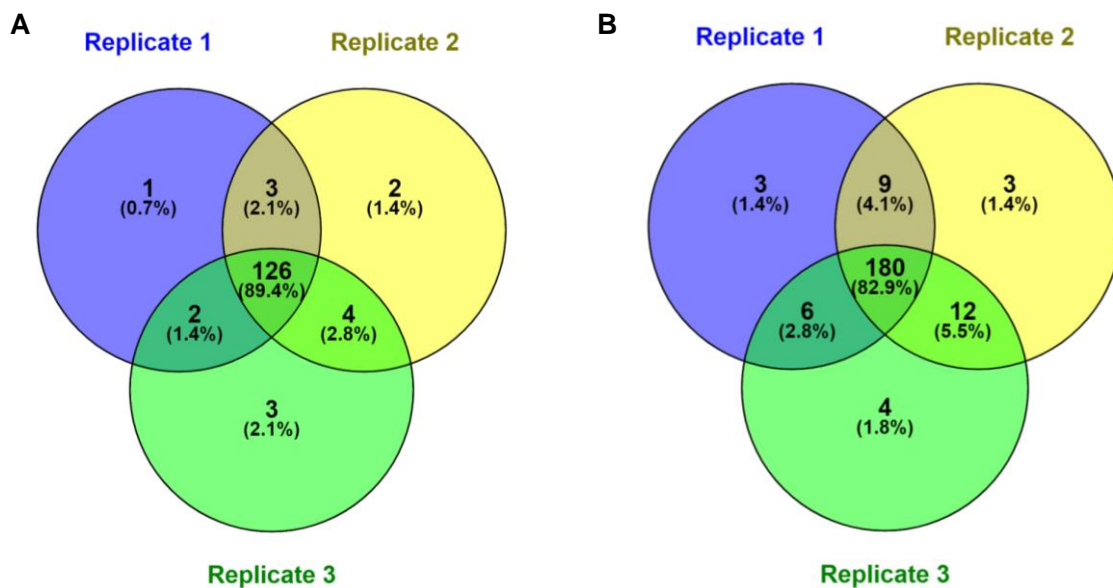
Figure 3.6: Lysosomal proteins detected in 3 biological replicates of proteomics datasets.
A. MaxQuant output 'ProteinGroups' file was compared with an in-house curated lysosomal protein list by Venny. B. MaxQuant output 'ProteinGroups' file was compared with Panther gene ontology lysosomal protein list by Venny (Oliveros 2007).

In the next step, proteomics dataset was analyzed for differentially expressed proteins upon U18666A treatment. For that, proteins which are common in at least two biological replicates were accepted with a quantified p-value below 0.05. Changes in protein expression level will be addressed as regulated in the rest of the text. For up- and downregulation, SILAC ratios +1 and -1 (log2 values) were applied as cutoff values. Based on this setup, 31 upregulated and 115 downregulated proteins were detected (Figure 3.7). From 31 up-regulated proteins, 3 proteins were known to be lysosomal, 7 proteins have a direct link with cholesterol metabolism (Table 3.1) and from 115 down-regulated proteins, 5 were found as lysosomal proteins (Table 3.2). Furthermore, a total of 118 kinases and 62 phosphatases were

identified, where only 1 kinase (Pnkp) and 2 phosphatases (Alpl, Pnkp) were found upregulated while 4 kinases (Dclk1, Jak1, Mapk15, and Tgfbr1) and 3 phosphatases (Ptprf, Ptprg, and Ppp5c) were found to be downregulated.
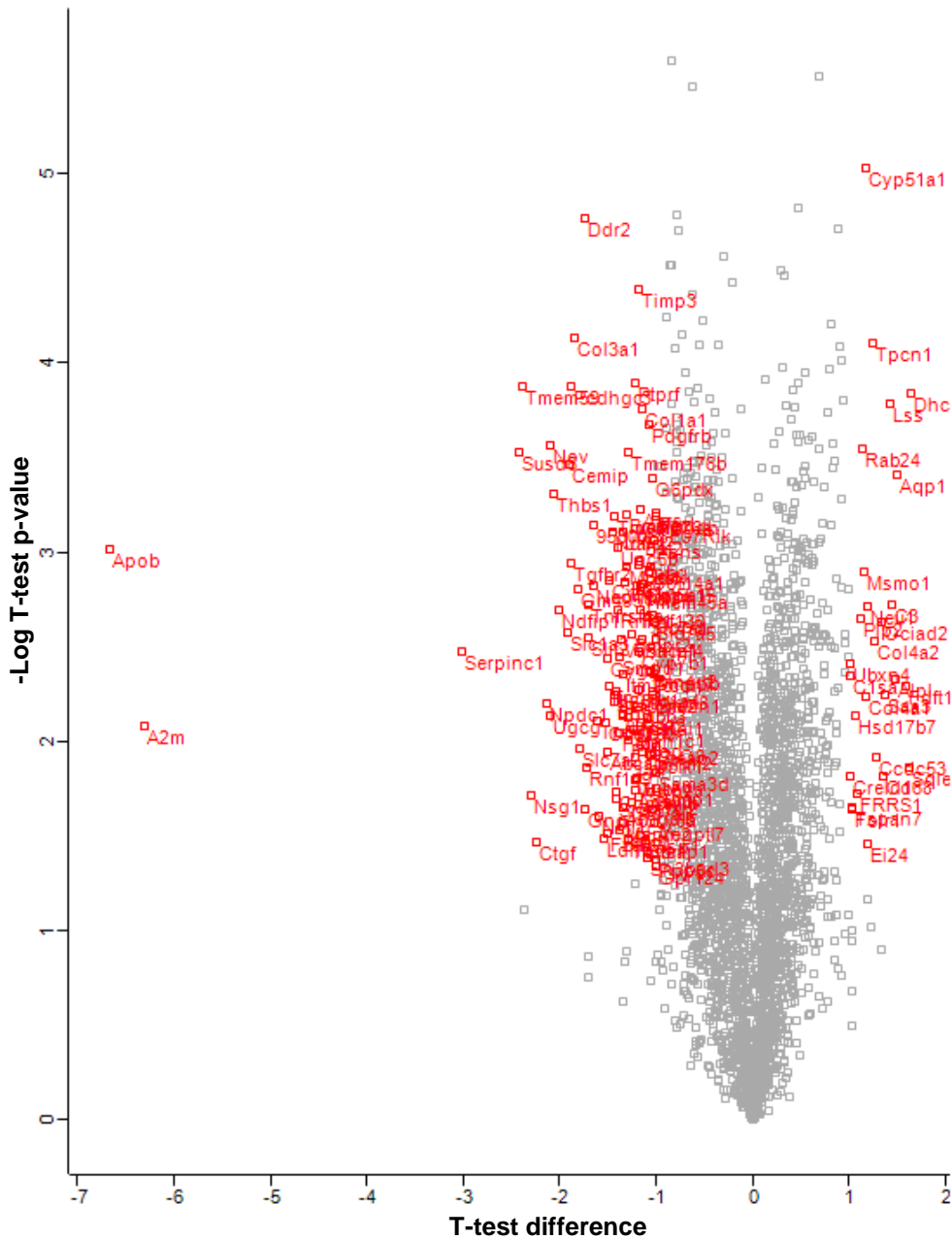
Figure 3.7: Volcano plot showing differentially expressed proteins upon U18666A treatment.
Proteins differentially expressed with the addition of U18666A to the culture medium were plotted here. Proteins which were detected in at least two biological replicates with p-values < 0.05 and had log2 transformed SILAC ratios higher or lower than +1 and -1 (arbitrary cut off), respectively were considered here. T-test differences of proteins were plotted against -Log T-test p-value using Perseus.

Table 3.1: Up-regulated proteins in the proteomics dataset.
The table shows up-regulated proteins upon U18666A treatment from MEF lysosomal proteomic dataset with more than 1 unique peptide and p-value < 0.05. Protein and gene names, log2 ratio (treatment/control) and the number of identified unique peptides are shown. Proteins which are related to cholesterol metabolism are marked as green and known lysosomal proteins are marked as yellow.

Protein Names	Gene Names	Log2 ratio (H/L)	Unique peptides
Delta(24)-sterol reductase	Dhcr24	1.6396	20
Squalene monooxygenase	Sqle	1.6251	13
Squalene synthase	Fdft1	1.5961	17
Aquaporin-1	Aqp1	1.4968	3
Alkaline phosphatase, tissue-nonspecific isozyme	Alpl	1.4882	11
Complement C3	C3	1.4460	58
Lanosterol synthase	Lss	1.4277	27
Serum amyloid A-3 protein	Saa3	1.3726	2
Isoform Short of Macrosialin	Cd68	1.3535	3
Gap junction alpha-1 protein	Gja1	1.3345	8
WASH complex subunit CCDC53	Ccdc53	1.2840	3
Collagen alpha-2(IV) chain	Col4a2	1.2588	10
Two pore calcium channel protein 1	Tpcn1	1.2422	8
Isoform 2 of Interleukin-1 receptor antagonist protein	Il1rn	1.2256	3
Sialidase-1	Neu1	1.1922	13
Dimethylaniline monooxygenase [N-oxide-forming] 2	Fmo2	1.1836	4
Lanosterol 14-alpha demethylase	Cyp51a1	1.1728	25
Collagen alpha-1(IV) chain	Col4a1	1.1660	5
Methylsterol monooxygenase 1	Msmo1	1.1502	8
Ras-related protein Rab-24	Rab24	1.1445	5
Proteolipid protein 2	Plp2	1.1239	2
Ferric-chelate reductase 1	FRRS1	1.0782	11
3-keto-steroid reductase	Hsd17b7	1.0628	10
Mitochondrial carrier homolog 1	Mtch1	1.0380	4
Folate receptor alpha	Folr1	1.0323	8
Tetraspanin-7	Tspan7	1.0311	5
Isoform 2 of Bifunctional polynucleotide phosphatase/kinase	Pnkp	1.0264	2
Charged multivesicular body protein 1b-1	Chmp1b1	1.0218	2
F-box only protein 6	Fbxo6	1.0213	4
Complement C1s-A subcomponent	C1sa	1.0180	10
Cysteine-rich with EGF-like domain protein 1	Creld1	1.0044	10

Table 3.2: Down-regulated lysosomal proteins in the proteomics dataset.
Down-regulated lysosomal proteins from U18666A treated MEFs with more than one unique peptide and p-value < 0.05.

Protein Names	Gene Names	Log 2 Ratio	Unique peptides
Hematopoietic progenitor cell antigen CD34	Cd34	-1.19	4
Sphingomyelin phosphodiesterase	Smpd1	-1.27	12
Integral membrane protein 2C	Itm2c	-1.35	10
Neuron-specific protein family member 1	Nsg1	-2.25	3
Transmembrane protein 59	Tmem59	-2.42	10

3.1.4 GO analysis of regulated proteins in the proteomics dataset

31 upregulated and 115 downregulated proteins were subjected to Panther GO analysis, where statistical overrepresentation test was performed for 'GO biological process complete' with a false discovery rate (FDR) of < 0.05. p-values were calculated using Fisher's exact test and ontologies were plotted against their p-values. Upregulated proteins were found to have an effect on cellular metabolic and biosynthetic processes with the lowest p-value (highest significance) for steroid metabolic process, secondary alcohol metabolic process and cholesterol metabolic process (Figure 3.8 (A)). In the case of 115 down-regulated proteins, in Figure 3.8 (B), the highest significance was found in the regulation of cell projection organization.
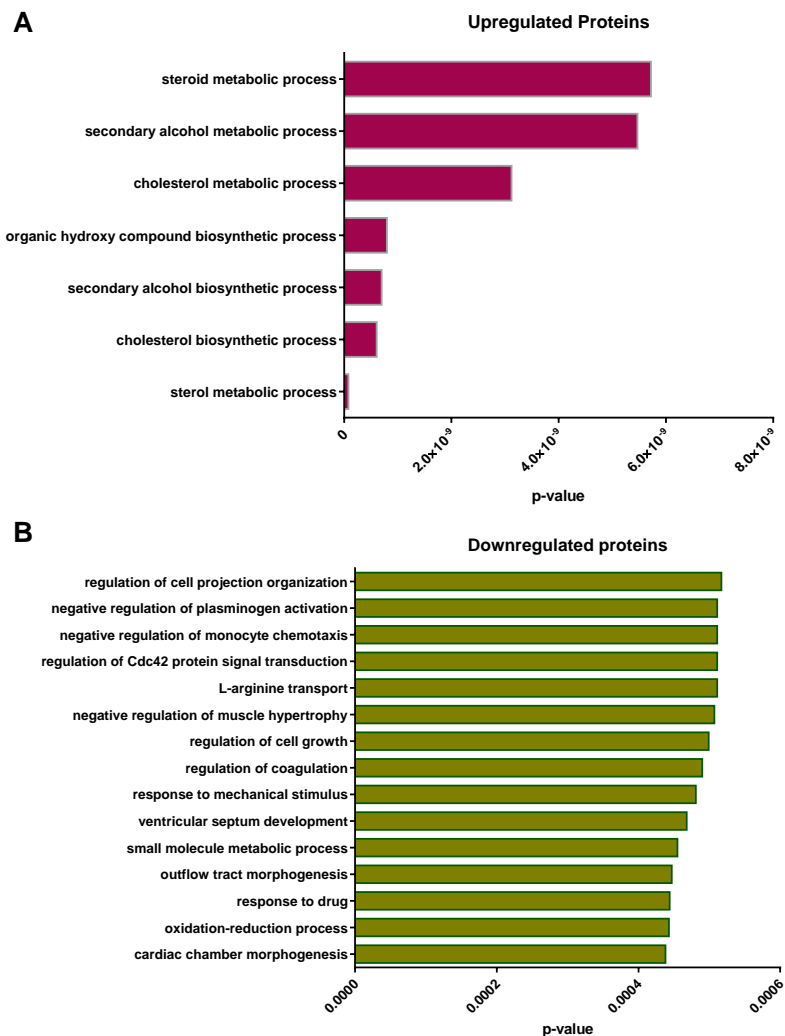
Figure 3.8: GO analysis of the differentially expressed proteins.
Gene ontology clustering was performed using PANTHER analysis (Mi et al. 2013, 2017). 31 upregulated proteins (A) or 115 downregulated proteins (B) were evaluated in the PANTHER software for the "GO biological process complete". Top enriched GO terms with a false discovery rate < 0.05 are shown here. Raw p-values were calculated using Fisher's exact test.

3.1.5 Lysosomal phosphoproteomic analysis of U18666A treated MEFs

In order to investigate whether lysosomal proteins are differentially phosphorylated in U18666A treated MEFs, mass spectrometry-based phosphoproteomics analysis was performed. For that purpose, 1.5 mg of combined lysosomal eluates from both U18666A (treatment) and DMSO (control) treated MEFs were enriched for phosphorylated peptides by double TiO_2 chromatography (Larsen et al. 2005). Then enriched phosphopeptides were desalted using StageTip desalting and analyzed by LC-MSMS (Figure 3.1). Mass spectrometry data were further evaluated by MaxQuant and Perseus.
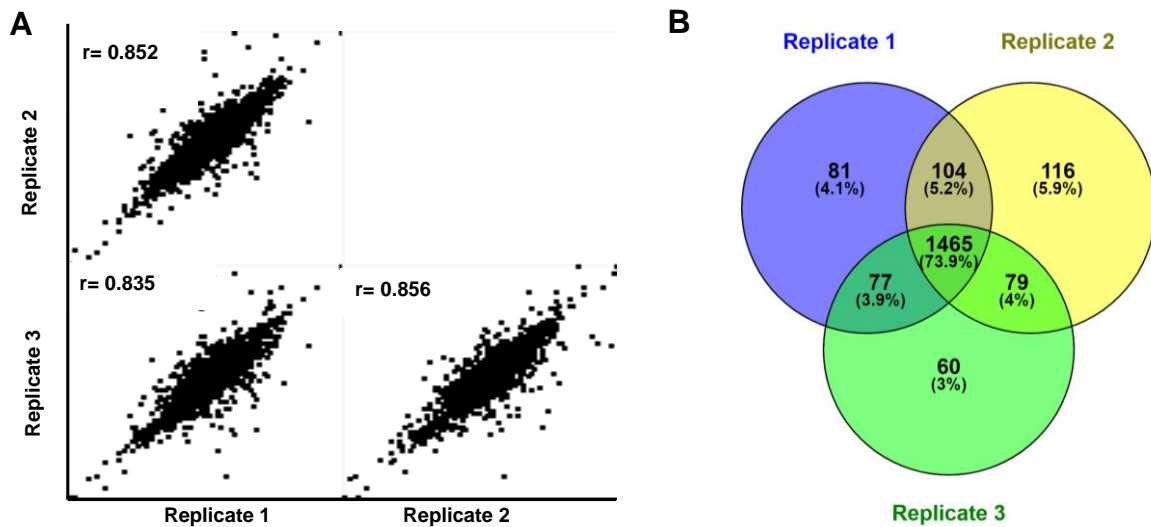
Figure 3.9: Experimental reproducibility of the lysosomal phosphoproteomic dataset.
A. Phosphoproteomics data were analyzed by MaxQuant and 'Phospho(STY)Sites' output of MaxQuant was further assessed by Perseus. Correlation between the three biological replicates was determined using Pearson's correlation coefficient (r). B. Venn diagram showing overlap between the total phosphoproteins found in three biological replicates. Online open-access tool Venny (Oliveros 2007) was used to generate Venn diagram.

From 3 biological replicates, 3046 proteins, 5730 phosphopeptides, and 7627 phosphorylation sites were detected. From a total of 3046 proteins, 2208 were found to be phosphorylated with 73 % phosphopeptide enrichment efficiency. Reproducibility among replicates was assessed using Pearson's correlation coefficient (r) where relatively high positive 'r' values confirmed high reproducibility of replicates (Figure 3.9 (A)). Around 74 % of phosphoproteins were found common in all 3 biological replicates whereas 13 % phosphoproteins were detected in only one replicate (Figure 3.9 (B)). Next, the distribution of Log2 SILAC H/L ratio data was assessed for total quantified phosphosites in each biological replicate and histogram in Figure 3.10 shows a normal distribution of SILAC H/L log2 ratio, as expected.
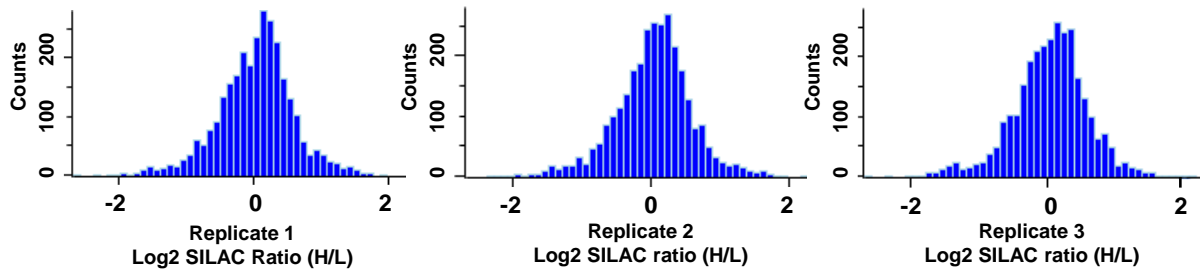
Figure 3.10: Distribution of SILAC H/L ratios of all 3 biological replicates.
SILAC H/L ratios from MaxQuant were converted to Log2 ratios and the distribution was calculated by using Perseus. In the histogram, the x-axis indicates Log2 SILAC (H/L) ratio and the y-axis shows the number of counts per bin.

Identified phosphopeptides were further analyzed for the presence of lysosomal and lysosomal-associated proteins by comparing both curated and Panther list. A total of 42 and 90 lysosomal plus lysosomal-associated proteins were detected as phosphorylated according to curated and Panther list respectively, out of which, 62 % and 67 % lysosomal plus lysosomal-associated proteins were found to be common in all 3 replicate (Figure 3.11).
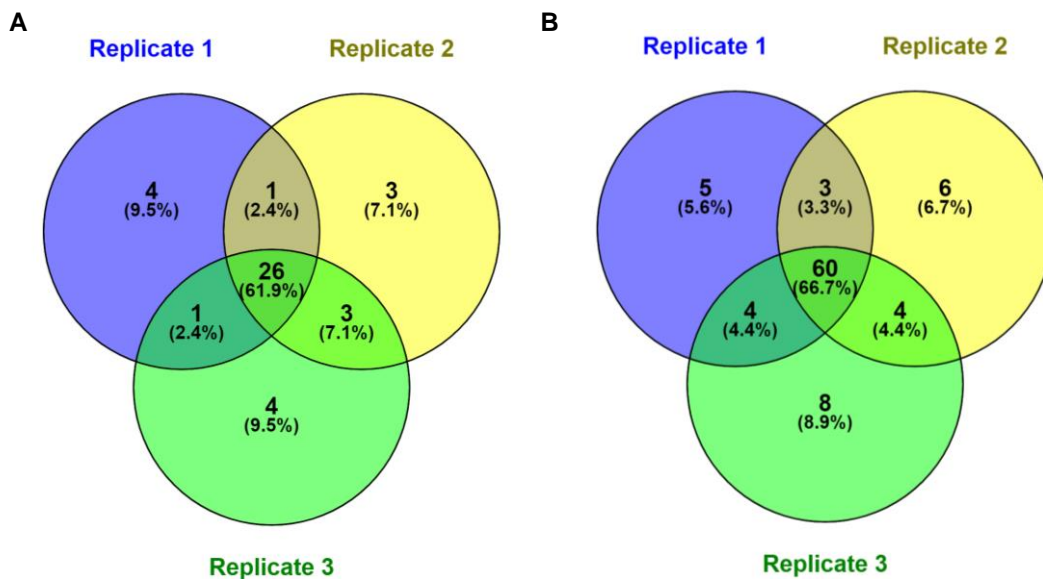
Figure 3.11: Lysosomal phosphosites detected in 3 biological replicates in phosphoproteomics dataset.
A. MaxQuant output 'phospho(STY)Sites' file was compared with an in-house curated lysosomal protein list and analyzed lysosomal phosphoproteins from 3 biological replicates by Venny. B. MaxQuant output 'phospho(STY)Sites' file was compared with Panther gene ontology lysosomal protein list and analyzed lysosomal phosphoproteins from 3 biological replicates by Venny (Oliveros 2007).

In the next step, 7627 detected phosphosites were further analyzed for up- and downregulated phosphosites. Based on localization probability (localization $p > 75\%$), 4821 phosphosites were detected as class I phosphosites where 123 and 122 phosphosites were up- and downregulated, respectively (Figure 3.12). These

phosphopeptides have SILAC log2 H/L ratios > 1 and were found in three biological replicates with p-values < 0.05. Up- and downregulated phosphosites were further analyzed for the presence of lysosomal phosphosites. Table 3.3 shows up-regulated and Table 3.4 shows down-regulated lysosomal phosphosites from U18666A treated phosphopeptides enriched lysosome samples.
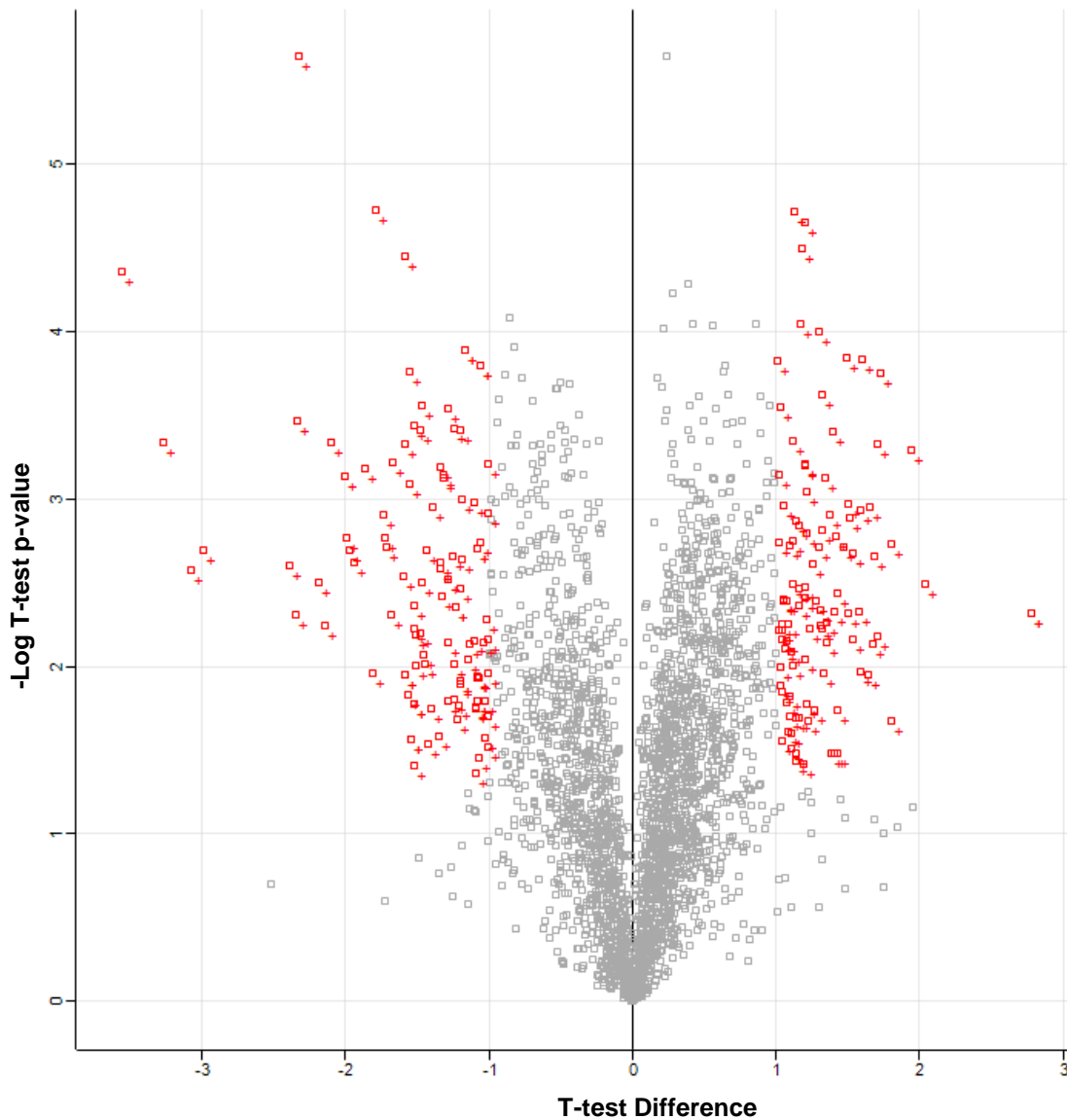
Figure 3.12: Volcano plot showing differentially regulated phosphopeptides from U18666A treated MEFs.
Phosphoproteomics dataset was analyzed for significantly regulated phosphopeptides. Phosphopeptides that were detected in at least two biological replicates with p-values less than 0.05 and had log2 transformed SILAC ratios higher or lower than +1 and -1 (the arbitrary cut off), respectively, were considered as regulated. T-test differences of phosphopeptides were plotted against -Log t-test P-value using Perseus.

Table 3.3: Up-regulated lysosomal phosphosites in the phosphoproteomics dataset.
11 lysosomal phosphosites were detected in the 123 upregulated phosphorylation sites from U18666A treated lysosomal phosphoproteomics dataset. Phosphorylation sites with localization probability > 75 % are shown below. 'Protein level ratio' shows the relative intensity of the corresponding protein in the proteomics dataset.

Gene Names	Position within protein	Localization prob	Phospho (STY) Probabilities	Ratio H/L (log2)	Protein level ratio (log2)
TPC1	767	1	TKS(1)DLSLK	1.7990	0.0146
Nbr1	701	0.81	DS(0.186)S(0.814)LKFALPEEGPR	1.7446	-0.2971
TM175	17	0.76	LQTEEQAVDSEGDS(0.758)S(0.242)LHR	1.6867	1.5286
MFSD1	20	1	READS(1)AVHGAPR	1.4906	1.4871
RAB9A	187	0.99	SEHLIQT(0.006)DT(0.994)VNLHR	1.3548	0.8134
TM175	18	1	LQTEEQAVDS(1)EGDSSLHR	1.2082	1.5286
VAMP8	5	1	MEEAS(0.998)GS(0.002)AGNDR	1.2002	1.1457
VAMP8	54	1	NKTEDLEAT(1)S(1)EHFK	1.1357	1.1457
RAB9A	185	1	SEHLIQT(0.998)DT(0.002)VNLHR	1.0721	0.8134
HGSNAT	202	1	LINS(1)ELGSPSR	1.0563	0.1631
Tmem106b	34	1	NGLVSS(1)EVHNEDGR	1.0555	0.3760

Table 3.4: Down-regulated lysosomal phosphosites in the phosphoproteomics dataset.
11 lysosomal phosphosites were detected in the 122 downregulated phosphorylation sites from U18666A treated lysosomal phosphoproteomics dataset. Phosphorylation sites with localization probability > 75 % are shown below. 'Protein level ratio' shows the relative intensity of the correspondent protein in the proteomics dataset.

Gene Names	Position within protein	Localization prob	Phospho (STY) Probabilities	Ratio H/L (log2)	Protein level ratio (log2)
Syt11	110	1	EVCES(1)PRKPIAK	-2.1829	-2.1462
NSG1	169	1	SVSPWMSVLS(1)EEK	-1.7230	-1.5935
PGFRB	747	1	DES(1)IDYVPMLDMK	-1.6005	-1.5251
Fgfr3	421	1	IARLS(1)S(1)GEGPVLANVSELELPADPK	-1.5479	-1.6612
Fgfr3	420	1	IARLS(1)S(1)GEGPVLANVSELELPADPK	-1.5479	-1.6612
FNBP1	296	0.98	T(0.004)VS(0.983)DNS(0.012)LSSSK	-1.4887	-1.0160
Abca2	1330	1	VSEEDQS(1)LENS(1)EADVK	-1.4428	-1.2950
Abca2	1326	1	VSEEDQS(1)LENS(1)EADVKESR	-1.3510	-1.2950
Rptor	722	0.96	SVS(0.027)S(0.961)Y(0.012)GNIR	-1.1909	-0.0340
AKTS1	92	1	AATATRPPGPPPAPQPPS(1)PAPS(1)PPPRPALAR	-1.0819	-0.5824
AKTS1	88	1	AATATRPPGPPPAPQPPS(1)PAPS(1)PPPRPALAR	-1.0819	-0.5824

3.1.6 GO analysis of differentially expressed phosphorylation sites

Using Panther Gene Ontology online tool, 68 phosphoproteins corresponding to 123 upregulated phosphosites were analyzed for 'GO biological process complete'. Figure 3.13 (A) shows GO terms with the lowest p-values (highly significant). Top three enriched GO terms were found to be 'retrograde transport, endosome to Golgi', 'membrane docking' and 'organelle membrane fusion'. GO analysis also shows, most

of the upregulated phosphoproteins were related to transportation, localization and intracellular organization (Figure 3.13 (A)). GO analysis was also performed for downregulated phosphosites for their role in biological processes. 90 phosphoproteins from 122 downregulated phosphosites were analyzed using Panther GO online tool, where top 3 GO terms include 'negative regulation of transport', 'regulation of localization' and 'transmembrane transport' (Figure 3.13 (B)). Out of 14 GO terms depicted in Figure 3.13 (B), 9 were related to the regulation of different cellular processes.
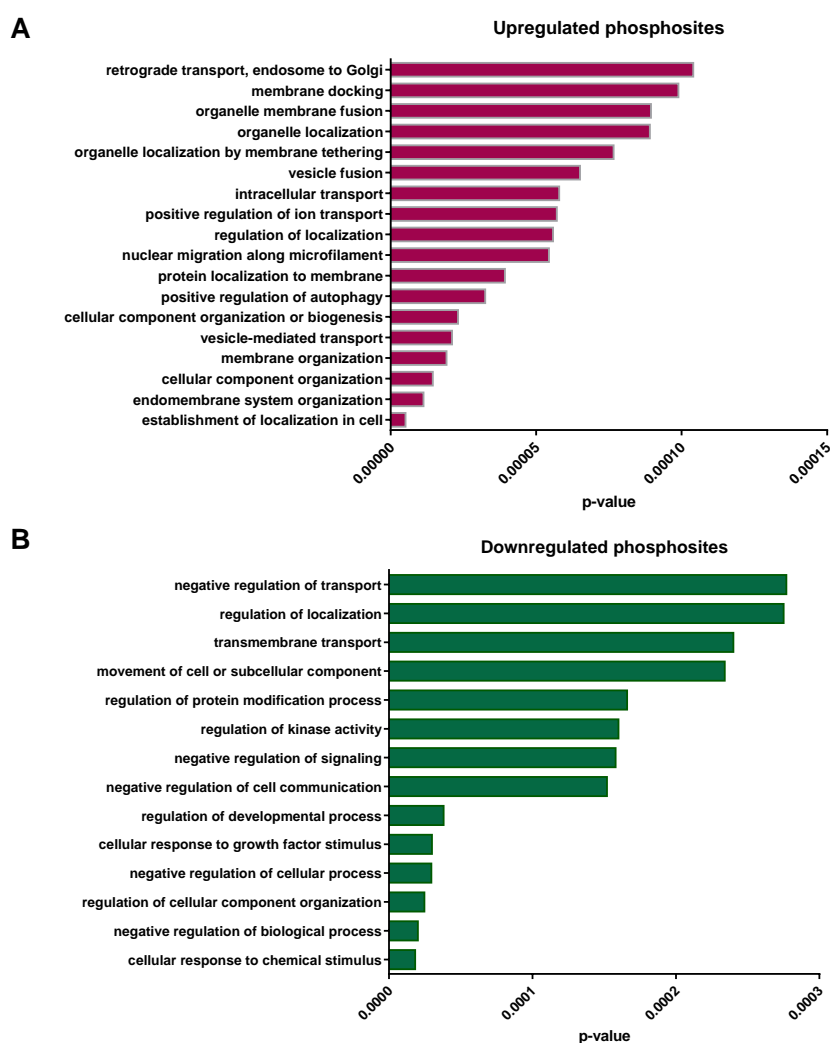
Figure 3.13: GO term enrichment analysis of differentially expressed phosphosites.
Gene ontology clustering was performed using PANTHER analysis (Mi et al. 2013, 2017). Phosphoproteins containing the 123 upregulated phosphosites (A) or 122 downregulated phosphosites (B) upon addition of U18666A to the culture medium in the phosphoproteomics dataset were subjected to the PANTHER software for the "GO biological process complete". Enriched GO terms with a false discovery rate < 0.05 are shown here. Raw P-values were calculated using the Fisher test. The diagram shows GO terms with the appertaining P-values (x-axis) in decreasing order.

3.2 Selection of candidate phosphoproteins and manual validation

Phosphoproteomics data analysis revealed 11 up-regulated and 11 down-regulated lysosomal phosphorylation sites upon U18666A treatment. Based on the clinical impact and literature review, HGSNAT (Heparan-α-glucosaminide N-acetyltransferase) was chosen for further investigation. Initially, mass spectrometry data were manually validated for the presence of phosphosites (Figure 3.14).
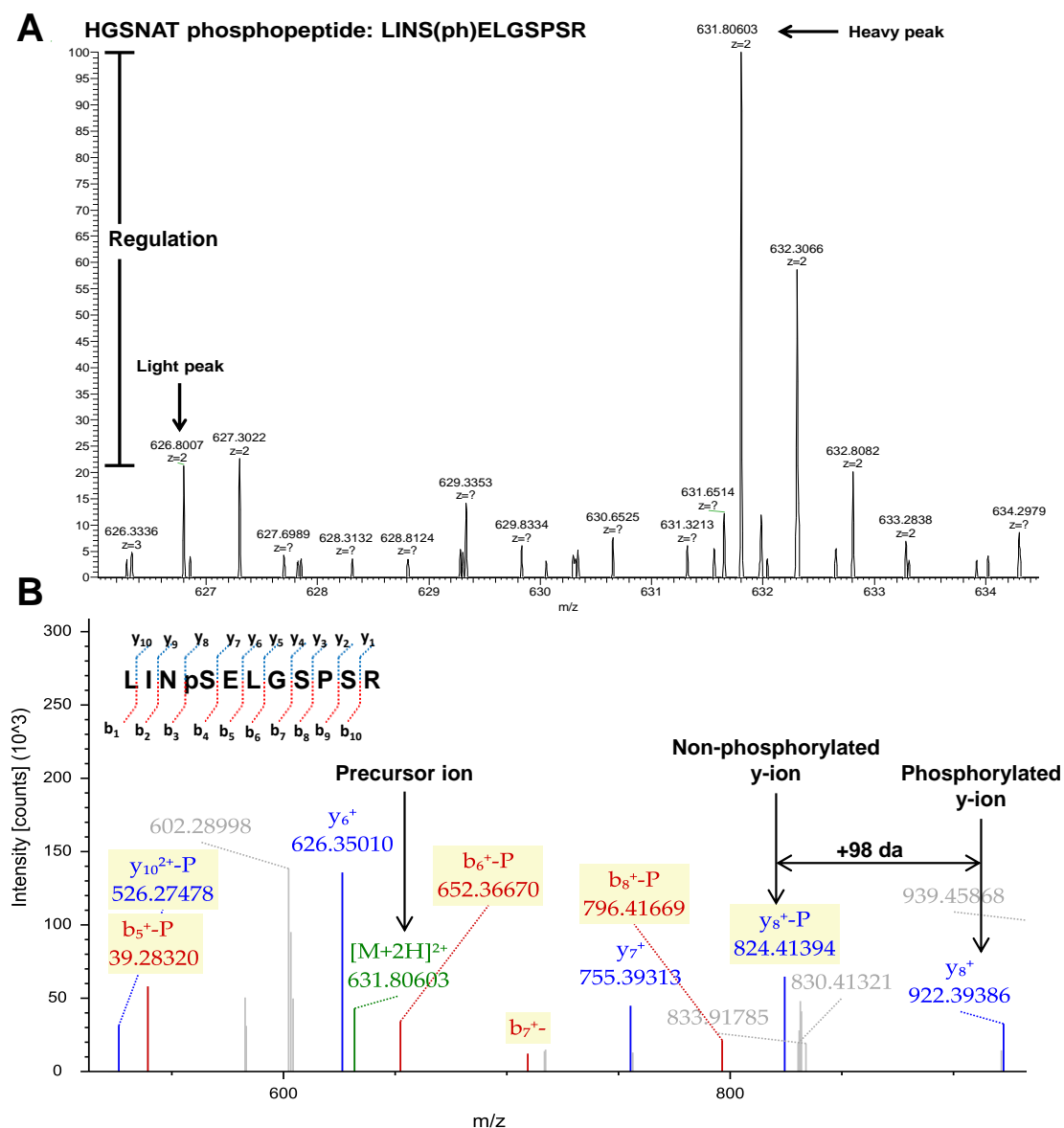
Figure 3.14: Manual validation of SILAC ratios using LC-MS survey scans and fragment spectrum.
To confirm results obtained from MaxQuant, LC-MSMS raw data were analyzed manually using Xcalibur (https://www.thermofisher.com/). A. Light and heavy spectra correspondent to the HGSNAT phosphopeptide and their intensities were compared by dividing relative abundance of the heavy peak to the light peak. B. Fragment spectrum containing phosphoserine 211. Green, red and blue colors are indicating precursor ions, b-ions, and y-ions respectively.

Localization probability of MaxQuant data showed 100 %, which is again confirmatory for the presence of phosphorylation sites on HGSNAT. Two phosphorylation sites, in case of human, Ser211, and Ser215 (Ser202 and Ser206, in case of the mouse) (see the amino acid sequences on Appendix 6.2) have been detected by mass spectrometry where serine 211 was found to be 2-fold up-regulated after U18666A treatment.

3.3 Effect of phosphorylation on HGSNAT enzyme activity

3.3.1 U18666A treatment regulates HGSNAT activity

HGSNAT-pS211 was found 2-fold up-regulated in U18666A treated lysosomal phosphoproteomics dataset, therefore the effects of U18666A treatment on HGSNAT enzyme activity was investigated. To that end, lysosomes were enriched from HEK293 cells with treatment of either DMSO (control) or 3 µg/ml U18666A for 24 hours. HGSNAT enzyme activity assay (Figure 3.15 (A)) was performed as described in 2.2.4.2 with 1 µg protein from enriched lysosomes. Figure 3.15 (B) shows U18666A treated HEK293 cells have 2-fold higher HGSNAT activity compared to control cells which were supplemented with DMSO, indicating U18666A treatment influences HGSNAT activity.
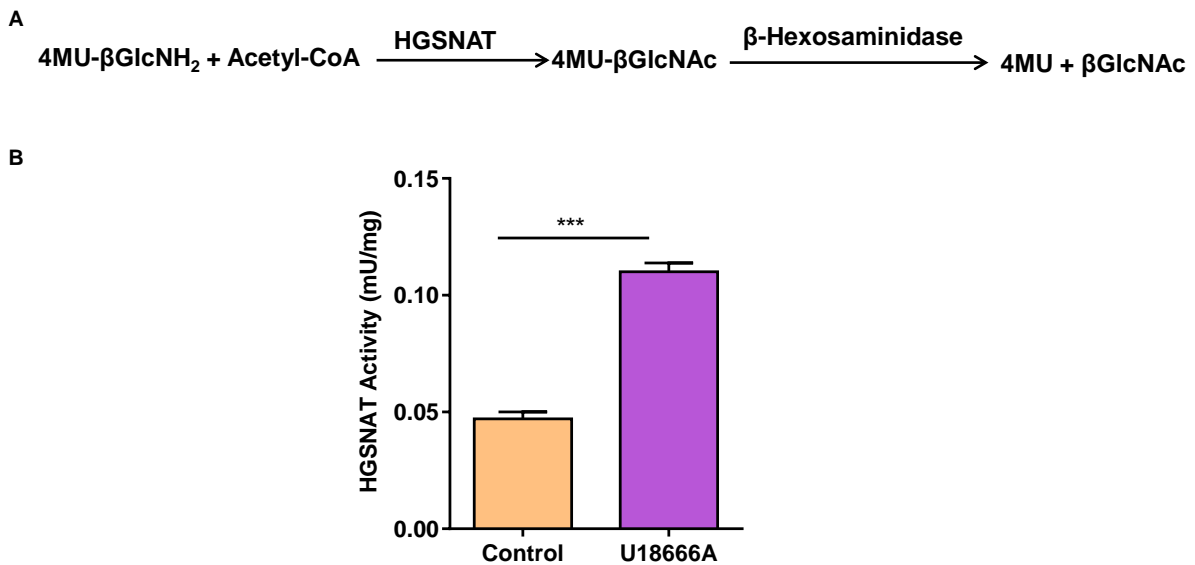
Figure 3.15: U18666A mediated up-regulation of HGSNAT activity.
A. The standard reaction for measuring HGSNAT activity. B. Lysosomes were enriched from control (DMSO) and U18666A (3 µg/ml) treated HEK293 cells and 1 µg protein from each sample was used for HGSNAT activity assay as shown in figure 3.15 (A). Shown are mean values ± SEM; n=3; ***: paired t-test with p-value 0.0004.

3.3.2 Impairment of heparan sulfate degradation affects HGSNAT activity

As HGSNAT plays a crucial role in the degradation of heparan sulfate in the lysosomal lumen, it was investigated whether impairment of heparan sulfate

degradation has any influence on HGSNAT activity. To that end, suramin was applied, which is an inhibitor of heparanase, thereby blocking both internalization and degradation of heparan sulfate (HS) (Nakajima et al. 1991). HEK293 cells were magnetite treated for lysosome enrichment and after 24 hours of pulse magnetite media was removed and cells were treated with 150 µM suramin for 24 hours (Mani et al. 2006). After lysosome enrichment they were lysed and subjected to HGSNAT activity assay. 1 µg protein from each sample was used and the assay was performed as in 2.2.4.2. Suramin treated samples showed a 2.5-fold decrease in HGSNAT activity in comparison to untreated control samples (Figure 3.16 (A)). Therefore, it suggests that blockage of heparanase and ultimately heparan sulfate degradation affects HGSNAT activity. Moreover, a separate HGSNAT activity assay was performed by combining both control and suramin treated samples to observe if suramin has any negative effect on lysosomes. The activity in combined samples showed a similar extent of decrease and increase in comparison to control and suramin treated samples respectively. This indicates suramin does not have any unusual effect on HGSNAT activity rather the decrease in activity compared to control samples is the real effect of suramin treatment (Figure 3.16 B).
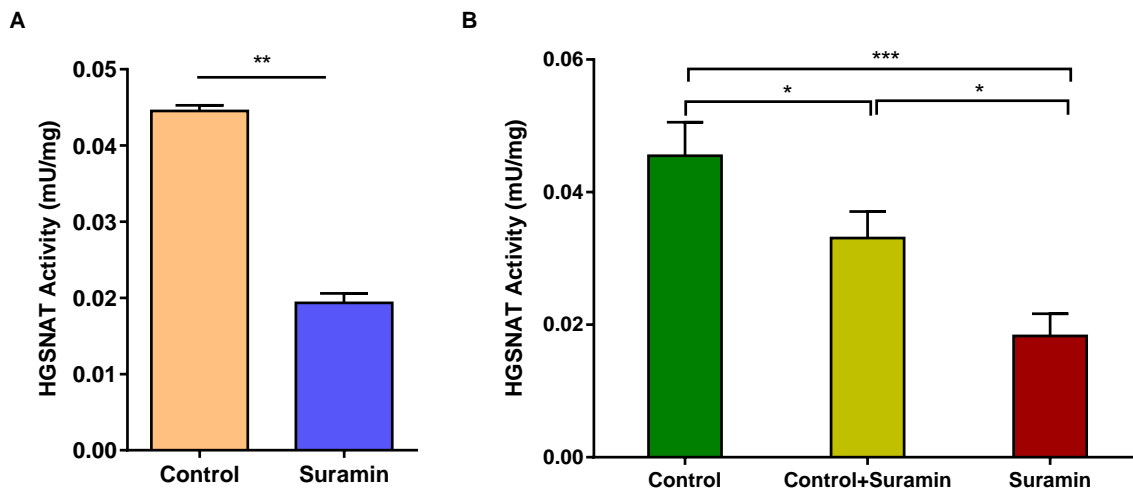
Figure 3.16: Suramin mediated down-regulation of HGSNAT activity.
A. Lysosomes were enriched from control or suramin (150 µM) treated HEK293 cells and 1 µg protein from each sample was used for HGSNAT activity assay. Shown are mean values ± SEM; n=3; **: paired t-test with p-value 0.0034. B. 1 µg lysosomes from each control and suramin treated HEK293 cells plus a combined sample containing 500 ng from each was subjected to HGSNAT assay. Shown are mean values ± SEM; n=3. One-way ANOVA with Tukey's multiple comparison test, *** = p-value 0.0006.

3.3.3 Dephosphorylation increases HGSNAT activity

Since phosphoproteomics data showed 2-fold upregulation of HGSNAT upon phosphorylation, therefore the role of phosphorylation on HGSNAT activity was

investigated. To that extent, 1 µg proteins from enriched lysosomes of HEK293 cells were treated with 1 unit Shrimp Alkaline Phosphatase (SAP) for 3 hours and then HGSNAT activity assay was performed on the treated and control samples as described in 2.2.4.2. Results in Figure 3.17 shows, upon SAP mediated dephosphorylation, the total activity of the HGSNAT enzyme has increased 2.4-fold from 0.05 mU/mg to 0.12 mU/mg in compare to untreated control samples.
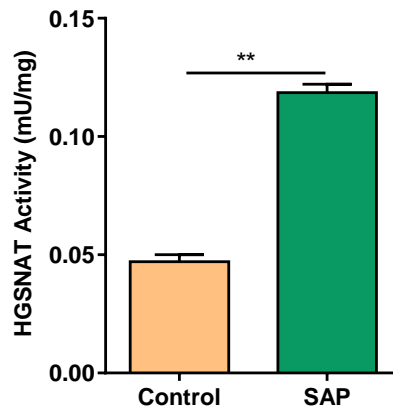
Figure 3.17: Dephosphorylation of HGSNAT by shrimp alkaline phosphatase.
1 µg enriched lysosomal samples were treated with or without 1-unit SAP and HGSNAT assay was performed for 17 h. Total activity was shown as mU/mg protein. Shown are mean values ± SEM; n=3; paired t-test with p-value 0.0063.

3.3.4 Generation of HGSNAT knock out HEK293 cells by CRISPR/Cas9

HEK293 cells were knocked out for HGSNAT gene using CRISPR/Cas9 method. To that end, HEK293 cells were co-transfected with crRNA targeting exon 7 of the HGSNAT gene (Figure 3.18 (A)), tracrRNA and Cas9 plasmid containing puromycin resistance gene. 48 hours post-transfection cell with media were supplemented with 2 µg/ml puromycin for another 48 hours. Then, cells were seeded under limiting dilution conditions and cultured for 2 weeks. Afterward, all single-cell clones were picked and duplicated, one duplicate was lysed to perform a locus-specific PCR and a subsequent second PCR was performed for attaching barcodes and sequencing adapters to them. Then PCR products were pooled and subjected to deep sequencing.

Figure 3.18 (A) shows the genomic location of HGSNAT and the crRNA sequence which was used for generating knock out cells. Red letters show the CRISPR target site and downstream of it is a PAM sequence (protospacer adjacent motifs) which is essential for Cas9 nuclease function. Figure 3.18 (B) shows results obtained from deep sequencing which were further analyzed by OutKnocker web tool (Schmid-Burgk et al. 2014). Each pie chart represents a single clone and its size refers to the

number of reads analyzed to assess the clone. Based on color code mutation status was defined. Mutations detected in three specific clones- clone 7, clone 23 and clone 38, shows 2 nucleotide deletions, 1 nucleotide insertion and 7 nucleotide deletion respectively (Figure 3.18 (C)), indicating they all are knock out cells.
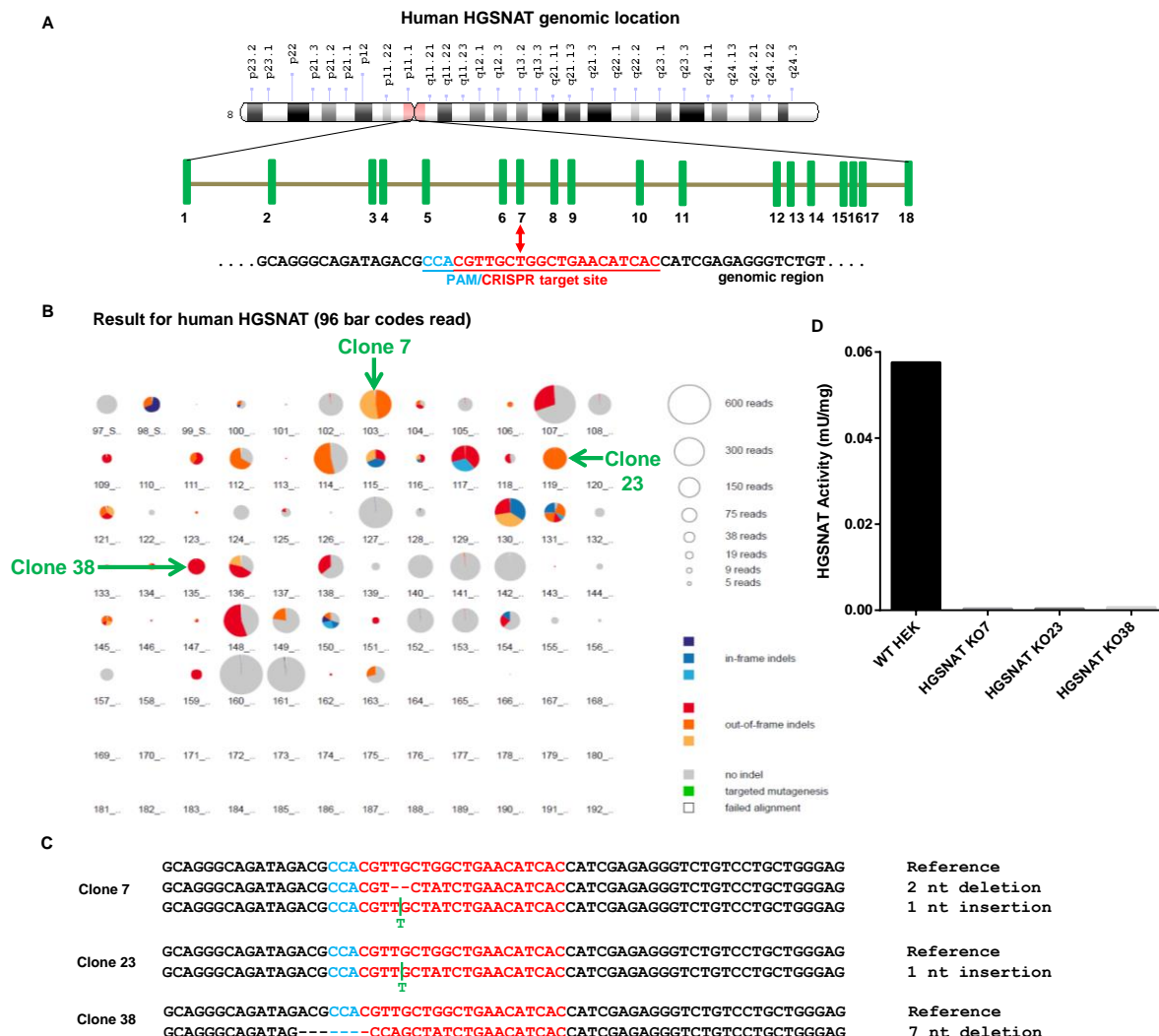
Figure 3.18: CRISPR/Cas9 mediated generation of HGSNAT-HEK293 knock out (KO) cell lines.
A. Illustration of the genomic locus of the human HGSNAT gene. Green bars showing exons and the red arrow highlights the target site of CRISPR which is magnified below. B. Results obtained from OutKnocker (Schmid-Burgk et al. 2014) web tool of HEK293 clones treated with a CRISPR targeting human HGSNAT. Every pie chart represents a clone, whereas the size of each chart corresponds to the number of reads that were analyzed to evaluate the clone. Colors of the individual pie areas indicate in-frame mutations (blue), out-of-frame mutations (red), or no indel (insertion or deletions) calls (gray). (C) Detected indel mutations of three knockout clones (7, 23 and 38) are shown. The PAM sequence is highlighted in cyan letters and red letters indicate the CRISPR target site. (D) HGSNAT enzyme activity with 10 µg proteins each from WT HEK293 cells (control) and knockout clones (clones 7, 23 and 38), which were shown with green arrows in figure 3.18 (B). The activity is shown as mU/mg.

Next, HGSNAT activity assays were performed to further confirm the clones are knockouts. In Figure 3.18 (D), it is shown that compared to the activity of WT

HEK293 cells, each of the three clones of HEK293 cells transfected with CRISPR method showed almost no activity (~165-fold decreased activity), which confirms that all these clones are HGSNAT knockout clones. To assess these clones further an antibody which detects endogenous HGSNAT proteins is needed.

3.3.5 Detection of endogenous HGSNAT by anti-HGSNAT antibody

Due to unavailability of an antibody which detects endogenous HGSNAT, it was decided to generate an antibody against HGSNAT protein. It is known that HGSNAT is proteolytically cleaved in the lysosomes to 27 kDa alpha-chain and 48 kDa beta-chain and both N- and C-terminal parts are connected by a disulfide bond (Fan et al. 2011). To generate an antibody against HGSNAT, N-terminal part (240 amino acids) of the protein was chosen due to the presence of the phosphopeptide (LINSELGSPSR) on that region. Therefore, 720 nucleotide base pairs of HGSNAT were cloned at the C-terminus of maltose-binding protein (MBP) followed by 6x Histidine tag (Figure 3.19 (A), Appendix 6.1 (A)). Protein was expressed in *E. coli* XL1 blue strain by IPTG induction and purified consecutively first by amylose resin and then nickel column using affinity chromatography technique (see 2.2.3.5) (Figure 3.19 (B)).

Purified protein was then sent to Pineda antibody service, Berlin (http://pineda-abservice.de/main.php) to inject a rabbit for generating antibody against HGSNAT. Antiserum received from Pineda was further purified by affinity chromatography using MBP-HGSNAT affinity column, where expressed MBP-HGSNAT fusion proteins were covalently coupled to amino link plus resin beads. The eluted pure antibody was run in SDS-PAGE and Coomassie-stained to visualize the heavy and light chain of the antibody (Figure 3.19 (C)). Finally, the activity of the antibody was observed by Western blot. IT was observed that the anti-HGSNAT antibody detects the endogenous alpha-chain of HGSNAT (~25 kDa) wherein HGSNAT KO cells this was missing (Figure 3.19 (D)). This also indicates HGSNAT KO cells were indeed true knock out cells. Furthermore, some non-specific bands at 35 kDa and 40 kDa were detected.
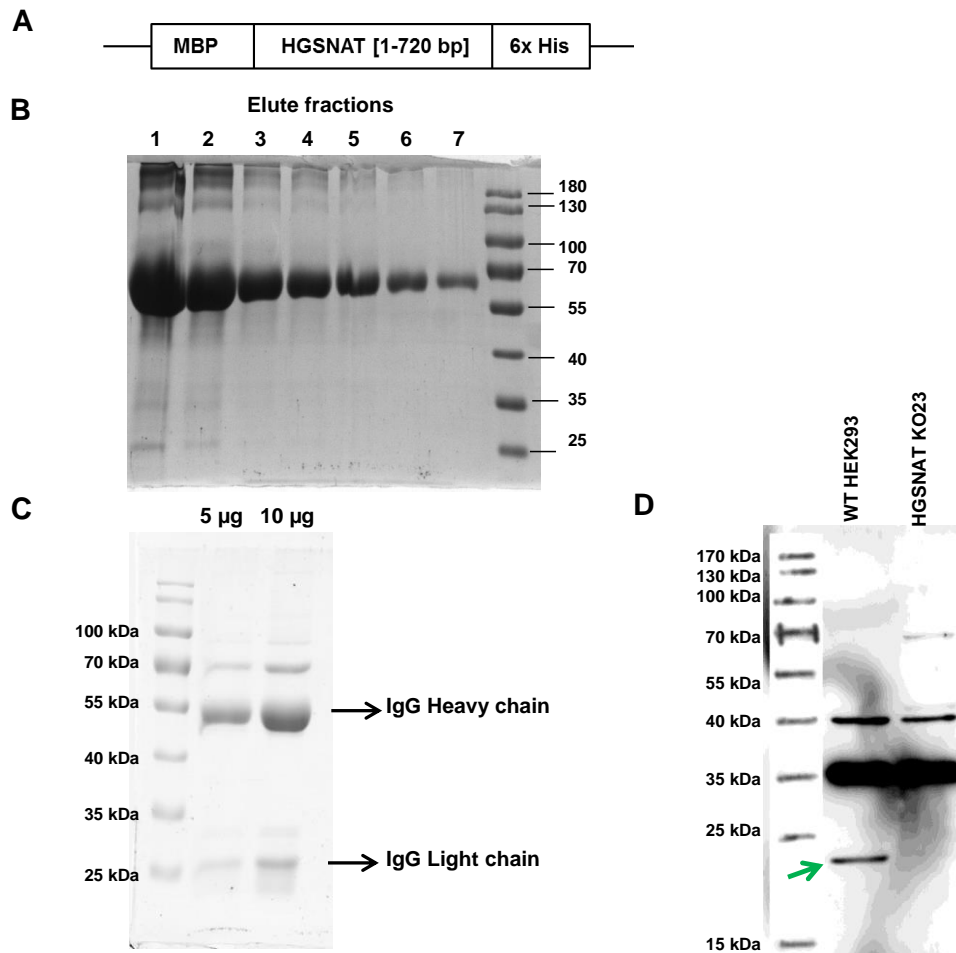
Figure 3.19: Production of anti-HGSNAT antibody against endogenous HGSNAT.
A. Scheme of the MBP-HGSNAT fusion protein. 720 bp of HGSNAT were fused to the C-terminus of maltose-binding protein (MBP) followed by 6x histidine tag. B. MBP-HGSNAT fusion protein expressed in *E. coli* was affinity purified and seven elute fractions were analyzed by SDS-PAGE. C. Immunoglobulin G (IgG) heavy (50 kDa) and light (25 kDa) chains were visualized in SDS-PAGE after affinity purification from anti-HGSNAT rabbit serum. D. 50 µg of protein from whole-cell lysates of each WT HEK293 and HGSNAT KO23 cells were subjected to Western blot analysis with affinity-purified rabbit anti-HGSNAT antibody and the blot was visualized using Fusion imaging system. The green arrow is showing the α-chain of HGSNAT.

Since the newly generated anti-HGSNAT antibody can detect endogenous HGSNAT, in order to normalize the HGSNAT activity U18666A (3 µg/ml) and suramin (150 µM) treated samples were subjected to western blot analysis using the newly generated anti-HGSNAT antibody for protein expression (Figure 3.20 (A)). Subsequently, the assay reactivity was normalized to the expressed protein values quantified using Fusion software. Figure 3.20 (B) shows, U18666A treatment increases the activity up to 2.8-fold and suramin treatment decreases the activity up to 4-fold in comparison to the control samples. Next, all samples were treated with SAP in order to observe its effect, since in endogenous samples it has increased the activity. Results show that

surprisingly in all samples irrespective of control or treated, dephosphorylation increases the activity ~2.5 fold (Figure 3.21).
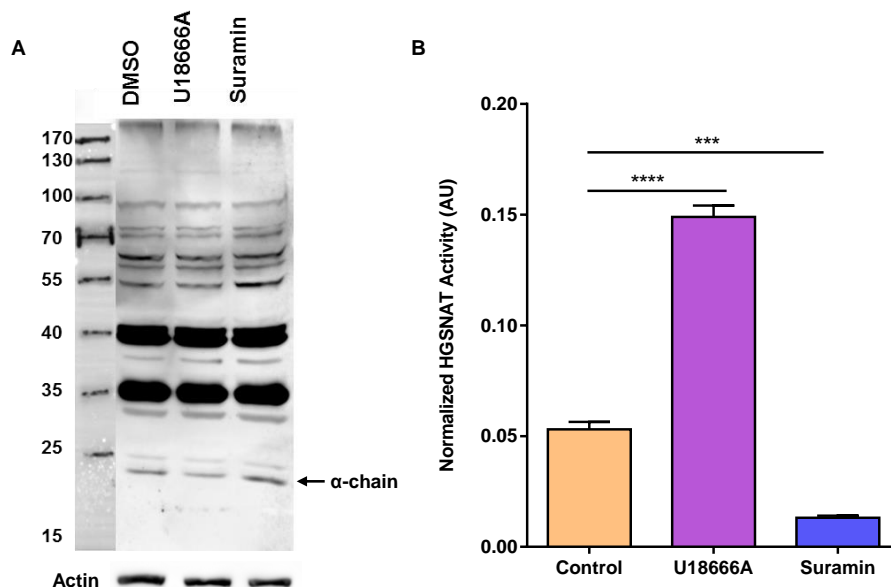
Figure 3.20: Effects of U18666A and suramin on HGSNAT protein expression and activity.
Lysosomes were enriched from HEK293 cells treated with either DMSO (control) or 3 µg/ml U18666A or 150 µM Suramin for 24 hours. 1 µg proteins from each sample were subjected to HGSNAT activity assay and 50 µg proteins from whole cells lysates were used for Western blotting using rabbit anti-HGSNAT antibody (1:500). After quantifying the relative amount of protein signal using Fusion, HGSNAT assay was normalized against the relative protein expression. Shown are mean values + SEM, n=3; one-way ANOVA, with Tukey's multiple comparison test; p-value control vs U18666A: ****: < 0.0001; control vs suramin: ***: 0.0005; AU: arbitrary unit.
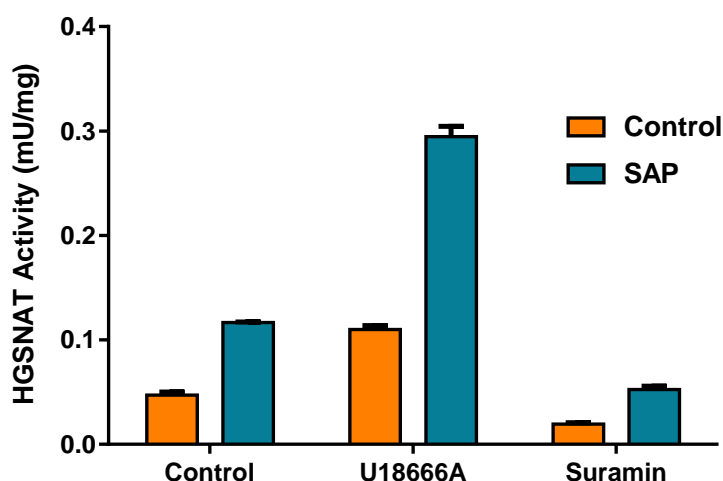
Figure 3.21: Dephosphorylation of U18666A and suramin treated samples.
Lysosomes were enriched from control, U18666A and suramin treated HEK293 cells and 1 µg protein from each sample was treated with or without 1-unit shrimp alkaline phosphatase (SAP) and performed HGSNAT activity assay for 17 hours. Activity is expressed as mU/mg protein. Shown are mean values + SEM, n=3.

3.3.6   Phosphomimetic amino acid substitution study

Phosphorylation can change the structure and conformation of proteins which can lead to an alteration of their activity, degradation rate, or of their interaction with other proteins (Krebs 1986). Phosphorylation of HGSNAT was found upregulated in the phosphoproteomic dataset; therefore, a phosphomimetic study was performed in which the candidate phosphorylated residues were mutated to either phosphomimetic glutamic acid (E)/ aspartic acid (D) or phosphoresistant alanine (A) residues (Figure 3.22 (A & B)) by site-directed mutagenesis PCR. With the due course, wild-type (WT) and 9 mutated HGSNAT cDNAs were transiently expressed in HEK293 cells and by using SPIONs techniques lysosomes were enriched from these cells. After lysis, the lysosomal proteins were subjected to an HGSNAT activity assay as well as Western blotting. In the following, the results of follow-up experiments with different mutants are presented.

A   **Phosphomimetics amino acid substitution**

Phospho-Serine (S)   Glutamic Acid (E)   Aspartic Acid (D)

Phosphoresistant amino acid substitution

Serine (S)   Alanine (A)

B
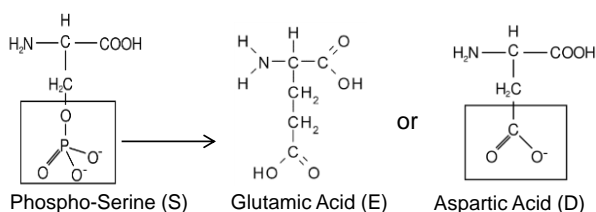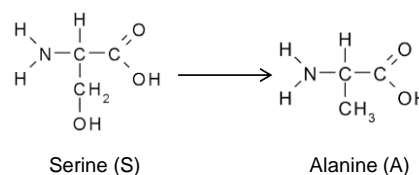
HGSNAT phosphopeptide	-LINS(p)ELGS(p)PSR-
S211E	[-LIN**E**ELGSPSR-]
S211A	[-LIN**A**ELGSPSR-]
S215E	[-LINSELG**E**PSR-]
S215A	[-LINSELG**A**PSR-]
S211E/S215E (EE)	[-LIN**E**ELG**E**PSR-]
S211E/S215A (EA)	[-LIN**E**ELG**A**PSR-]
S211A/S215A (AA)	[-LIN**A**ELG**A**PSR-]
S211A/S215E (AE)	[-LIN**A**ELG**E**PSR-]
S211D/S215D (DD)	[-LIN**D**ELG**D**PSR-]

Figure 3.22: Phosphomimic study of HGSNAT.
A. The structure of phosphoserine (S), glutamic acid (E), aspartic acid (D), serine (S) and alanine (A). B. HGSNAT phosphopeptide and its different mutants. Serine 211 and Serine 215 were mutated to either singly glutamic acid (S211E or S215E) or alanine (S211A or S215A) or doubly to glutamic acid plus alanine (S211E/S215E (EE), S211E/S215A (EA), S211A/S215A (AA), S211A/S215E (AE)) or aspartic acid (S211D/S215D (DD)) by site-directed mutagenesis PCR.

3.3.6.1 Optimization of HGSNAT assay and protein expression

The fluorimetric enzymatic assay (Voznyi et al. 1993), which was used in this study for the activity determination of HGSNAT, requires the sequential action of two lysosomal enzymes, HGSNAT, and β-hexosaminidase (Figure 3.15 (A)). HGSNAT acetylates the substrate, 4-methylumbelliferyl β-D-glucosaminide (4MU-βGlcNH$_2$) to 4MU-βGlcNAc and then endogenous β-hexosaminidase cleaves off 4MU. This free fluorescent 4MU reflects HGSNAT activity. Lysosomes from normal cells usually contain a very low amount of HGSNAT whereas there is an excess of β-hexosaminidase compared to HGSNAT. However, when cells were transfected with HGSNAT constructs, the amount of endogenous HGSNAT levels increases but β-hexosaminidase remains on the basal level. Therefore, the production of free fluorescent 4MU from reaction intermediate 4MU-βGlcNAc is not possible due to insufficient β-hexosaminidase activity. To keep the assay in the linear range, either the level of overexpression can be reduced by transfecting cells with less plasmid DNA or pure β-hexosaminidase can be added to the assay reaction.
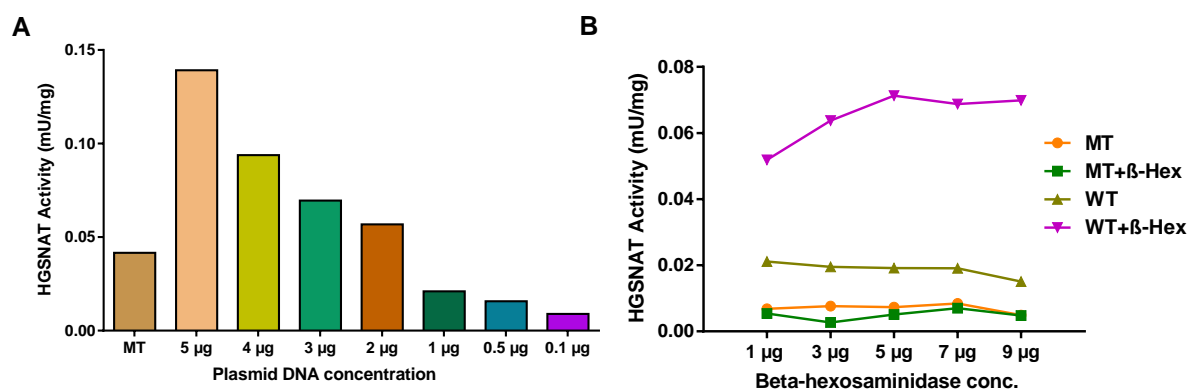
Figure 3.23: Optimization of HGSNAT expression and HGSNAT activity assays.
A. Lysosome enrichment was performed from HEK293 cells transfected with different amounts of WT HGSNAT plasmid. HGSNAT activity assay was performed with 1 µg protein from each sample expressed in mU/mg. B. Lysosomes from mock-transfected (MT) and WT HGSNAT transfected samples were subjected to HGSNAT activity assay by adding different amounts of β-hexosaminidase. Activities were calculated as mU/mg of total protein in the sample.

To apply the first one, HEK293 cells were transiently transfected with different amounts of WT HGSNAT constructs (from 5 µg to 0.1 µg in a 10 cm cell culture dish). Lysosomes were enriched from these cells and 1 µg protein from each sample was subjected to an HGSNAT activity assay and results showed that the activity of transfected cells was ~4-fold higher than the endogenous sample, in case of 5 µg HGSNAT plasmid (Figure 3.23 (A)). However, it was also observed that 2 µg DNA expressed HGSNAT slightly higher than the endogenous sample. Proceeding from

this result, it was decided to transfect one 10 cm cell culture dish of HEK293 cells with 1.5 µg plasmid DNA from each HGSNAT construct.

On the other hand, in order to add β-hexosaminidase to the assay reaction, pure β-hexosaminidase extracted and purified from fungus *Pichia pastoris* was received as a kind gift from Dr. Bojarova, Prague, CZ (Slámová et al. 2012). Subsequently, the enzyme was titrated in different concentrations (1, 3, 5, 7 and 9 µg) to the assay reaction. It was observed that the activity has increased from 1 to 5 µg and then it stayed steady (Figure 3.23 (B)). Therefore, it was decided to add 5 µg of the pure β-hexosaminidase enzyme to the assay reaction when transfected samples will be used. In the case of untransfected samples, no β-hexosaminidase is needed, since endogenous β-hexosaminidase is sufficient to convert all intermediate products (4MU-βGlcNAc) to final free fluorescent 4MU.

3.3.6.2 Expression of HGSNAT mutants on HGSNAT KO HEK293 cells

In order to reduce endogenous background and observe a clear effect of the HGSNAT phosphomutants, CRISPR/Cas9 mediated HGSNAT KO HEK293 cells were used. Lysosomes were enriched from these cells after transient transfection with all HGSNAT constructs in pcDNA3.1 (Figure 3.24 (A), see plasmid map in Appendix 6.1 (B)) and effects were observed by HGSNAT activity assay (Figure 3.24 (C)) as well as a Western blot (Figure 3.24 (B)). Although transfection was carried out with the optimized DNA concentration for close to basal level expression, it was seen that the expression of all constructs was extremely higher. Furthermore, β-hexosaminidase was provided to the assay reaction but the overexpression was too high that endogenous as well as exogenous β-hexosaminidase was not sufficient to convert all reactions intermediates. Hence, the HGSNAT activity assay results were not conclusive in this aspect (Figure 3.24 (C)).

A

HGSNAT KO HEK293 cells ⟶ Transfection ⟶ 24h Magnetite treatment ⟶ 24h Chase ⟶ Lysosome enrichment

B

C
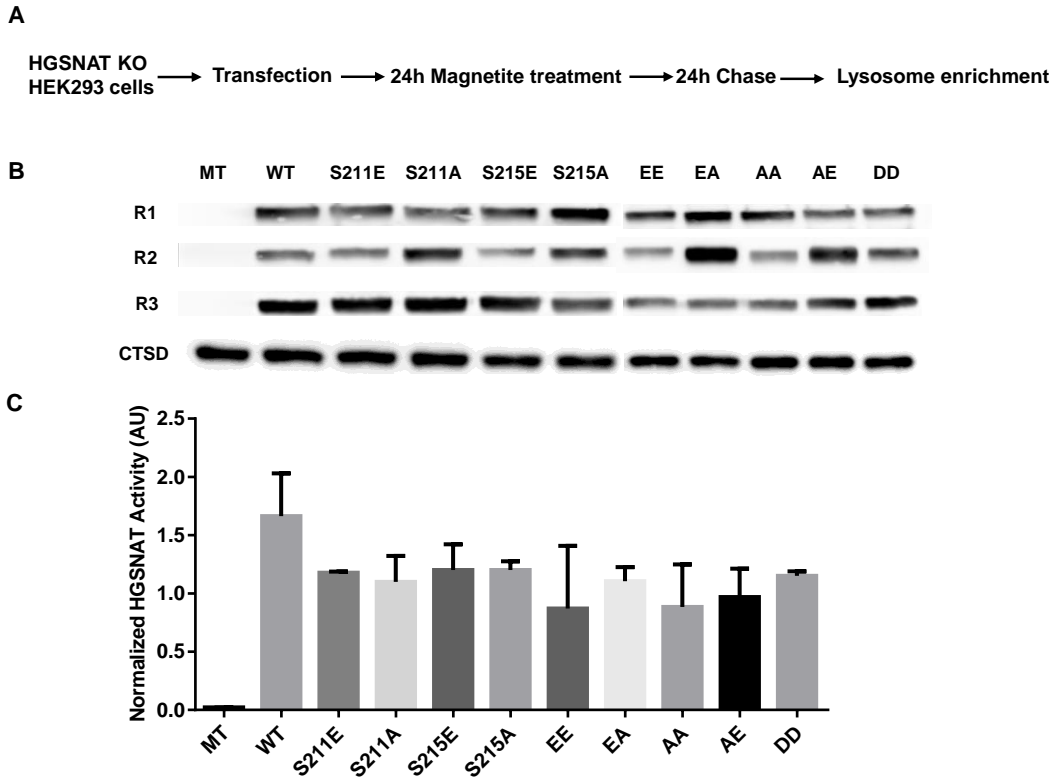
Figure 3.24: Expression of HGSNAT mutants on HGSNAT KO HEK293 cells.
A. Lysosomes were enriched from HGSNAT KO HEK293 cells transfected with HGSNAT constructs. B. 10 µg protein of each sample was subjected to Western blot using anti-FLAG (1:5000) or cathepsin D (CTSD) (1:1000). CTSD is shown for one replicate. C. 200 ng proteins from all samples were used for HGSNAT activity assay. Specific activities were calculated as mU/mg and the signal intensity of the FLAG in each sample was normalized against the signal intensity of CTSD. Shown are mean values + SEM, n=3; one-way ANOVA, with Tukey's multiple comparison test; p-value < 0.05; MT: Mock transfected, AU: arbitrary unit.

3.3.6.3 HGSNAT Tet-On cells for endogenous expression

Even though after transfecting HGSNAT KO cells with HGSNAT mutants, the expression level was still very high. Therefore, it was decided to convert HGSNAT KO cells into a Tet inducible system by using the piggyBac transposon method (Li et al. 2013). In this method, titrating doxycycline concentration helps in reducing protein expression level close to the endogenous level. For that purpose, HGSNAT KO cells were stably co-transfected with a piggyBac transposon plasmid containing rtTA gene (Appendix 6.1 (D)) and a transposase plasmid (Appendix 6.1 (E)) in several transposons to transposase ratios (5:1, 2.5:1 and 1:1). The cells were then selected with supplementation of 300 µg/ml G418 antibiotic (a doxycycline derivative). Meanwhile, all HGSNAT constructs (WT and 9 phosphomutants) were cloned into the pTRE3G plasmid (Appendix 6.1 (C)) where the protein expression is controlled by a TRE3G promoter in the presence of doxycycline antibiotics. In order to assess their activity, each of these newly generated HGSNAT Tet-On cells were transiently

transfected with WT HGSNAT plasmid with (1 µg/ml) or without doxycycline supplementation. Lysosomes were enriched from these cells and 20 µg proteins from each sample were subjected to Western blot analysis. It was observed that stable cells made with transposon to transposase ratio of 2.5:1 expressed the protein better in comparison to the other two ratios with the presence of G418. However, without G418 supplementation proteins didn't express in case of all 3 samples (Figure 3.25 (A)). Next, 1 µg of protein from the same samples was used for the HGSNAT activity assay (Figure 3.25 (B)) which showed that the cells made with 2.5:1 of transposon to transposase ratios had the highest activity of 0.3 mU/mg protein in comparison to 5:1 and 1:1 cell with 0.16 and 0.21 mU/mg, respectively. Additionally, the cells which were not selected with G418 did not show any activity. Therefore, it was found that HGSNAT KO cells had been converted to HGSNAT Tet-On cells and since cells with 2.5:1 (transposon: transposase) ratios showed the highest activity, it was decided to use these cells for follow-up experiments.
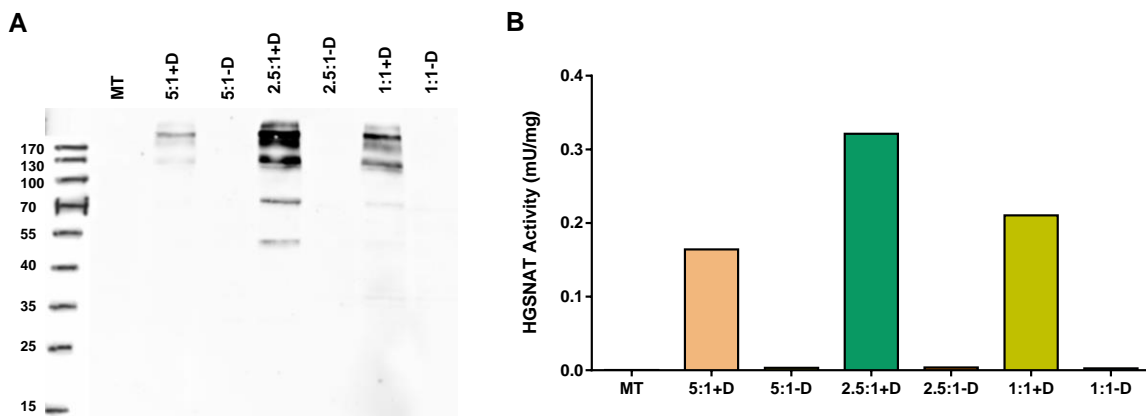
Figure 3.25: Evaluation of newly generated HGSNAT Tet-On cells.
A. Newly generated three different HGSNAT Tet-On cell lines were transfected with WT HGSNAT plasmid with (1 µg/ml) or without doxycycline and lysosomes were enriched from these cells. 20 µg proteins from each sample were subjected to Western blot using anti-FLAG antibody (1:5000). B. 1 µg proteins from each sample were assayed for HGSNAT activity and total enzymatic activity is expressed as mU/mg.

Newly generated HGSNAT Tet-On cells were first titrated for doxycycline concentration in order to achieve a protein expression close to the endogenous level. Therefore, the HGSNAT Tet-On cells were transfected with 6 µg of WT HGSNAT plasmid with no doxycycline (control) or different concentrations (0.5, 0.05, 0.005, 0.0005, and 0.00005 µg/ml) of doxycycline. Moreover, wild type HEK293 cells were used as a positive control in order to compare them with the endogenous HGSNAT activity. Then, after enriching lysosomes from these cells Western blot analysis (Figure 3.26 (A)) as well as the HGSNAT assay was performed. It was found that,

even 0.00005 µg/ml doxycycline results in an HGSNAT-expression, which is 2.9-fold higher than the endogenous HGSNAT activity (Figure 3.26 (B)). Therefore, it was decided to use a lower concentration which was selected as 0.5 pg/ml of doxycycline in the cell culture medium to express HGSNAT constructs.
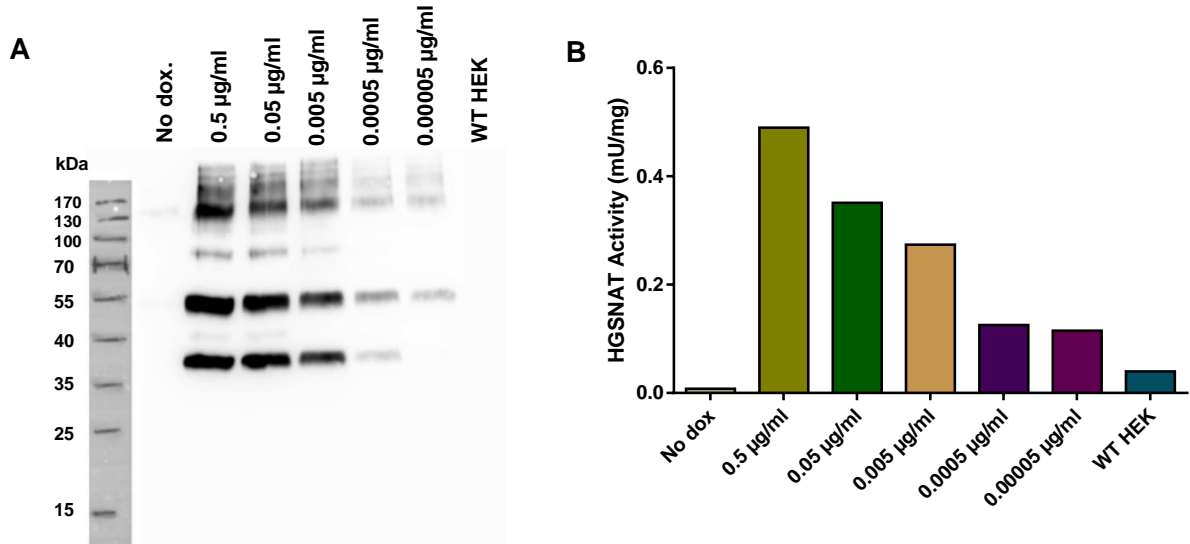
Figure 3.26: Doxycycline concentration determination for HGSNAT expression.
A. HGSNAT Tet-On HEK293 cells were transfected with pTRE3G-HGSNAT-WT and expressed under the control of the different concentration of doxycycline. After lysosome enrichment, 10 µg proteins from each sample were used for Western blot with anti-FLAG antibody (1:5000). B. 1 µg proteins from each sample were assayed for HGSNAT activity and expressed in mU/mg of proteins.

Next, HGSNAT Tet-On HEK293 cells were transfected with WT and all 9 mutants, where protein expression was controlled by adding 0.5 pg/ml doxycycline to the culture media. The normalized activity of HGSNAT mutants showed a decrease in activity in all samples compare to WT samples, where only S215E, EA, and DD mutants showed a statistically significant decrease in activity with 1.79-, 2.34- and 1.99-fold respectively compared to WT samples (Figure 3.27 (C)). According to our hypothesis, S211A should reduce the activity while S211E should increase, as this site was regulated on phosphorylation. Since S215 was not regulated in mass spectrometry analysis, we could not foresee its function. However, the results are conflicting to our hypothesis and also not clear about which site has regulated the phosphorylation.

A

HGSNAT Tet-On HEK293 cells → Transfection + 0.5 pg/ml doxycycline → 24h Magnetite treatment → 24h Chase → Lysosome enrichment

B

C
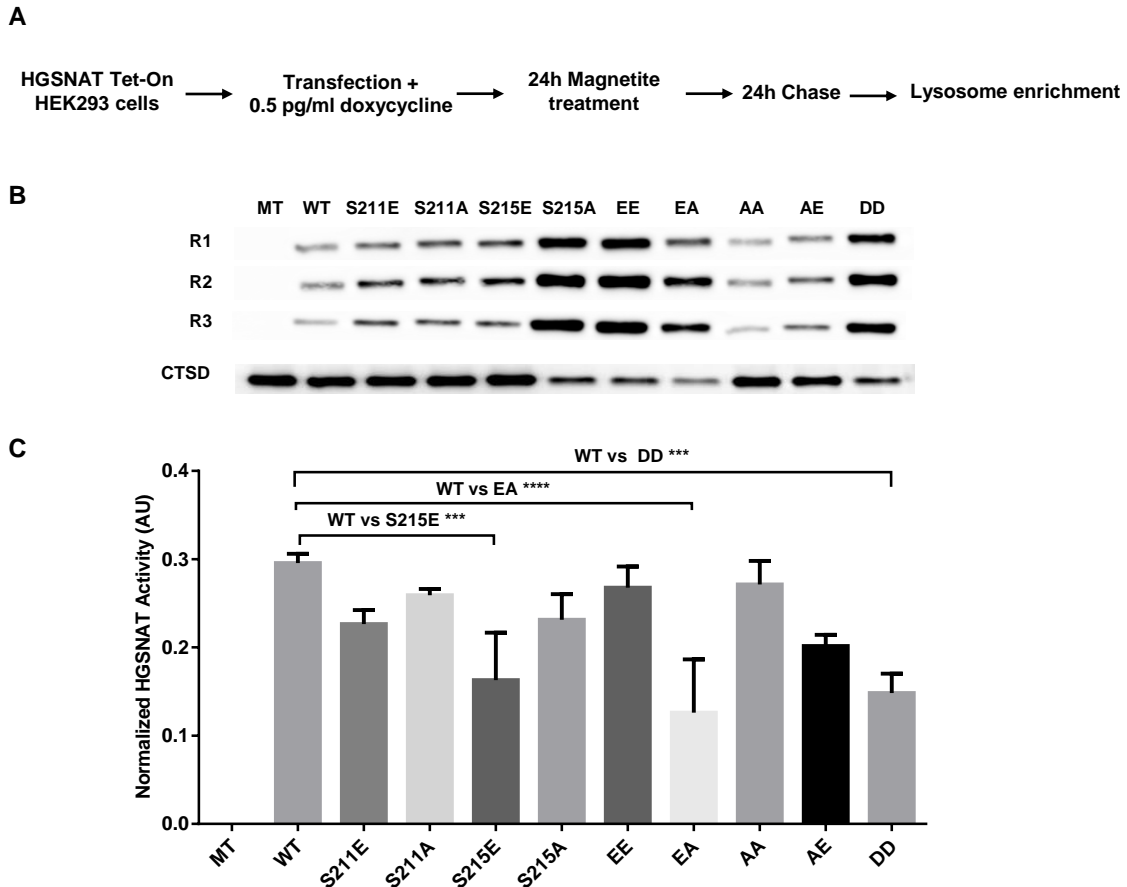
Figure 3.27: Effects of phospho-mutants on HGSNAT Tet-On cells.
A. HGSNAT Tet-On HEK293 cells were transiently transfected with pTRE3G-HGSNAT constructs and lysosomes were enriched. B. 10 µg protein of each sample was used for Western blotting with anti-FLAG (1:5000) or cathepsin D (1:1000) antibodies. Blots were visualized and quantified using Fusion imaging system. C. 500 ng proteins from each sample were used for HGSNAT activity assay and activity was normalized against the protein expression. Shown are mean values + SEM, n=3; one-way ANOVA, with Tukey's multiple comparison test; p-value < 0.05; AU: arbitrary unit.

Moreover, all double mutant transfected samples were dephosphorylated using SAP and surprisingly, all of them were showing a decrease in activity upon dephosphorylation (Figure 3.28 (A)) which is completely contrary to what was found when investigating the endogenous samples (Figure 3.17). In another experiment, the cells overexpressing HGSNAT were treated with U18666A in order to examine whether there is a change in HGSNAT activity. A slight increase in HGSNAT activity was observed to a similar extent in WT as well as AA and DD mutants after U18666A treatment, but not significantly (Figure 3.28 (B)), whereas there was a significant increase in activity in endogenous samples (Figure 3.20 (B)).
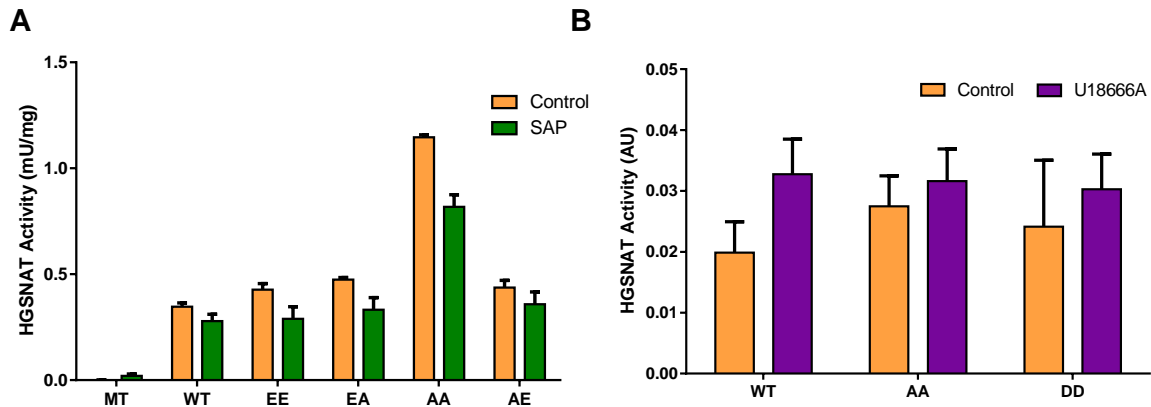
Figure 3.28: SAP and U18666A treatment of transfected samples.
A. Lysosomes were enriched from overexpressed samples and HGSNAT assay was performed with SAP treatment. B. Lysosomes were enriched from U18666A treated overexpressed cells and subjected to HGSNAT assay. In both A and B, specific activities were calculated as mU/mg. n=3, shown are mean values ± SEM.

3.4 Subcellular localization of HGSNAT phosphomutants

The effects of HGSNAT phosphomutants on their subcellular localization were studied by immunofluorescence microscopy. NIH3T3 Tet-On cells were transfected with constructs expressing each of the phosphomutants under the control of doxycycline, after fixing and permeabilization cells were probed with anti-FLAG antibody to identify the HGSNAT and anti-Lamp2 antibody as a lysosomal marker. WT and all mutants showed a distinct punctate staining characteristic for lysosomal targeting. Accordingly, WT and all phosphomutants co-localized with the lysosomal marker Lamp2, as indicated by the yellow color on the merged image (Figure 3.29)

Next, different fractions of samples during lysosome enrichment were investigated for HGSNAT subcellular localization. To that extent, input (cell homogenate) fractions containing all cellular organelles except nuclear pellet, flow-through fractions from the magnetite column which should contain organelles other than lysosomes and elute fractions from the magnetic column containing enriched lysosomes were tested for the presence of HGSNAT enzyme by Western blot and HGSNAT enzyme activity assay. It has been detected that WT and HGSNAT phosphomutants are localized in the lysosome enriched samples (Figure 3.30 (A & B)) showing that mutants have no effect on HGSNAT localization.
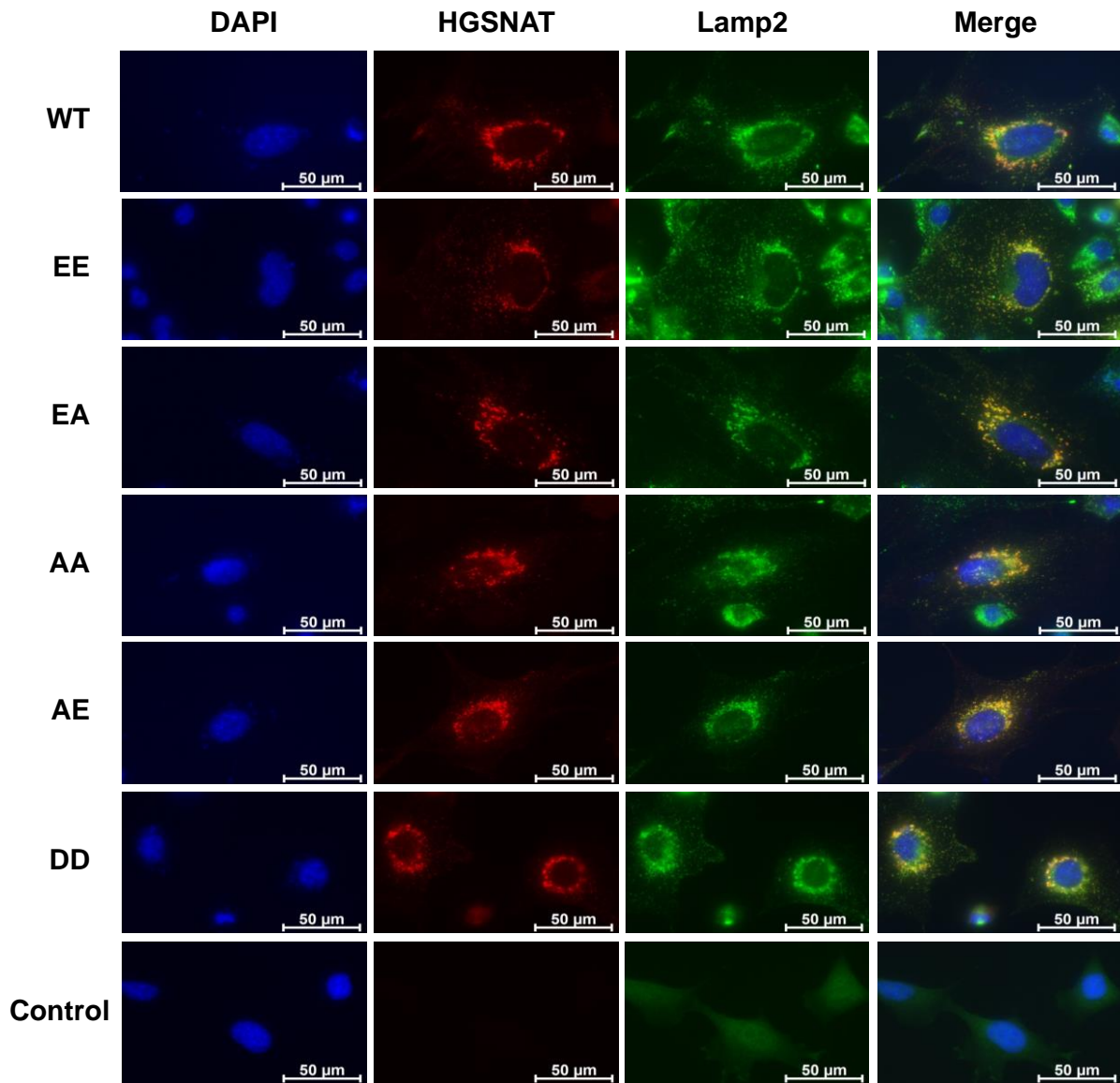
Figure 3.29: **Localization of HGSNAT phosphomutants expressed in NIH3T3 Tet-On cells by immunofluorescence microscopy.**
NIH3T3 cells were seeded on 12 mm coverslips and transfected with respective HGSNAT constructs. Anti-myc (1:400) and anti-Lamp2 (1:100) and primary antibodies were employed to identify HGSNAT and lysosomes, respectively. Immunocytochemistry images were acquired using Axiovert 200M equipped with an AxioCamMR3 (60x objective). Filter sets 38, 43 and 49 were used for detecting Cy3 (red fluorescence, HGSNAT), Alexa 488 (green fluorescence, Lamp2), and DAPI (blue), respectively.
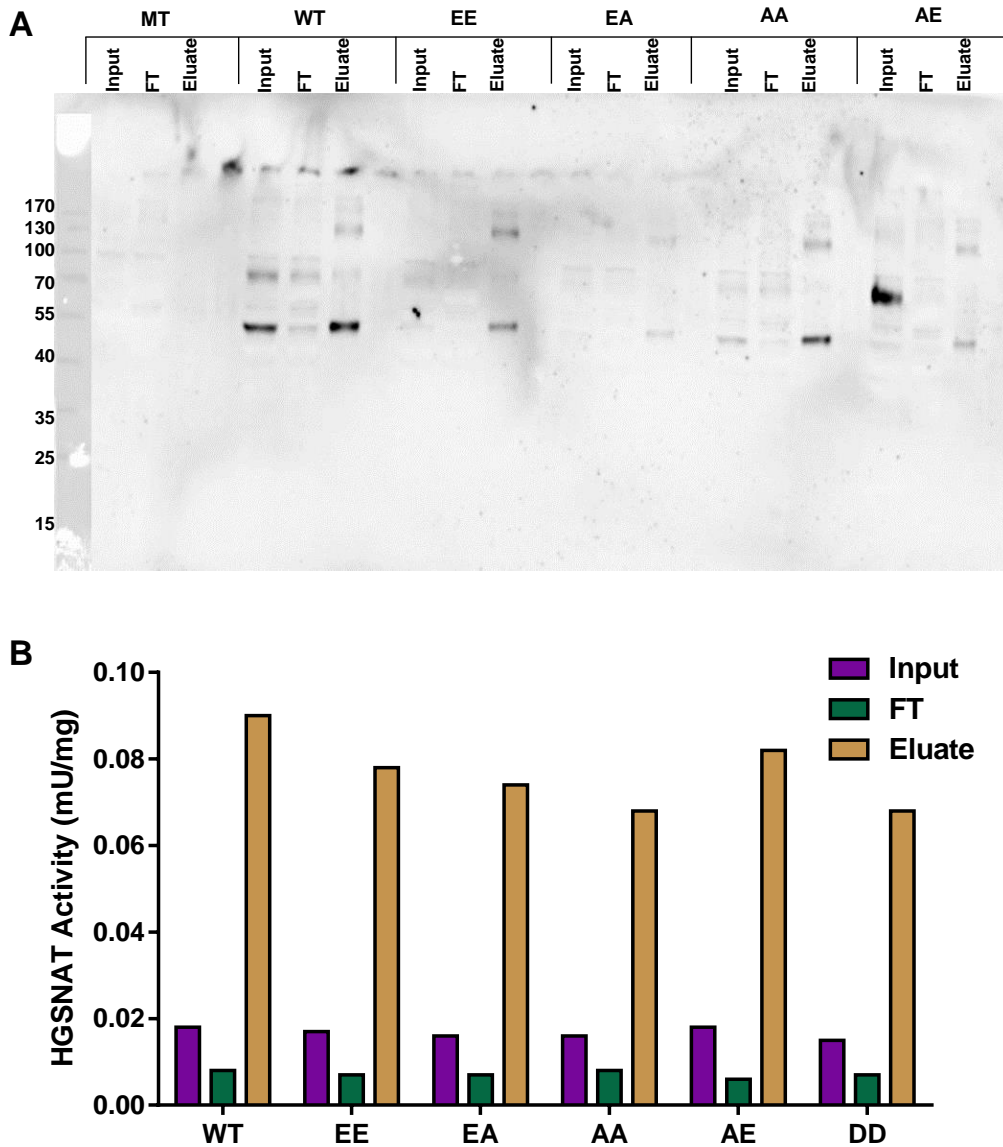
Figure 3.30: Intracellular distribution of WT and phosphomutants of HGSNAT.
A. An equal percentage of input, flow-through (FT) and eluate fractions of lysosome enrichment from MT (mock-transfected), WT (wild-type) and double mutants transfected HGSNAT KO cells were used for Western blotting with anti-FLAG antibody (1:5000). B. HGSNAT activity assay was performed with an equal amount of proteins from the same samples. Activities were expressed as mU/mg of total protein in the sample.

3.5 Identification of HGSNAT and its interaction partners

Since HGSNAT activity assay didn't show any difference within the phospho-mutants, it was speculated that probably any activating protein or interaction partner influences HGSNAT phosphorylation. Therefore, with the aim of detecting HGSNAT interaction partners, two SILAC-based co-immunoprecipitation experiments were performed: in one experiment, endogenous HGSNAT and its interacting proteins were precipitated using anti-HGSNAT antibody and in another one, overexpressed HGSNAT along with its interaction partners were precipitated using myc-Trap beads.

In both of the techniques, first HGSNAT and its interaction partners were precipitated by using an affinity purification method. Then precipitated proteins were identified by mass spectrometry using an Orbitrap Fusion Lumos mass spectrometer.

3.5.1 Detection of proteins co-immunoprecipitating with endogenous HGSNAT

Proteins interacting with HGSNAT were precipitated using anti-HGSNAT antibody. In this approach, SILAC heavy labeled WT HEK293 cells and light labeled HGSNAT KO HEK293 cells were used, where endogenous HGSNAT and its interaction partners were precipitated from heavy labeled WT HEK293 cells whereas, HGSNAT KO cells worked as a control due to the absence of HGSNAT gene. This approach helps to identify real interacting proteins and allows separating proteins binding unspecifically to the affinity matrix (background) from interacting proteins by heavy to light ratios. Background proteins will show heavy and light peptide signals of equal intensity, whereas, true interactors of HGSNAT should have significantly higher heavy peptide signals.
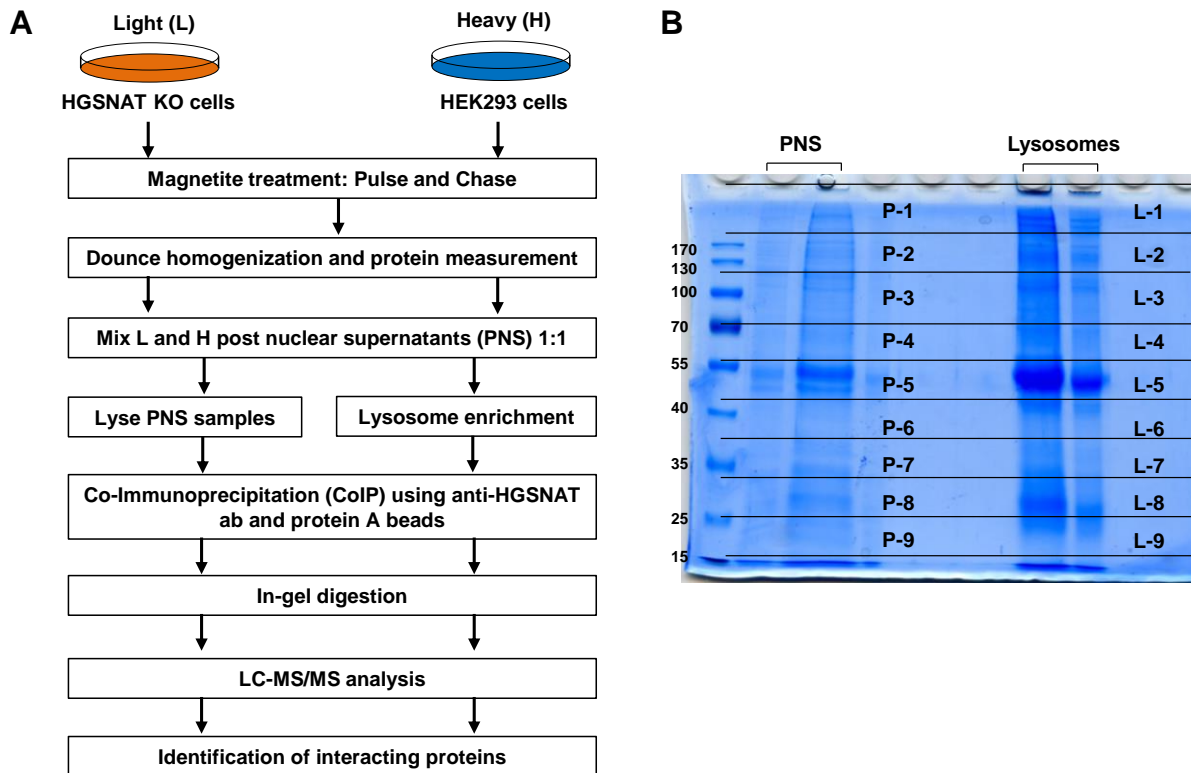
Figure 3.31: Workflow of HGSNAT SILAC CoIP with anti-HGSNAT antibody.
A. Schematic workflow for HGSNAT CoIP using PNS and lysosome samples. From SILAC heavy (WT HEK293) and light (HGSNAT KO) cells, post-nuclear supernatant (PNS) and enriched lysosome samples were co-immunoprecipitated using rabbit polyclonal anti-HGSNAT antibody and protein A beads. B. Co-immunoprecipitated proteins were separated on SDS-PAGE and all nine fractions were excised from the gel containing PNS and lysosome samples and subjected to LC-MS/MS analysis.

3.5.1.1   SILAC CoIP to detect HGSNAT interaction partners in PNS

Post-nuclear supernatants (PNS) samples were processed as shown in Figure 3.31 (A) followed by in-gel digestion of 9 fractions (Figure 3.31 (B)) and then analyzed by mass spectrometry. Generated raw data were further analyzed by MaxQuant and Perseus. Protein identification and quantification were performed by MaxQuant whereas, in Perseus, statistical analysis was performed using one-sample t-test (shown as volcano plots) and Pearson's correlation test (shown as multiple scatter plot). The results are depicted in Figure 3.32 and Table 3.5.

A total of 2086 protein groups were identified and quantified in all three replicates of PNS samples. Correlation between three biological replicates was assessed using Pearson's correlation coefficient (r) (Puth et al. 2014) and determined as 0.885, 0.869 and 0.871 between replicate 1&2, 1&3 and 2&3 respectively (Figure 3.32 (A)). This relatively high Pearson's correlation coefficient indicates high reproducibility among the replicates. It's also observed that more than 84 % proteins are common in all three replicates whereas only 6.4 % proteins were detected in only one replicate, as shown in Venn diagram (Figure 3.32 (B)). After performing a one-sample t-test, log2-transformed average H/L ratio (log2 fold change) of all identified and quantified protein groups was depicted on a volcano plot against their -log10-transformed p-value (-log10 p-value) (Figure 3.32 (C)). The protein groups enriched in the heavy labeled (bait-containing) samples were expected to have a high log2 fold change (=enrichment) in combination with a high -log10 p-value (=significance) which appears on the upper right part of the plot. In contrast, the background proteins present in both samples (control and bait) were expected to cluster around log2 fold change of '0'. For SILAC CoIP experiments presented in this thesis, the threshold for a protein group to be considered significantly enriched was set to log2 fold change ≥ 1 (fold change ≥ 2) and -log10 p-value > 1.301 (p-value < 0.05).

Statistical one-sample t-test shows that out of a total of 2086 proteins 91 proteins were significantly enriched in PNS samples with a log2 fold change ≥ 1 together with a good level of confidence (-log10 p-value > 1.301) (Figure 3.32 (C)). These protein groups sorted by their log2 fold change are shown in Table 3.5.
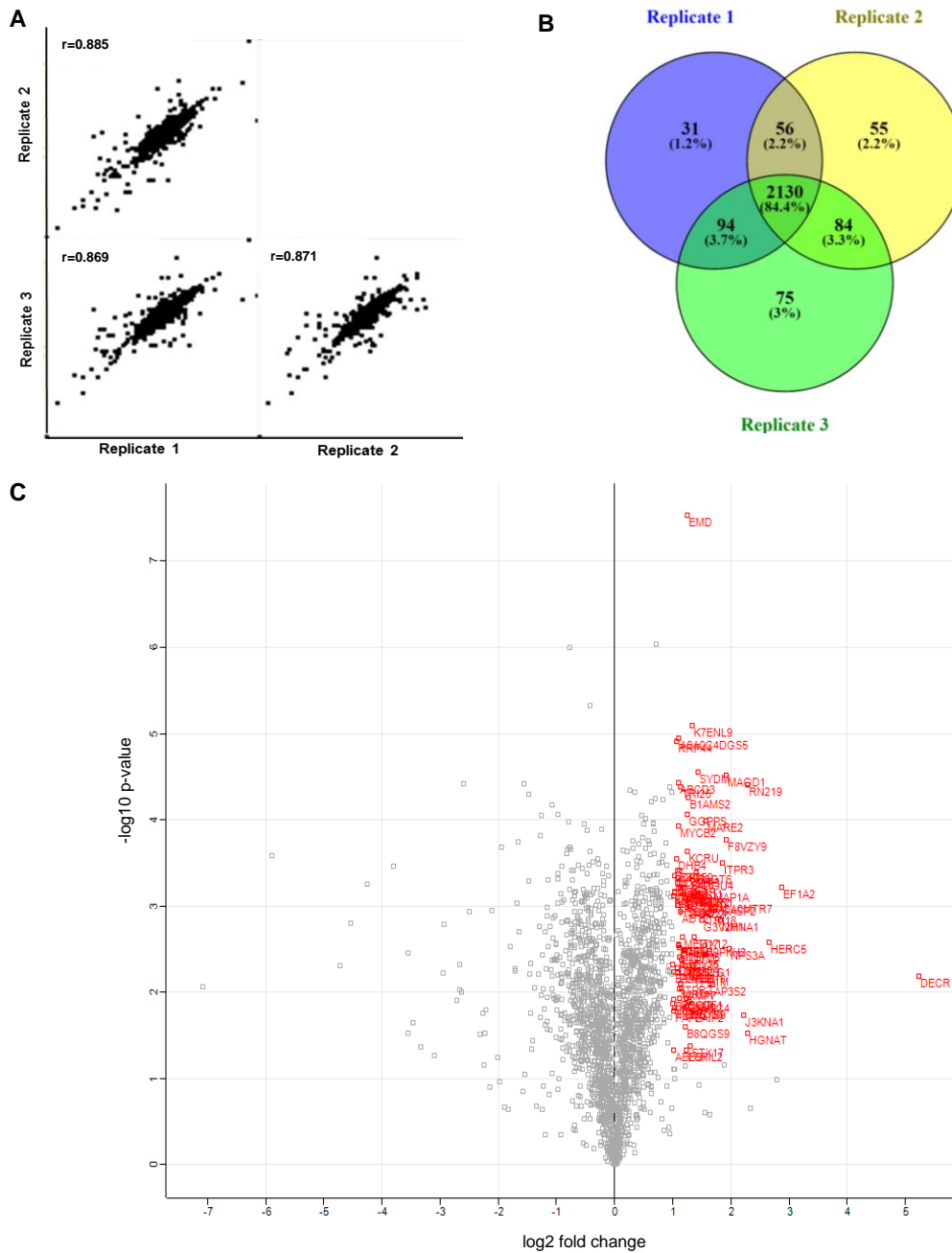
Figure 3.32: HGSNAT interaction partner analysis from co-immunoprecipitated PNS samples.
A. Experimental reproducibility for total proteins obtained from 3 biological replicates of HGSNAT CoIP of PNS samples. MS raw data were analyzed by MaxQuant and 'ProteinGroups' output of MaxQuant was further assessed by Perseus. Correlation between the three biological replicates was determined using Pearson's correlation coefficient (r). B. Venn diagram showing overlap between the total proteins found in three biological replicates. Venn diagrams were created online using open-access tool Venny (Oliveros 2007). C. Volcano plot of all protein groups identified and quantified in all 3 replicates. After performing a one-sample t-test against 0, resulting -log10 transformed p-values (-log10 p-value) of each protein groups were plotted against their log2 transformed average H/L ratios (log2 fold change). Significantly enriched protein groups having a log2 fold change ≥ 1 (fold change ≥ 2) and –log10 p-value > 1.301 (p-value < 0.05) are highlighted in red.

Table 3.5: Protein groups significantly enriched in PNS samples after pulldown from HEK293 and HGSNAT KO cells.

Shown are all protein groups identified with a log2 fold change ≥ 1 and -log10 p-value >1.301. Protein groups are listed by their respective gene names and sorted by their log2 fold change in descending order. Protein groups associated with lipid metabolism are highlighted in red.

No.	Gene Names	log2 fold change	-log10 p-value	Unique peptides	No.	Gene Names	log2 fold change	-log10 p-value	Unique peptides
1	DECR	5.2337	2.1867	6	47	CCD51	1.2355	1.9348	8
2	EF1A2	2.8743	3.2149	4	48	MED30	1.2282	1.8340	2
3	KI13B	2.7861	0.9815	2	49	TCF25	1.2243	2.4816	8
4	HERC5	2.6571	2.5821	5	50	RPC6	1.2238	1.1407	3
5	IFT74	2.3289	0.6541	4	51	PKP2	1.2147	1.6034	3
6	HGSNAT	2.2942	1.5271	2	52	RPR1A	1.1996	3.1441	8
7	RN219	2.2858	4.4074	21	53	IDH2	1.1792	3.2155	4
8	KIF5A	2.2150	1.7280	2	54	MED17	1.1595	2.6463	6
9	NPS3A	1.9734	2.5043	3	55	DSG2	1.1589	2.3274	7
10	MAGD1	1.9233	4.5128	3	56	MRM1	1.1568	2.0483	2
11	IRS4	1.8893	1.1618	14	57	MED20	1.1556	2.3771	4
12	ITPR3	1.8498	3.5044	16	58	HAX1	1.1370	3.2157	7
13	NMNA1	1.8162	2.8418	3	59	AIFM1	1.1368	3.1664	17
14	GASP2	1.7956	3.0184	2	60	ADT2	1.1345	2.9366	5
15	MAP1A	1.6925	3.1808	27	61	TRI25	1.1305	4.3786	16
16	AP3S2	1.6776	2.0835	2	62	ITPR2	1.1252	2.1016	13
17	CHMP2B	1.6305	0.5774	2	63	PEX13	1.1196	1.9262	4
18	TRI18	1.5629	2.9144	9	64	DPM1	1.1139	2.5233	7
19	MARE2	1.5611	3.9953	4	65	CA050	1.1129	3.4150	4
20	SERPINH1	1.5587	2.5432	3	66	ATPG	1.1066	2.5597	4
21	AMOT	1.5491	0.6112	20	67	MYCB2	1.1048	3.9259	53
22	RNF213	1.5357	3.0440	9	68	GOLGA2	1.1009	4.9430	5
23	UBIP1	1.4450	0.9189	2	69	ABCD3	1.0990	4.4256	17
24	P5CR1	1.4418	3.1156	12	70	ACADM	1.0958	2.5382	9
25	SYDM	1.4380	4.5535	20	71	WDR6	1.0943	3.2039	11
26	B3GT6	1.3997	3.3876	4	72	HSDL2	1.0905	3.0109	8
27	P5CR2	1.3707	3.1165	5	73	RT31	1.0851	3.4120	7
28	STX12	1.3608	2.6434	3	74	IF4A3	1.0838	3.2633	21
29	BAIP2	1.3543	1.7782	6	75	PLOD3	1.0823	3.3313	10
30	GSLG1	1.3473	2.3151	6	76	H10	1.0750	2.2858	2
31	XPP3	1.3417	3.3574	2	77	RRP44	1.0676	4.9060	17
32	RMC1	1.3282	5.0872	2	78	DHB4	1.0672	3.5446	32
33	TFB1M	1.3269	2.1962	6	79	GCSP	1.0398	0.9980	2
34	ACOX1	1.3051	2.9997	7	80	DOC11	1.0382	3.3537	15
35	STX17	1.3044	1.3753	3	81	TCTP	1.0352	3.1524	6
36	DNAJB6	1.2899	1.8982	4	82	MED6	1.0326	1.8252	3
37	TMLH	1.2877	2.3151	2	83	CYBP	1.0316	3.1682	13
38	CTNNB1	1.2769	3.3106	6	84	CDC123	1.0273	1.1047	2
39	MED18	1.2768	2.2372	4	85	B4DLN1	1.0180	1.9166	3
40	CTNL1	1.2721	3.0856	2	86	FAF2	1.0159	1.7767	2
41	SEPT6	1.2693	4.2613	3	87	ABC3F	1.0125	1.3229	2

N o.	Gene Names	log2 fold change	-log10 p-value	Unique peptides
42	KCRU	1.2581	3.6357	14
43	DYHC2	1.2535	2.9826	8
44	EMD	1.2524	7.5362	16
45	GGPPS	1.2519	4.0663	12
46	CNOT8	1.2396	1.3232	5

No.	Gene Names	log2 fold change	-log10 p-value	Unique peptides
88	MED8	1.0104	3.1219	5
89	SCMC1	1.0090	1.1131	3
90	PISD	1.0065	1.8711	5
91	ESYT2	1.0044	2.3172	8

3.5.1.2   SILAC CoIP to identify HGSNAT interaction partners in lysosomes

In this experiment, lysosome samples were prepared as shown in Figure 3.31 (A) and after in-gel digestion of all co-immunoprecipitated samples (Figure 3.31 (B)), mass spectrometry analysis was performed followed by data analysis using MaxQuant and Perseus.

In the MaxQuant analysis, 2576 total proteins were detected with valid values in all three biological replicates. They were evaluated for their correlation between all replicates and Pearson's correlation coefficient (r) shows 0.886, 0.888 and 0.872 between replicate 1&2, 1&3 and 2&3 respectively as shown in Figure 3.33 (A), which indicates high reproducibility among the replicates. It's also observed that around 86 % of the proteins were common in all three replicates whereas only 5.9 % of the proteins were detected in only one replicate, as shown in Venn diagram Figure 3.33 (B). Next, the log2-transformed average H/L ratio (log2 fold change) of protein groups from one-sample t-test was depicted on a volcano plot against their -log10 p-value (Figure 3.33 (C)). Statistical one-sample t-test demonstrated that, out of a total 2576 proteins, 76 proteins were significantly enriched in immunoprecipitated lysosomes samples with a log2 fold change ≥ 1 in combination with a significant level of confidence (-log10 p-value > 1.301) (Figure 3.33 (C)). These protein groups sorted by their log2 fold change are given in Table 3.6.

From these 76 significantly enriched proteins, 4 proteins were found to be related to the heparan sulfate metabolism: HGSNAT, Beta-1,4-galactosyltransferase 7 (B4GALT7), Beta-1,3-galactosyl-transferase 6 (B3GALT6) and Glypican-3 (GPC3) and they are marked as bold in Table 3.6. Additionally, 12 proteins were detected to be associated with lipid metabolism which is marked as red in Table 3.6.
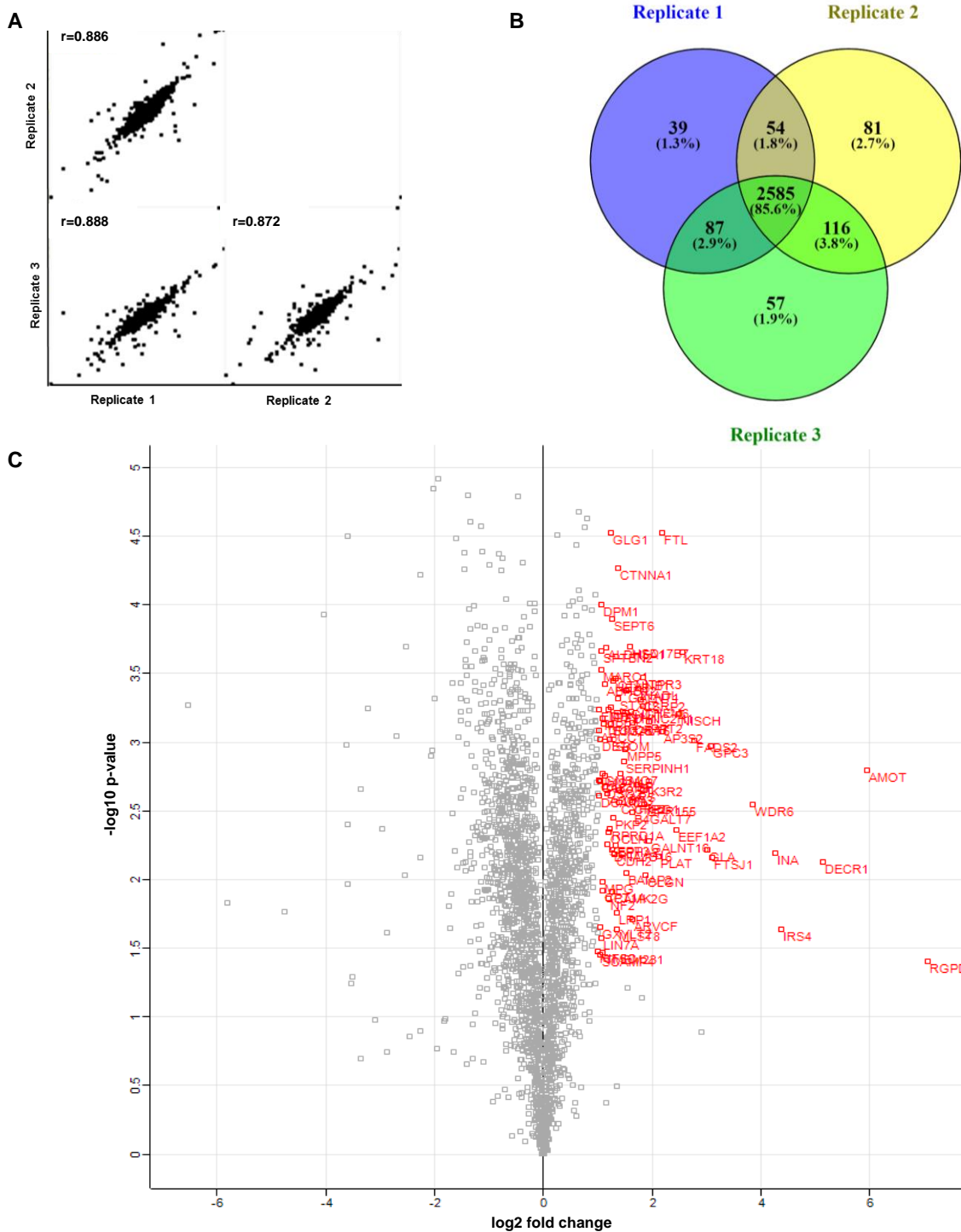
Figure 3.33: HGSNAT interaction partner analysis from co-immunoprecipitated lysosomes.
A. Experimental reproducibility for total proteins obtained from 3 biological replicates. The correlation was determined using Pearson's correlation coefficient (r). B. Venn diagram showing overlap between the total proteins found in three biological replicates (Oliveros 2007). C. Volcano plot of all protein groups identified and quantified in all 3 replicates. Significantly enriched protein groups having a log2 fold change ≥ 1 (fold change ≥ 2) and -log10 p-value > 1.301 (p-value < 0.05) are highlighted in red.

Table 3.6: Protein groups significantly enriched in lysosome samples.
Shown are all protein groups identified with a log2 fold change ≥ 1 and -log10 p-value >1.301. Protein groups are listed by their respective gene names and sorted by their log2 fold change in descending order. Protein groups associated with lipid metabolism are highlighted in red and with heparan sulfate metabolism in **bold.**

No.	Gene Names	Log2 fold change	-log10 p-value	Unique peptides	No.	Gene Names	Log2 fold change	-log10 p-value	Unique peptides
1	AMOT	5.9417	2.7934	45	39	KIF13B	1.3409	0.4919	2
2	DECR1	5.1432	2.1271	7	40	EPHA2	1.3357	2.2489	2
3	IRS4	4.3726	1.6393	15	41	CTNNB1	1.3341	3.4651	18
4	INA	4.2616	2.1916	24	42	CDH2	1.3087	2.1862	8
5	WDR6	3.8475	2.5448	17	43	HEXB	1.2798	3.4483	6
6	FTSJ1	3.1073	2.1565	2	44	PKP2	1.2777	2.4521	10
7	**GPC3**	**3.0840**	**2.9714**	**2**	45	CAMK2G	1.2707	1.9072	4
8	GLA	3.01720	2.2191	5	46	SEPT6	1.2604	3.8955	4
9	**HGSNAT**	**2.9103**	**0.8927**	**4**	47	**B3GALT6**	**1.2476**	**3.1332**	**8**
10	FADS2	2.7819	3.0159	4	48	GLG1	1.2457	4.5212	18
11	NISCH	2.4824	3.2056	6	49	ERGIC1	1.2359	3.2498	7
12	EEF1A2	2.4502	2.3597	3	50	STOM	1.2250	3.0196	9
13	FTL	2.1765	4.5275	7	51	RPRD1A	1.2230	2.3708	2
14	PLAT	2.1218	2.1648	3	52	NF2	1.2061	1.8641	11
15	GALNT16	1.9444	2.2842	2	53	DBN1	1.1973	3.2373	24
16	CLGN	1.8665	2.0329	2	54	ACOX3	1.1807	2.6280	3
17	ITPR3	1.8361	3.4673	22	55	SEPT3	1.1714	2.2570	3
18	THEM6	1.8327	3.2584	7	56	ARGLU1	1.1604	0.3770	2
19	CSRP2	1.7924	3.3081	7	57	ALDH5A1	1.1595	3.6856	5
20	INADL	1.7508	3.3943	6	58	MON1B	1.1385	2.7576	11
21	PIK3R2	1.7320	2.6935	2	59	DSG2	1.1331	2.6785	21
22	CRAT	1.636	3.1402	13	60	ABHD12	1.1227	3.4252	6
23	**B4GALT7**	**1.6281**	**2.4897**	**4**	61	TRIM25	1.1011	3.1352	6
24	DYNC2H1	1.6219	3.2188	29	62	TMEM231	1.0982	1.4679	2
25	HSD17B7	1.5930	3.6951	3	63	CPT1A	1.0948	1.9217	4
26	NOS1AP	1.5471	1.2084	3	64	MPG	1.0867	1.9809	3
27	BAIAP2	1.5346	2.0462	2	65	MPP2	1.0791	3.1746	5
28	GALNT1	1.5205	3.3729	7	66	LIN7A	1.0771	1.5751	3
29	MPP5	1.5116	2.9488	20	67	MARC1	1.0685	3.5236	8
30	SERPINH1	1.4798	2.8606	8	68	DPM1	1.0652	4.0016	13
31	FTH1	1.4585	3.2244	2	69	SPTBN2	1.0623	3.6646	33
32	LMO7	1.4120	2.7714	4	70	GXYLT2	1.0552	1.6551	2
33	CYFIP2	1.4028	2.5614	8	71	C18orf8	1.0541	2.7109	14
34	YARS	1.3935	2.6507	11	72	DIS3	1.0349	3.0188	4
35	STX12	1.3725	3.3179	9	73	DCAKD	1.0277	2.6083	8
36	CTNNA1	1.3701	4.2641	26	74	CACYBP	1.0246	2.7247	11
37	LRP1	1.3589	1.7563	3	75	EMD	1.0232	3.2374	16
38	MLST8	1.3409	1.6400	3	76	ABCC1	1.0195	3.0847	5

Finally, enriched proteins from both PNS and lysosome co-immunoprecipitated samples were analyzed for common proteins and 13 proteins were found enriched with a log2 fold change ≥ 1 in both samples.

Table 3.7: 13 common proteins between PNS and Lysosome samples

Protein Names	Gene Names
Angiomotin	AMOT
Insulin receptor substrate 4	IRS4
WD repeat-containing protein 6	WDR6
Heparan-α-glucosaminide N-acetyltransferase	HGSNAT
Inositol 1,4,5-trisphosphate receptor type 3	ITPR3
Serpin H1	SERPINH1
Syntaxin-12	STX12
Catenin beta-1	CTNNB1
Plakophilin-2	PKP2
Septin 6	SEPT6
Desmoglein-2	DSG2
Dolichol-phosphate mannosyltransferase subunit 1	DPM1
Emerin	EMD

3.5.2 Identification of proteins co-immunoprecipitating with myc-tagged HGSNAT

In this study, overexpressed myc-tagged HGSNAT was used as a bait to identify proteins co-immunoprecipitating with it using myc-trap agarose beads. To that purpose, SILAC heavy (H) and light (L) labeled HGSNAT KO cells were transfected with WT HGSNAT and medium (M) labeled HGSNAT KO cells were transfected with SCARB2 (Figure 3.34 (A)). SCARB2 is a lysosomal transmembrane protein which was used as a negative control in this study. Since our mass spectrometry data detected HGSNAT as phosphorylated after U18666A treatment, therefore to observe changes in HGSNAT interaction partners in this study U18666A was used as a stimulus. Therefore, 24 h post-transfection, heavy and medium labeled cells were treated with U18666A for 24 h and light cells were supplied with DMSO (control). In order to identify weak and transient interactors, PFA-crosslinking was performed in between cell harvesting and lysis. Then, myc-Trap agarose beads were employed to isolate transiently expressed HGSNAT along with its interaction partners from SILAC labeled HGSNAT KO cell lysates. Captured proteins were eluted from the beads after boiling at 40 °C for 10 min with 2x SDS sample buffer and resolved using 10 % SDS-PAGE. In order to keep disulfide bridges intact, no reducing agent was used in the sample buffer. Subsequently, proteins were run 1 cm on the separating gel (Figure 3.34 (B)) because of lower complexity samples and were then subjected to tryptic in-

gel digestion. Digested peptides after StageTip desalting were analyzed on an Orbitrap Fusion Lumos mass spectrometer and evaluated further using MaxQuant and Perseus.
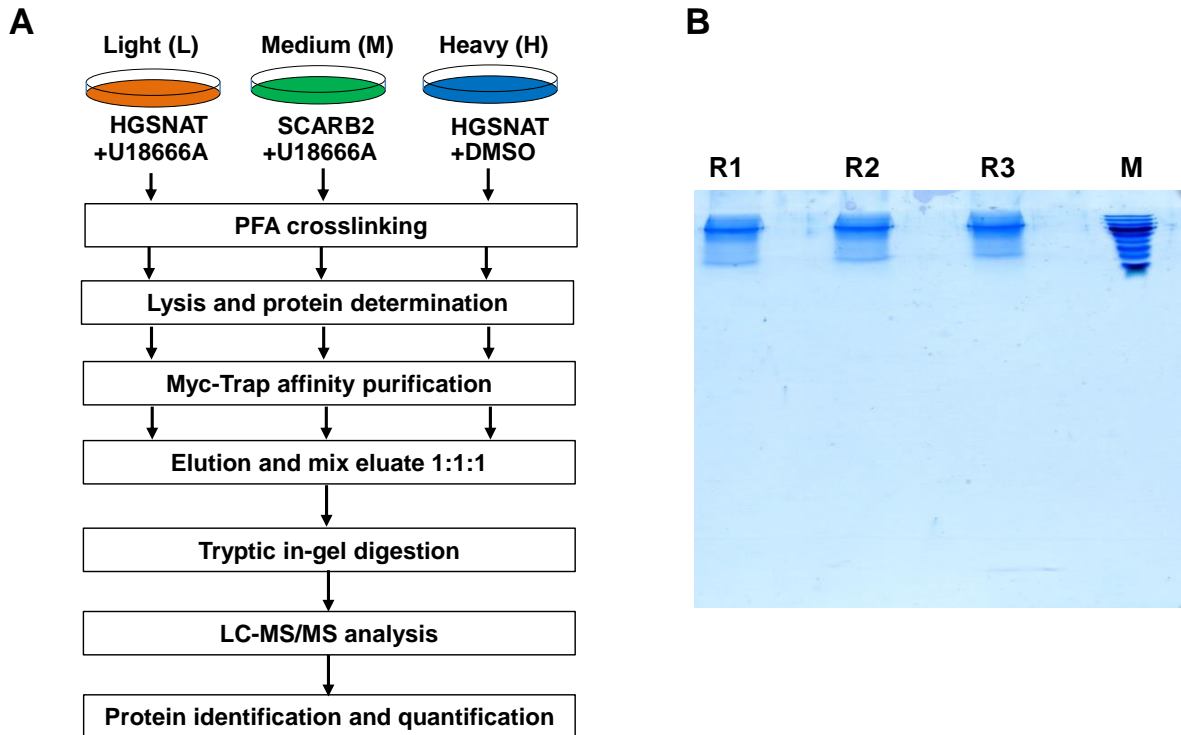
Figure 3.34: Workflow for co-immunoprecipitation of HGSNAT interaction partners.
A. SILAC labeled HGSNAT KO HEK293 cells were transfected and treated with U18666A as above. 24 hours post-treatment, cells were PFA cross-linked and lyzed with 1 % NP-40 buffer. Then they were subjected to affinity purification using Myc-Trap agarose beads. Eluate was used for tryptic in-gel digestion followed by mass spectrometry analysis. B. The eluate from three biological replicates was run in SDS-PAGE until 1 cm, gels were excised and samples were prepared for mass spectrometry.

The results of H/L values were expected to reflect the effect of U18666A treatment on HGSNAT interaction partners and H/M values were expected to show potential candidates for HGSNAT interaction partners, whereas M/L values were considered as background. Using MaxQuant, a total of 1722 proteins were identified as well as quantified with valid values in all three biological replicates. Pearson correlation coefficient (r) for each condition varies 0.7 to 0.903 indicating high correlation among replicates (Figure 3.35). Moreover, the Venn diagram depicted that around 77 % of the proteins were found common among all three replicates of H/L, H/M and M/L conditions (Figure 3.35). Subsequently, a statistical one-sample t-test was performed against '0' to detect enriched proteins in each condition. The results from this one-sample t-test were shown as volcano plots in which a log2-fold change (H/L or H/M) was depicted against respective -log10 p-value (Figure 3.36 and Figure 3.37). At this

point, it needs to be mentioned that the proteins found as regulated on M/L samples (background) were filtered out from the regulated H/L and H/M values before plotting. The statistical one-sample t-test found that 29 proteins were significantly enriched on HGSNAT treated (H) versus HGSNAT control (L) samples (Figure 3.36) whereas 21 proteins were enriched in HGSNAT (H) versus SCARB2 samples (Figure 3.37). As can be seen in Table 3.8, there were 29 enriched proteins found in the H/L samples whereas Table 3.9 illustrates the 21 enriched proteins detected in H/M samples, and they were sorted by their log2-fold change in descending order.
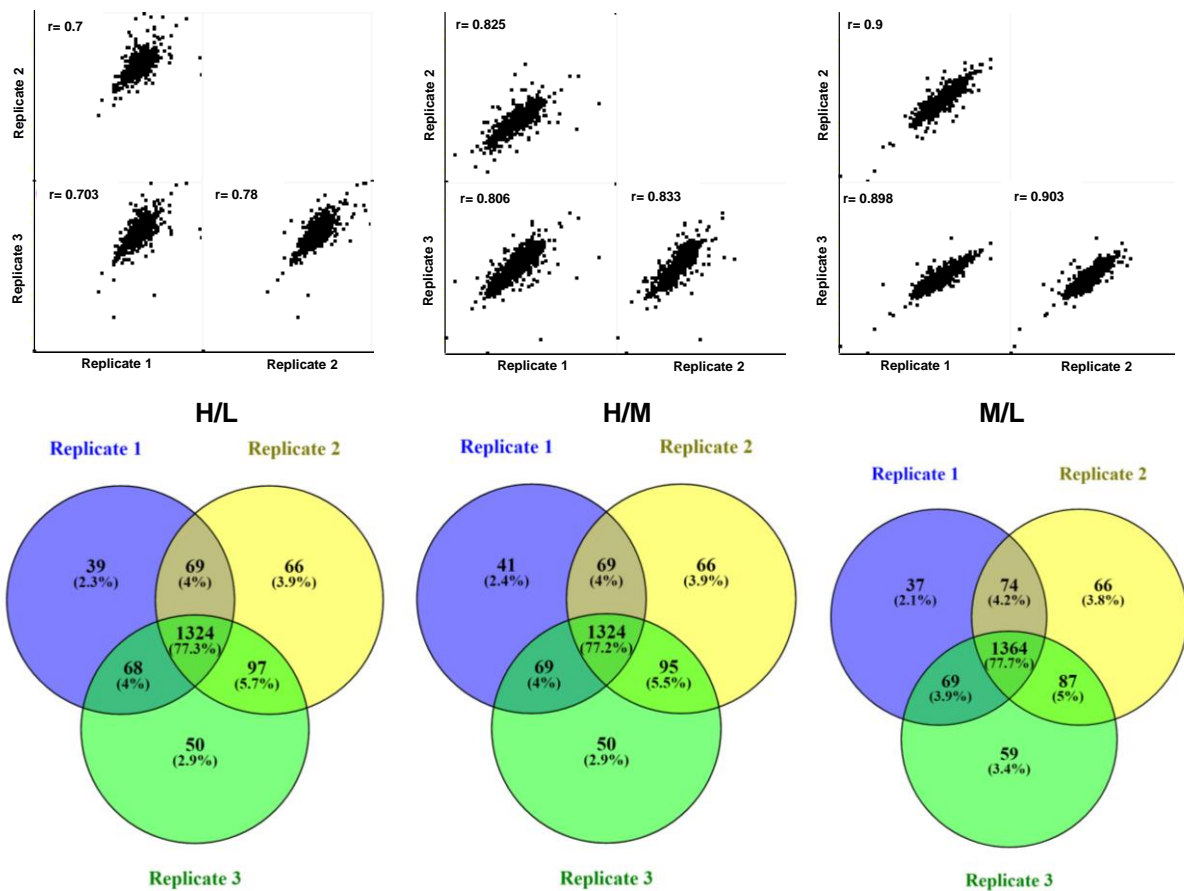
Figure 3.35: Comparing the reproducibility among different replicates of HGSNAT CoIP experiment.
MS raw data from CoIP experiment were analyzed by MaxQuant and 'ProteinGroups' output of MaxQuant was further assessed by Perseus. Correlation between the three biological replicates was calculated using Pearson's correlation coefficient (r). Venn diagram is showing overlap between the total proteins found in three biological replicates. Venn diagrams were created online using open-access tool Venny (Oliveros 2007).
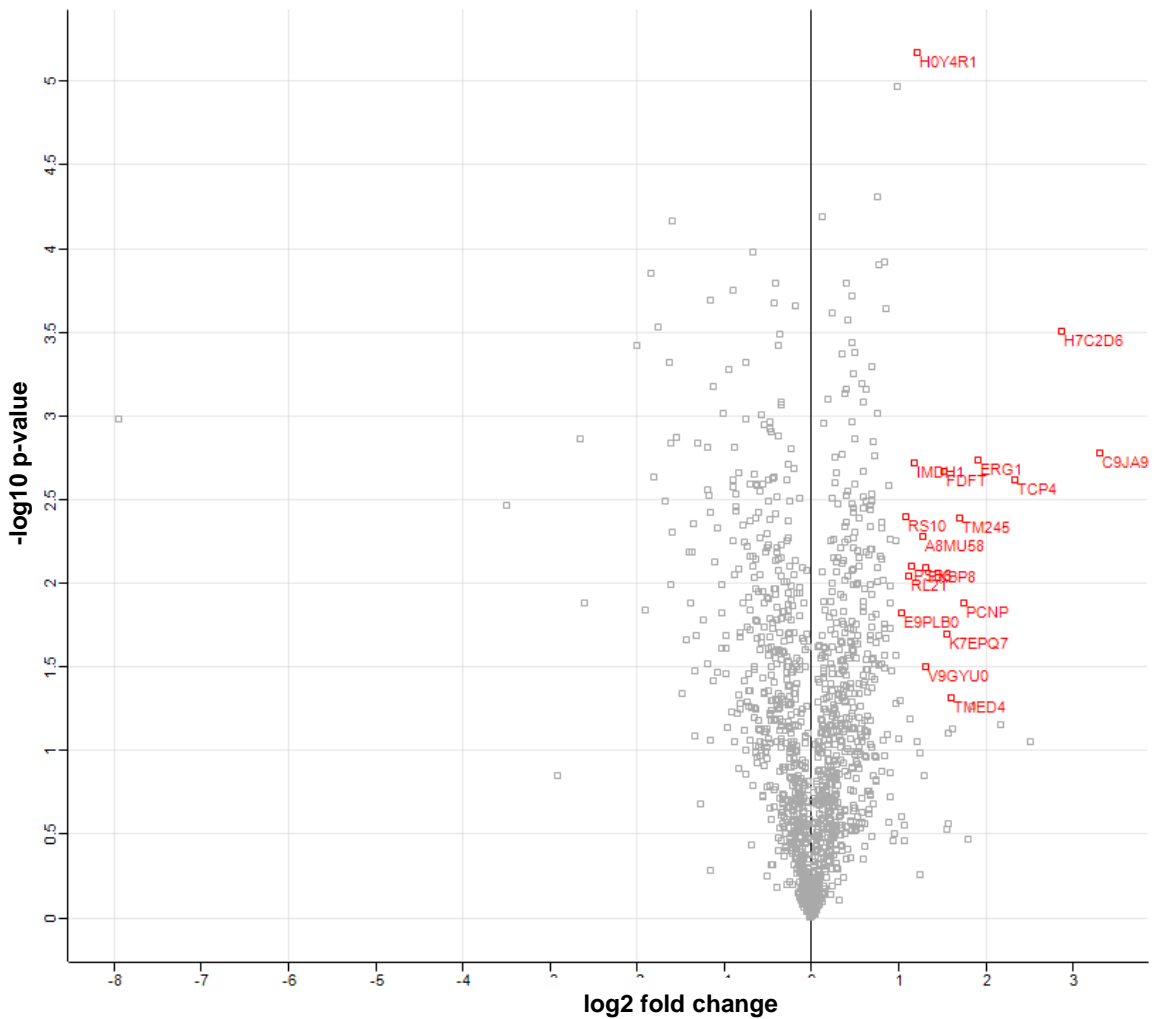
Figure 3.36: Volcano plot showing differentially enriched proteins on U18666A treated HGSNAT CoIP samples (HGSNAT treated vs. HGSNAT control).
Detected proteins in the Myc-Trap based co-immunoprecipitation dataset were searched for the differentially enriched proteins upon addition of U18666A to the culture medium. Proteins detected in all three biological replicates with p-values less than 0.05 and log2 transformed SILAC ratios higher than 1 (the arbitrary cut off) were considered as differentially enriched. Log2 fold changes (x-axis) of proteins were plotted against -log10 p-value (y-axis) using Perseus (Tyanova et al. 2016).

In samples transfected with HGSNAT and treated with U18666A (H/L), the enriched proteins contain four ribosomal proteins (RL21, RS10, RS25, and RPL15) and two proteins that are related to the cholesterol metabolism (Squalene monooxygenase (ERG1) and Squalene synthase (FDFT)), whereas in H/M samples, mostly ribosomal proteins are enriched.

Table 3.8: Protein groups significantly enriched in WT HGSNAT transfected (U18666A vs. DMSO) and PFA cross-linked HGSNAT KO cells after pulldown with Myc-Trap beads.

Shown are protein groups identified in all three biological replicates with a log2 fold change ≥ 1 and -log10 p-value >1.301. Protein groups are given by their respective gene names and sorted by their log2 fold change in descending order.

No.	Gene Names	log2 fold change (H/L)	-log10 p-value	Unique peptides	No.	Gene Names	log2 fold change (H/L)	-log10 p-value	Unique peptides
1	MPC2	2.5024	1.0519	2	16	TM165	1.2429	0.2564	2
2	ITM2B	2.1637	1.1586	3	17	IMDH2	1.2125	5.1756	23
3	ERG1	1.9098	2.7403	11	18	RPS16	1.2112	1.0578	4
4	SNX3	1.8481	1.2641	2	19	IMDH1	1.1856	2.7177	15
5	IDHP	1.7933	0.4713	6	20	PSB6	1.1437	2.0983	3
6	PCNP	1.7545	1.8858	4	21	RPL15	1.1243	1.1880	3
7	ITM2C	1.6127	1.1333	4	22	RL21	1.1208	2.0463	2
8	TMED4	1.6064	1.3120	4	23	RS10	1.0732	2.3994	7
9	PGP	1.5740	1.1081	3	24	PMGT2	1.0668	0.4630	2
10	ALG5	1.5729	0.5663	8	25	SPTC2	1.0662	0.5521	5
11	ZCCHV	1.5512	0.5302	2	26	RBM4B	1.0339	1.8242	2
12	FDFT	1.5235	2.6685	11	27	EI24	1.0334	0.6021	2
13	FKBP8	1.3104	2.0979	3	28	I2BP2	1.0170	1.2985	2
14	YIPF5	1.2995	0.8541	2	29	RS25	1.0031	1.0670	3
15	AIMP2	1.2692	2.2794	2					

Table 3.9: Protein groups significantly enriched in WT HGSNAT vs. WT SCARB2 transfected and PFA cross-linked HGSNAT KO cells after pulldown with Myc-Trap beads.

Shown are protein groups identified in all three biological replicates with a log2 fold change ≥ 1 and -log10 p-value >1.301. Protein groups are given by their respective gene names and sorted by their log2 fold change in descending order.

No.	Gene Names	log2 fold change (H/M)	-log10 p-value	Unique peptides	No.	Gene Names	log2 fold change (H/M)	-log10 p-value	Unique peptides
1	HGSNAT	6.2831	1.7996	13	12	RPS16	1.2928	1.1173	4
2	IDHP	1.9805	0.5169	6	13	RPL23A	1.2450	1.8890	3
3	PCNP	1.6953	1.8916	4	14	RL4	1.1128	2.4109	6
4	PGP	1.6254	0.7808	3	15	RLA0	1.1110	2.5779	8
5	QTRD1	1.6029	1.4939	3	16	RL12	1.0868	2.2032	4
6	DJB12	1.5533	1.5219	6	17	VTA1	1.0684	0.6277	3
7	PSB6	1.5183	2.1547	3	18	RL5	1.0564	1.5042	5
8	SNX3	1.5060	0.9851	2	19	RS3A	1.0370	2.7213	9
9	MAP4	1.5024	0.9112	4	20	PAIRB	1.0125	2.8499	10
10	ZCCHV	1.3951	0.4284	2	21	RL3	1.0037	1.5998	10
11	RL27	1.2929	2.6811	3					
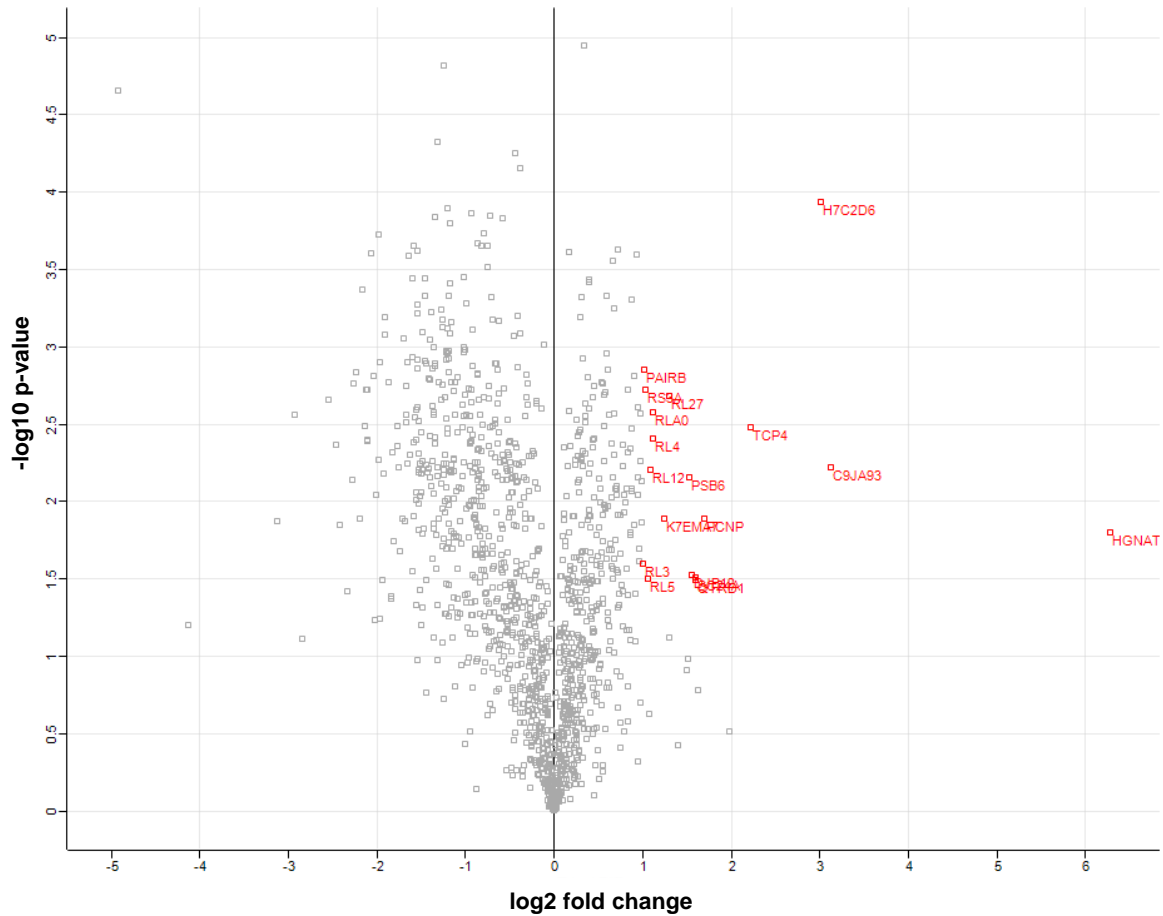
Figure 3.37: Volcano plot showing differentially enriched proteins on transiently expressed HGSNAT CoIP samples (HGSNAT vs SCARB2).
Detected proteins in the Myc-Trap based co-immunoprecipitated dataset were searched for the differentially enriched proteins on the HGSNAT transfected HGSNAT KO cells in compare to SCARB2 transfected HGSNAT KO cells. Proteins detected in at all three biological replicates with p-values less than 0.05 and log2 transformed SILAC ratios higher than 1 (the arbitrary cut off) were considered as differentially enriched. Log2 fold changes (x-axis) of proteins were plotted against –log10 p-value (y-axis) using Perseus (Tyanova et al. 2016).

Furthermore, HGSNAT protein was assessed for interaction partners using two online databases: STRING (https://string-db.org/) v-11.0 (Szklarczyk et al. 2019) and COXPRESdb (http://coxpresdb.jp) (Obayashi et al. 2019). STRING consortium database identified ten functional partners for HGSNAT (Figure 3.38 (A)) and COXPRESdb database detected six genes which are directly connected to HGSNAT (Figure 3.38 (B)). For analysis, both database use high-throughput data from large scale studies. Unfortunately, no similarities were found while comparing these data with the enriched interaction partner's dataset for HGSNAT in this study.

Finally, enriched proteins from both conditions (H/L and H/M) were searched for common proteins and there were 7 proteins that could be found common in-between

these 2 conditions. However, no common proteins were found in-between HGSNAT CoIP and myc-trap CoIP experiments.

Table 3.10: Seven common proteins between H/L and H/M samples.

Protein Names	Gene Names
Sorting nexin-3	SNX3
Isocitrate dehydrogenase [NADP], mitochondrial	IDHP
PEST proteolytic signal-containing nuclear protein	PCNP
Phosphoglycolate phosphatase	PGP
Zinc finger CCCH-type antiviral protein 1	ZCCHV
40S ribosomal protein S16	RPS16
Proteasome subunit beta type-6	PSB6
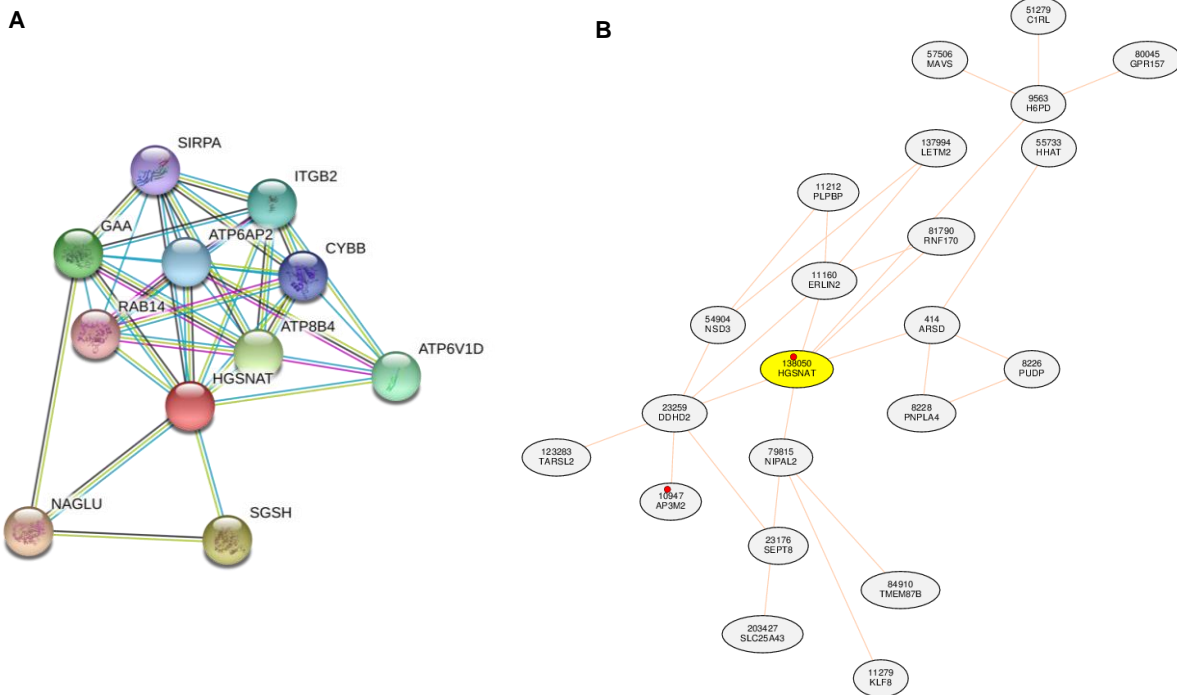
Figure 3.38: Predicted HGSNAT interaction networks.
A. HGSNAT protein was searched for interaction partners in STRING consortium database, v-11.0 (https://string-db.org/). Edges represent protein-protein interaction with different colors. Known interactions are marked as magenta and pink, where magenta shows interactors from the curated database and pink shows experimentally determined interactors. Predicted interaction marked as green, red and blue, where green shows interactors in gene neighborhood, red: gene fusions and blue: gene co-occurrence. B. HGSNAT was searched for genes coexpressing with it in COXPRESdb (http://coxpresdb.jp) database. Co-expressed gene networks in COXPRESdb are drawn based on the rank of correlation. Use of correlation rank partially normalizes the density of genes in correlation space.

4. Discussion

4.1 Proteomic and phosphoproteomic studies on U18666A treated MEFs

4.1.1 Induction of NPC by U18666A treatment

In order to investigate the alterations of proteome and phosphoproteome of NPC disease, an *in vitro* model of NPC was induced by treating cells with U18666A. Prior to investigating the disease, it was necessary to understand if U18666A can induce the phenotype accurately or it creates artificial effects. U18666A is an inhibitor of lysosomal transmembrane protein NPC1, which transports cholesterol from late endosomes or lysosomes to the endoplasmic reticulum (ER) and cytoplasm (Liscum and Faust 1989; Cenedella 2009). However, several other models have also been developed for studying NPC disease. The very first study was performed by Pentchev et al., where they used fibroblasts from 20 NPC patients and detected that intracellular cholesterol esterification was defective in those cells (Pentchev et al. 1985). Apart from patient cells, NPC1 knock out hepatocytes showed 5-fold higher cholesterol content in comparison to WT hepatocytes (Kulinski and Vance 2007). Knock out mouse models for NPC1 (Miyawaki et al. 1982) and NPC2 (Nielsen et al. 2011) were also used for studying NPC disease. However, in knockout cells due to the compensatory mechanism, it is not possible to investigate early events. On the other hand, U18666A mediated lysosomal cholesterol accumulation is a well-established model for studying the NPC disease (Lange et al. 2002). The mechanism by which U18666A inhibits cellular cholesterol trafficking is unclear, but several studies showed that U18666A interacts directly with NPC1 protein, its transport and lysosomal sequestration (Neufeld et al. 1999; Sugimoto et al. 2001). Cenedella et al. showed an increase in fluorescence emission from U18666A treated-cells compared to control cells while stained with filipin (Cenedella 2009). Although the total free cholesterol amount increases upon adding U18666A, it has been found that cholesterol accumulates mainly in the lysosomes and other cell compartments undergo cholesterol deficiency (Lange et al. 2002). Treating cells with U18666A also triggers a broad spectrum of alterations in proteins and their modifications. If it is considered critically, U18666A mediated NPC model is an ideal model because it doesn't create any genetic interference, thereby reduces potential artifacts. Moreover, by treating cells in different time-points with U18666A, early events of the disease and related protein alterations can also be observed, which is not possible in case of knock out cells or patient fibroblasts

4.1.2 Enrichment of lysosomes

In the current study, lysosomes were enriched from SILAC labeled MEF cells using superparamagnetic iron oxide nanoparticles (SPIONs) (Thelen et al. 2017), which is a very efficient method for lysosome enrichment. Walker et al. also have performed lysosome enrichment from different cell lines using the SPIONs technique, where they have compared enrichment efficiency from these cell lines (Walker and Lloyd-Evans 2015). They observed that Chinese hamster ovary (CHO) cell line and primary murine astrocytes have the greatest yield with the purification of 300-500 µg total lysosomal proteins from 2x T75 flasks. We have purified ~250 µg total lysosomal proteins from 1x 10 cm dish of MEF cells. One should consider that total yield and purity varies from cell line to cell line and also depends upon several factors e.g. culture media, optimal cell seeding number, coating of culture plates, pulse and chase period, type of dounce homogenizer and so on. Therefore, in order to have better yield, it needs to be optimized the purification condition of lysosome enrichment.

Especially while studying LSDs, the pulse-chase period is crucial to optimize since endocytic trafficking abnormalities are common in these diseases (Choudhury et al. 2005). In the case of NPC, in addition to fluid-phase endocytosis defects, delivery of materials between late endosomes and lysosomes is significantly retarded (Lloyd-Evans et al. 2008). Therefore, chase time needs to be extended from 24 h to 48 h. In this dissertation project, SPIONs technique is used for enriching lysosome fractions. However, in this dissertation project, storage was induced during the last 24 hours of the chase. Therefore, abnormal endocytosis is not really an issue in this case.

Several other methods are also available for subcellular fractionation and lysosome enrichment, each of them having their own advantages and disadvantages. Sucrose density centrifugation is the earliest method for organelle enrichment (Hirst and Cox 1976). However, in sucrose and other gradient media, organelles like lysosomes, peroxisomes, and mitochondria have similar and partially overlapping densities. Therefore, full separation of these organelles based on their density is very challenging. Nevertheless, using several techniques the resolution of lysosomes can be significantly improved, e.g. treatment of animals with Triton WR1339 results in density shift of liver lysosomes (Wattiaux et al. 1963) or by using dextran (Arai et al. 1991). Mitochondria become less dense and swell by adding 1 mM $CaCl_2$ to PNS and improve their separation from lysosomes in Percoll gradients (Arai et al., 1991).

Treatment of cultured skin fibroblasts with superparamagnetic magnetite/dextran nanoparticles results in the isolation of lysosomal membranes (Diettrich et al. 1998). For enriching lysosomes immunoprecipitation via tagged lysosomal membrane proteins has been used recently. Initially, overexpression of lysosomal membrane protein LAMP1 with RFP-FLAG tag was used to enrich lysosomes (Zoncu et al. 2011), thereafter TMEM192 with HA-tag version was utilized for the proteomic and metabolomic analysis of isolated lysosomes (Abu-Remaileh et al. 2017).

4.1.3 Lysosomal proteomics analysis from U18666A treated MEFs

Mass spectrometry has become a powerful tool in recent years for quantitative and qualitative characterization of thousands of macromolecules from complex samples (Aebersold and Mann 2003). A simple and quantitative approach, stable isotope labeling by amino acids in cell culture (SILAC) has also been used recently in combination with mass spectrometry which allowed it to be more accurate (Ong et al. 2002). In SILAC method, proteins are metabolically incorporated with stable isotopes of labeled amino acids ($^{13}C/^{15}N$ of arginine and/or lysine) in cultured cells. Therefore, a mass shift occurs between samples from different culture conditions resulting in proper quantification and comparison between them.

In this thesis project, in order to explore the alterations of proteins in proteome and phosphoproteome level, SILAC-based mass spectrometry analysis was performed on enriched lysosome fractions from U18666A treated MEFs. Proteomics dataset revealed a total of 3592 proteins with 142 known lysosomal proteins. Out of 3592 total proteins, only 31 proteins were significantly upregulated and 115 were downregulated, where only 7 lysosomal proteins were found in the upregulated proteins and 7 in the down-regulated proteins. Dehghani et al. have performed similar studies with U18666A treated MEFs but on whole cell lysates. From three biological replicates, they have found 5502 total proteins, 74 of them were differentially regulated (66 up- and 8 downregulated) (Dehghani 2019). These 66 upregulated proteins were compared with the 35 up-regulated proteins from the current study and only 7 proteins were found common in between these two studies. Interestingly, these 7 proteins are the proteins which are related to cholesterol metabolism (Table 3.1, green). Dataset from Dehghani was also evaluated for the presence of lysosomal proteins and found 2 lysosomal proteins as upregulated on U18666A treated MEF whole cell lysates. No lysosomal protein was downregulated in the proteomics dataset.

Besides the SILAC approach, other mass spectrometry-based proteomics studies have also been performed on NPC models. Quantitative mass spectrometry analysis of tandem mass tag (TMT) labeled peptides from NPC1-mutated human fibroblasts found 4308 distinct proteins where 61 proteins were upregulated on mutated cells compared to WT (Rauniyar et al. 2015). These 61 proteins were compared to our 35 upregulated proteins and only 3 proteins were found common: Delta (24)-sterol reductase (DHCR24), Gap junction alpha-1 protein (GJA1) and Cysteine-rich with EGF-like domain protein 1 (CRELD1). This inconsistency probably due to the use of two different models- U18666A mediated NPC model and NPC1^{I1061T}-fibroblasts. However, it has been shown that NPC1^{I1061T} mutated proteins can localize to lysosomal membrane and functions properly by transferring cholesterol from lysosomes to ER (Gelsthorpe et al. 2008). Therefore, U18666A induced NPC model is a more efficient model for studying altered proteome than NPC1-mutant cells.

In another study with CRISPR/Cas9 mediated NPC1 KO HeLa cells, lysosomes were isolated using SPIONs technique and proteomics and lipidomics analysis were accomplished by label-free quantification method (Tharkeshwar et al. 2017). From enriched lysosomes samples, in total 2279 proteins have been identified with upregulation of 31 proteins compared to WT cells. They argued that SPIONs is a promising and efficient resource tool for organelle isolation. Moreover, SPIONs is compatible with subsequent omics analysis, providing lower complexity and higher resolution of an organelle compared to a whole cell or tissue. Proteomics analysis of NPC$^{-/-}$ mouse brain with dimethyl labeling showed that CTSA, PPT1, and SCPEP1 were significantly elevated in NPC1 knockout mouse brains. These enzymes are lysosomal proteins, which is in line with the important role of NPC1 in lysosomes (van Rooden et al. 2018).

In a separate study, lysosomal proteins containing the specific mannose 6-phosphate modification were enriched from WT, *Npc1*$^{-/-}$ and *Npc2*$^{-/-}$ mutant mouse brains followed by mass spectrometry analysis (Sleat et al. 2012). Lysosomal proteins involved in lipid catabolism including prosaposin and the two subunits of β-hexosaminidase were increased in both forms of NPC, probably representing a compensatory cellular response to the accumulation of glycosphingolipids. Several other lysosomal proteins were significantly altered, including proteases and glycosidases. Changes in lysosomal protein levels resembled similar alterations in

activities and transcript levels. Understanding the rationale for such changes may provide insights into the pathophysiology of NPC.

4.1.3.1 Regulated proteins in proteomics dataset

To investigate the biological importance and role in the pathophysiological mechanism of NPC disease, differentially regulated proteins were subjected to GO analysis. GO analysis of 34 upregulated proteins showed that, these proteins are related to cholesterol metabolic and biosynthetic process, whereas 126 downregulated proteins showed their effect on the regulation of different cellular processes like organization, transportation, and cell growth. 7 from 34 upregulated proteins have a direct effect on the cholesterol biosynthetic pathway. This result is consistent with the findings of large scale phosphoproteomics study on U18666A treated MEFs, where upregulated proteins also showed an effect on cholesterol biosynthetic pathway (Dehghani 2019). It has been shown that the ER of U18666A treated cells experiences cholesterol deficiency, which leads to the inhibition of cholesterol synthesis enzyme degradation (Lange et al. 1999). Dhcr24 is the most differentially expressed protein in proteomics dataset, which catalyzes the ultimate step of cholesterol biosynthesis by conversion of desmosterol to cholesterol. Other than this, Dhcr24 also influences in modulating oxidative stress (Zerenturk et al. 2013). Next upregulated protein, Sqle catalyzes the first oxygenation reaction of cholesterol synthesis by converting non-sterol intermediate squalene to 2,3(S)-oxidosqualene (Padyana et al. 2019).

Interestingly, alkaline phosphatase, tissue-nonspecific isozyme (AlpI) is the 5th upregulated protein in the proteomics data set. Intralysosomal cholesterol accumulation upregulates alkaline phosphatase thereby resulting dephosphorylation of certain proteins and modulates their functional properties. We also have identified flotillin-1 (Flot1) in U18666A treated samples, which is consistent with the findings that upon cholesterol storage cells compensate for a load of excess cholesterol through exocytosis of flotillin positive vesicles (Strauss et al. 2010).

4.1.4 Lysosomal phosphoproteomics analysis of U18666A treated MEFs

Several studies on NPC phenotype models have investigated that, protein phosphorylation plays a crucial role in NPC disease (Garver et al. 1999; Sawamura et al. 2001; Xu et al. 2010; Tamari et al. 2013). All of these studies have been performed on NPC homozygous and heterozygous mouse models or on NPC1 knock out cell lines or patient fibroblasts. However, no phosphoproteomics study has

performed to date on enriched lysosomes from U18666A mediated NPC models for investigating phosphorylation mediated protein alterations in NPC disease.

From three biological replicates, in total 7024 phosphorylation sites were found where 69 % of the phosphosites are considered as class I phosphosites based on their localization probability. Statistical analysis found 245 phosphosites as regulated (123 up- and 122 down-regulated) on U18666A treated MEF lysosomes. In comparison to our study, large scale phosphoproteomics study on whole cell lysates from U18666A induced NPC models detected 12881 phosphosites. They have found 288 phosphosites as regulated (120 up- and 168 down-regulated) (Dehghani 2019). Analysis and further investigation on these potential phosphosites may provide insights into NPC pathophysiology.

However, phosphoproteomics studies on whole cell lysates demand sample fractionation due to its high complexity before running them in the mass spectrometer. Whereas in the case of enriched lysosomes no fractionation is necessary, which is time-saving, less laborious as well as cost-effective. Furthermore, due to the low abundance of certain lysosomal proteins, there is a high possibility to fail in identifying them in whole-cell lysates, which can be mitigated by organelle-specific omics study.

Phosphorylation studies on homozygous and heterozygous NPC1 mice liver found that annexin II is phosphorylated by PKC (protein kinase C) and caveolin-1 by casein kinase IIα. They also have found altered expression of several kinases in these mouse models assuming that they are involved in the pathophysiology of NPC (Garver et al. 1999). Using a similar model, another study detected phosphorylation of tau protein at Ser-396 and Ser-404 by mitogen-activated protein kinase (MAPK) in the brains of NPC homozygous mice (Sawamura et al. 2001). Moreover, siRNA driven knockdown of NPC1 and NPC2 results in inhibition of the mammalian target of rapamycin (mTOR), indicating that cholesterol trafficking is essential for mTOR activation (Xu et al. 2010). It has been shown that expression or activation of PKCε in NPC1 cells dramatically decreases the amount of stored cholesterol and restores cholesterol transport out of endocytic vesicles (Tamari et al. 2013). These results suggested that PKC may be targeted in future efforts to develop therapeutics for NPC1 disease. Interestingly, phosphorylation studies discovered that insulin signaling pathway is impaired in the brain of NPC mice (Ong et al. 2012a). They found lower expression of insulin receptor substrate 2 (IRS2) and almost non-

detectable expression of insulin receptor substrate 1 (IRS1) in this NPC mouse model, which is associated with the loss of expression for the regulatory p85 subunit of phosphatidylinositol 3-kinase (p85/PI3K).

4.1.4.1 Regulated phosphoproteins in phosphoproteomics dataset

In our phosphoproteomics data, we found 11 phosphorylation sites which belong to eight different lysosomal proteins. Gene ontology analysis of these regulated phosphosites showed that they are found to be involved in protein transportation, protein and/or organelle localization and intracellular organization. This has similarities with the findings of Dehghani et al., where GO analysis of regulated phosphoproteins showed an effect on protein localization and organization of cellular components and organelles. They concluded that U18666A treatment and following perturbation of cholesterol trafficking assists proteins and organellar mislocalization and hence phosphorylation helps proteins to adjust with the new conditions. Their findings also support the previous studies on cholesterol depleted and NPC cells where localization was altered in case of amyloid-beta A4 protein (Ehehalt et al. 2003), beta-secretase 1 (Simons et al. 1998) and presenilin-1 (Wahrle et al. 2002).

The eight up-regulated phosphoproteins are: two-pore calcium channel protein 1 (TPC1), Next to BRCA1 gene 1 (Nbr1), Endosomal/lysosomal potassium channel TMEM175 (TMEM175), Major facilitator superfamily domain-containing protein 1 (MFSD1), Ras-related protein Rab-9A (Rab9A), Vesicle-associated membrane protein 8 (Vamp8), Heparan-alpha-glucosaminide N-acetyltransferase (HGSNAT) and Transmembrane protein 106B (TMEM106B).

Two pore calcium channel protein 1 (TPC1) was the most regulated lysosomal phosphoprotein in the dataset. It is a nicotinic acid adenine dinucleotide phosphate (NAADP) receptor that may function as one of the major voltage-gated Ca^{2+} channels (VDCC) across the lysosomal and endosomal membrane (Guo et al. 2016). Nbr1 (neighbor of BRCA1 gene 1) is an autophagy receptor containing LC3- and ubiquitin (Ub)-binding domains and acts for selective autophagosomal degradation of ubiquitinated targets (Kirkin et al. 2009). It has been shown that phosphorylation of NBR1 at Thr586 residues by GSK3 modulates the formation of protein aggregates (Nicot et al. 2014). However, in the current study, we found Ser701 as phosphorylated. Tmem175 is an organelle-specific potassium channel specifically responsible for potassium conductance in endosomes and lysosomes. By forming a potassium-permeable leak-like channel it regulates the luminal pH stability and is

essential for autophagosome-lysosome fusion (Cang et al. 2015). TMEM175 has been shown to influence α-synuclein phosphorylation and autophagy (Jinn et al. 2019). Mitophagy formed by a protein complex of unc-51 like kinase 1 (Ulk1), Rab9, receptor-interacting serine/threonine-protein kinase 1 (Rip1), and dynamin-related protein 1 (Drp1) protects the ischemic heart. It has been shown that phosphorylation of Rab9 at Ser179 by Ulk1 is crucial for this mitophagy where knockin of Rab9 (S179A) showed to abolish mitophagy (Saito et al. 2019). We have detected phosphorylation of Rab9 in two sites- Thr185 and Thr187. Another regulated phosphoprotein is Vamp8, it is a lysosomal SNARE protein (soluble N-ethylmaleimide-sensitive factor-attachment protein receptors) which directly controls lysosome-autophagosome membrane fusion through its interaction with STX17 and SNAP-29 proteins (Itakura et al. 2012). It has been found that VAMP8 is phosphorylated by PKC at T54 residue which is in the cytoplasmic domain of this protein; therefore, affects its interaction with other proteins such as STX17 and SNAP-29 (Malmersjö et al. 2016). They also have found phosphorylation at T48 and S55, whereas we have found Vamp8 phosphorylated at T54 as well as S5.

Furthermore, 150 kinases were found to be phosphorylated in the dataset with different extent. Top three regulated kinases are mitogen-activated protein kinase kinase kinase 20 (Map3K20), dual serine/threonine and tyrosine-protein kinase (Dstyk) and casein kinase I isoform epsilon (KC1E). In NPC mouse model, MAPK was found regulated which also supports our finding (Sawamura et al. 2001). We also detected A-kinase anchor protein 5 (AKAP5) in the dataset. AKAP is a heterogeneous family of scaffolding proteins defined by its ability to directly bind protein kinase A (PKA). AKAPs tether PKA to specific subcellular compartments and they bind further interaction partners to create local signaling hubs (Dema et al. 2015).

4.2 Heparan-α-glucosaminide N-acetyltransferase (HGSNAT)

Mass spectrometry data detected phosphorylation at two serine residues (either S211 or S215) of HGSNAT, but no doubly phosphorylated peptides were detected. Phosphorylation at S211 (pS211) showed 2-fold up-regulation of the peptide (-LINSELGSPSR-) after U18666A treatment, however, protein level remained constant. Spectra for both heavy and light peptides were further manually evaluated for proper validation of the data, the results of which confirm the accuracy.

Furthermore, different online prediction tools and databases were searched for additional evidence about HGSNAT phosphorylation at these residues. Surprisingly, high throughput proteomics studies and curated databases on PhosphositePlus (https://www.phosphosite.org) (Hornbeck et al. 2015) also annotated these two phosphorylation sites on HGSNAT, however, no low throughput studies have found phosphorylation of HGSNAT. High-throughput data showed, Ser211 has been detected in various mass spectrometry-based phosphoproteomic studies (Olsen et al. 2010; Wu et al. 2012; Reinartz et al. 2014; Stuart et al. 2015), while Ser215 has been annotated by others (Wilson-Grady et al. 2013; Parker et al. 2015; Sacco et al. 2016; Robles et al. 2017). Since there is no low-throughput data on these residues and as we validated HGSNAT phosphorylation by phosphoproteomic analysis, it was chosen as a candidate phosphoprotein to further investigate if phosphorylation has any effect on the regulation of HGSNAT activity.

4.2.1 U18666A mediated regulation of HGSNAT activity

U18666A treatment on HEK293 cells showed a 2-fold increase in HGSNAT activity in enriched lysosome samples in comparison to the control samples. This increase in activity indicated that intracellular cholesterol accumulation might have an effect on the HGSNAT activity. It is known that the exit of recycling cholesterol from late endosomes/lysosomes is defective in Niemann-Pick Type C (NPC) disease (Vanier 2015). Simultaneously, the trafficking route of the recycling glypican-1 (Gpc-1), a heparan sulfate proteoglycan (HSPG), also involves the late endosomes and lysosomes. During recycling through intracellular compartments, the HS side chains of Gpc-1 are deaminatively degraded by nitric oxide (NO) derived from preformed S-nitroso groups in the core protein and forms anMan-containing HS oligosaccharides (Vilar et al. 1997). It has been shown that the trafficking of Gpc-1 through the endosomal compartments in inhibited by the cationic steroid U18666A. Concurrently, NO-dependent Gpc-1 degradation is also defective in U18666A treated fibroblasts (Mani et al. 2006). Mani et al. observed that in NPC1 fibroblasts the formation of anMan-containing HS oligosaccharides was reduced to ~10 % of the level in normal fibroblasts. The reason might be that in NPC1 cells there are few NO-sensitive sites in Gpc-1 HS or because there is defective NO release from SNO groups in Gpc-1 or because Gpc-1 is poorly S-nitrosylated. However, exogenous addition of ascorbate to these cells triggers deaminative cleavage of HS in SNO-containing Gpc-1 (Fransson et al., 2004). Furthermore, when they treated normal fibroblasts with

U18666A, they observed that anMan-containing HS oligosaccharide formation was reduced to ~20 % of the level in untreated cells. However, exposure to ascorbate increased the level of anMan-containing oligosaccharides to ~35 % of the level in untreated fibroblasts (Mani et al. 2006). They also noticed that, although treatment with 3 µg/mL of U18666A for 8 h was sufficient to cause cholesterol accumulation, inhibition of HS degradation required treatment with 10 µg/mL of U18666A for 16 h, presumably because growing fibroblasts have an exceptionally active NO-dependent HS degradation.

Besides U18666A, further experiments with another cationic amphiphile imipramine (80 µM for 18 h) did not inhibit the formation of anMan-containing degradation products, although it generated an accumulation of cholesterol. Estrone (3 µg/mL for 16 h), which is not cationic but has a C-17 keto group similar to that of U18666A, caused neither cholesterol accumulation nor inhibition of NO-dependent HS degradation. Depletion of total cellular cholesterol by treatment with 10 mM methyl-β-cyclodextrin for 2 h did not affect the deaminative HS degradation (Mani et al. 2006). All these findings indicated that intralysosomal cholesterol accumulation blocks NO-dependent Gpc-1 degradation. Therefore, it was assumed that increased HGSNAT activity on NPC in our study might have an effect on the degradation of endosomal HS side chains of Gpc-1.

4.2.2 Impairment of heparan sulfate degradation affects HGSNAT activity

During recycling, Gpc-1 is enzymatically processed by heparanase which degrades HS on both SNO-free and SNO-containing Gpc-1 (Fransson et al. 2004). HGSNAT is crucial for heparan sulfate degradation by acetylating its non-reducing terminal α-glucosamine residues (Fan et al. 2006; Hrebícek et al. 2006). As we found HGSNAT differentially phosphorylated, it was decided to impair HS degradation and observe its effect on HGSNAT activity. Therefore, suramin, an inhibitor of heparanase was used which blocks both internalization and degradation of HS (Nakajima et al. 1991). It was found that suramin causes a decrease in HGSNAT activity by 2.5-fold in comparison to the untreated control samples. This indicates, if suramin blocks heparanase, then endoglycosidase degradation of heparan sulfate from HSPGs to oligosaccharide side chains is ultimately impeded. Since heparan sulfate degradation occurs in a strictly sequential manner by 9 different lysosomal enzymes, the degradation process cannot be initiated in suramin treated cells. Therefore, non-reducing glucosamine residues will not be available for acetylation by HGSNAT. It

was shown by others that, upon U18666A mediated arrest of cellular cholesterol recycling, NO-dependent deaminative HS degradation is defective but, in such case, heparanase-catalyzed HS degradation is not affected (Mani et al. 2006). In our data, U18666A treated samples had higher HGSNAT activity, as heparanase is not blocked therefore HS degradation can continue and HGSNAT can acetylate non-reducing glucosamine residue of HS, whereas as soon as heparanase is blocking by suramin the HGSNAT activity is reduced drastically.

Moreover, it has been shown that blocking of heparanase by suramin doesn't interfere with NO-dependent degradation (Cheng et al. 2002). We can postulate that, since we are inhibiting heparanase degradation of HS by suramin, more Gpc-1 HS will be available for deaminative cleavage by NO. We could also prove this by radioactive labeling, where cells feeding with $^{35}SO_4$ results in ^{35}S labeling of HS, which might be another approach to detect the amount of fully degraded, partially degraded or undegraded HS in the cell. If HGSNAT activity is an exemplar for HS degradation, we would expect that suramin mediated lower HGSNAT activity would generate an increased amount of undegraded or partially degraded HS.

4.2.3 Dephosphorylation increases HGSNAT activity

Phosphoproteomic data showed 2-fold up-regulation of HGSNAT phosphopeptides. Consequently, dephosphorylation of enriched lysosomes by shrimp alkaline phosphatase (SAP) treatment results in ~2.5-fold up-regulation of HGSNAT activity in all samples either untreated or treated with U18666A and suramin. This result is contradictory to what we have found in mass spectrometry data. The reason why we see an increase in activity in all conditions upon SAP mediated dephosphorylation is not clear. Moreover, by adding SAP we are not only removing all the phosphate groups from all HGSNAT phosphosites but also from other enzymes and proteins. Therefore, this could have unexpected effects on HGSNAT activity. However, it is still not clear if phosphorylation has any direct influence on HGSNAT activity.

4.2.4 Effect of phosphorylation on HGSNAT activity

It is known that phosphorylation as a post-translational modification regulates the activity of various enzymes (Beg et al. 1978; Hurley et al. 1990; Liu et al. 2019). To observe whether phosphorylation affects HGSNAT activity, phosphomimetic studies were performed on S211 and S215 with all possible combinations of E/D/A (either single or double mutations). Overexpression of the HGSNAT in HGSNAT KO cells resulted in saturation of the rate-limiting enzyme β-hexosaminidase. Therefore,

HGSNAT activity determination was not linear and not accurate in such cases. We, therefore, optimized the assay. It is known that overexpression of proteins can disrupt downstream regulation, balanced gene dosage, disturbs protein folding and complex assembly, thereby generating artifacts in the experiments (Gibson et al. 2013). Therefore, expressing modified proteins at almost native level is the best approach to study their effects.

To that end, Tet-inducible system was used for expression control, where the application of tetracycline or its derivatives like doxycycline can control protein expression to the basal level (Gossen and Bujard 1992; Das et al. 2016). Knocking out HEK293 Tet-On cells for HGSNAT gene by CRISPR/Cas9 method did not work in our hand. Probably use of two antibiotics, doxycycline for maintaining Tet-On cells and puromycin for CRISPR selection, cross-reacted with each other and interrupted the function of Cas9 nuclease. It has been shown that a tetracycline-inducible CRISPR/Cas9 system repressed the malignant behavior of bladder cancer cells by controlling the expression of Cas9 and simultaneously targeting two oncogenic lncRNAs (Peng et al. 2018). However, they used a tet-inducible system to express Cas9 which differs from our experimental set-up. HGSNAT KO cells were converted successfully to Tet-inducible cell line by using the piggyBac transposon system (Li et al. 2013), which is a very efficient method to integrate gene of interest into the cellular genome. Expression of all HGSNAT mutants on HGSNAT Tet-On cells showed a significant decrease in activity with S215E, EA and DD mutants, which is not conclusive. Since E and D mimic phosphorylation and A mimics phosphoresistant conditions, decreased activity on S215E and DD showed phosphorylation doesn't have any direct influence on HGSNAT activity. More importantly, it is not necessarily always the case that E or D mimics phosphorylation.

However, different studies have shown that missense mutation affects HGSNAT activity. Feldhammer et al. studied 17 different missense mutations of HGSNAT located on 8 of its 11 transmembrane segments as well as luminal and cytosolic domains. All these mutants resulted in misfolded HGSNAT protein with negligible HGSNAT residual activity (Feldhammer et al. 2009b).

Surprisingly, in contrast to endogenous samples, treatment of expressed double mutants with SAP showed reduced activity in all samples. If S211 or S215 is phosphorylated, then SAP treatment should not change in activity particularly in the case of serine substituted HGSNAT. This led towards the hypothesis that, there is

-100-

another activating factor for HGSNAT, which is rate-limiting. SAP treatment dephosphorylates the activating factor but not the HGSNAT. Therefore, overexpression of HGSNAT suppresses the amount of the activating factor resulting in no HGSNAT activity.

Moreover, treating all transfected cells with U18666A showed increase in activity with a similar extent in the case of all samples which is also surprising to us. Because if phosphorylation is really involved in activity regulation and U18666A is involved in this process, then phosphomimetic or phosphoresistant HGSNAT should not show any change in activity compared to WT HGSNAT. Therefore, it indicates probably phosphorylation is not the stimuli for HGSNAT activity regulation.

Furthermore, recently we have started developing two phosphospecific antibodies for HGSNAT in rabbit using the epitope HGSNAT phosphopeptides 'LINpSELGSPSR' and 'LINSELGpSPSR'. This is underway and we believe this antibody will further confirm the presence of phosphorylation on any of the residues.

4.2.5 Effect of phosphorylation on HGSNAT subcellular localization

It is well known that phosphorylation of proteins alters their subcellular localization (Blenis and Resh 1993; Puertollano et al. 2018). Since no effect of HGSNAT phosphorylation on its activity regulation was found, it was further postulated that it might affect HGSNAT localization. Therefore, NIH3T3 Tet-On cells were transfected with either WT or phosphomutants and performed immunofluorescent microscopy. It showed all constructs co-localize with the lysosomal marker indicating phosphorylation of HGSNAT doesn't alter its lysosomal localization. Similar findings were also observed by western blot and HGSNAT activity assay from different fractions of lysosome enrichment where lysosome fractions contain HGSNAT expression and retain the highest activity respectively. Whereas, expression of 17 missense mutations of HGSNAT identified in MPSIIC patients in cultured human fibroblasts and COS-7 cells showed that all mutations caused misfolding of the enzyme and thereby abnormal glycosylation. Immunofluorescence microscopy revealed that they were not targeted to the lysosome but retained in the endoplasmic reticulum (ER) (Feldhammer et al. 2009b). Fedele and Hopwood showed, all of 20 missense mutant variants of HGSNAT-Myc co-localized with the lysosomal marker LysoTracker Red DND-99 (Fedele and Hopwood 2010). In another study, in order to investigate lysosomal colocalization of HGSNAT enzyme, singly tagged HGSNAT (HGSNAT-His8FLAG) was compared with doubly tagged HGSNAT (myc-tag on N-

terminus before signal peptide; myc-HGSNAT-His8FLAG), where they found c-terminally tagged protein co-localizes with lysosomal marker β-Glucocerebrosidase, while doubly tagged protein localized in ER (Fan et al. 2011). Therefore, they suggested that retention of signal peptide prevents intracellular transport of the precursor to the lysosome and proteolytic processing into the mature α- and β-chains, but it doesn't prevent the activity of the enzyme. Based on obtained results from this dissertation, it was found that phosphomutants both -mimic and -resistant does not have any effect on HGSNAT enzymatic activity and lysosomal localization.

4.2.6 Identification of HGSNAT interaction partners

HGSNAT, the multi-pass lysosomal transmembrane protein, plays a crucial role in intralysosomal degradation of heparan sulfate but a little is known about which proteins interact with HGSNAT in order to facilitate its function and/ intracellular transport (Fedele et al. 2018). Moreover, as no effect of phosphorylation was found on HGSNAT activity, we postulated that other proteins interacting with HGSNAT might influence or regulate its phosphorylation. Therefore, in this study, probable candidates interacting with HGSNAT were identified by mass spectrometry analysis by two different approaches.

In the first approach, to immunoprecipitate, HGSNAT interactors with anti-HGSNAT antibody, post-nuclear supernatant (PNS) and lysosome samples were used. This was performed because at the same time the scenario of the intracellular interaction of HGSNAT can be captured in two different states and can also be compared if there any changes. Since HGSNAT is a lysosomal transmembrane enzyme, in lysosome enriched samples true interactors can be identified which can then be further confirmed by their presence in PNS samples. The control of this experiment was SILAC light labeled HGSNAT KO HEK293 cells and as these samples were combined with SILAC heavy labeled WT HEK293 cells, H/L ratios should provide true interactors with the removal of all backgrounds. Moreover, since the precipitation was done in untransfected cells, no artifacts from overexpression can influence the result (Gibson et al. 2013).

On the other hand, in the second approach, the SILAC triplex method was used, using this method at the same time interactors of HGSNAT on U18666A treated vs. control can be identified. SCARB2 was used as a control in this experiment, which is a lysosomal transmembrane protein required for normal biogenesis and maintenance of lysosomes and endosomes (Gonzalez et al. 2014). Proteins interacting with

SCARB2 samples can be filtered out from HGSNAT enriched data to identify the true interactors for HGSNAT. Furthermore, PFA crosslinking (Klockenbusch and Kast 2010) was performed in this approach, where proteins in close vicinity get covalently linked, therefore intermediate and weak interactors can be identified in this way which is normally lost during the washing steps of the affinity purification.

By using both approaches, 91 significantly enriched candidates were detected in PNS samples and 76 from lysosome samples whereas 29 and 21 candidates were identified from HGSNAT treated vs untreated and HGSNAT vs SCARB2 samples respectively from three independent biological replicates. Proteins related to lipid metabolism were found enriched in U18666A treated HGSNAT samples. Probably, the cellular machinery upon excess cholesterol deposition regulates HGSNAT to interact with proteins related to lipid metabolism and somehow influence on the catabolism of deposited unesterified cholesterol in the lysosome. Prediction analysis for HGSNAT interaction partner using STRING and COXPRESdb online tools showed no similarities between their predicted proteins and enriched proteins from this study. However, one should consider that these predictions are based on data generated from high throughput analysis but no direct study for HGSNAT interaction partners.

13 proteins were found common in between PNS and lysosome samples which are Angiomotin (AMOT), Insulin receptor substrate (IRS4), WD repeat-containing protein 6 (WDR6), Heparan-alpha-glucosaminide N-acetyltransferase (HGSNAT), Inositol 1,4,5-trisphosphate receptor type 3 (ITPR3), Serpin H1 (SERPINH1), Syntaxin-12 (STX12), Catenin beta-1 (CTNNB1), Plakophilin-2 (PKP2), Septin 6, isoform CRA (SEPT6), Desmoglein-2 (DSG2), Dolichol-phosphate mannosyltransferase subunit 1 (DPM1) and Emerin (EMD).

Angiomotin (AMOT) expresses predominantly in endothelial cells, thereby regulates endothelial cell migration and tube formation. It also plays a role in the assembly of endothelial cell-cell junctions (Troyanovsky et al. 2001). Insulin receptor substrate (IRS4) is a cytoplasmic protein which plays a key role in insulin signaling. It has been shown that IRS4 interacts with WD repeat-containing protein 6 (WDR6) (Chiba et al. 2009) and NISCH (Sano et al. 2002). Both IRS4 and WDR6 proteins get highly enriched in lysosome samples compared to PNS samples, when immunoprecipitated with the anti-HGSNAT antibody. It has been also shown that there is a cross-talk between insulin signaling and cholesterol metabolism (Suzuki et al. 2010). In NPC[nih]

mice, lower GSK3β phosphorylation was observed, which is associated with the impairment of insulin signaling including lower Akt phosphorylation as well as decreased expression of p85/PI3K, IRS1, and IRS2 (Ong et al. 2012b). Since HGSNAT was regulated by phosphorylation on NPC induced cells and we found IRS4 is immunoprecipitated with HGSNAT, there might be a correlation between IRS4 and HGSNAT.

Moreover, the only known interaction partner of HGSNAT is ALG-2-interacting protein X (ALIX), a protein that associates with the endosomal sorting complexes required for transport (ESCRT) components and which is implicated in the targeting of proteins to intraluminal vesicles of multivesicular bodies (Fedele et al. 2018). They found HGSNAT in the exosomes from human urine and cultured medium conditioned with HEK293T cells, co-localizing with ALIX. Therefore, proteins related to ESCRT complex or multivesicular bodies might have interaction with HGSNAT.

Furthermore, we found HGSNAT regulated in U18666A treated samples and it is known that defective cholesterol transportation pathway blocks the trafficking of GPC-1 (Mani et al. 2006), which constitutes heparan sulfate side chain. Therefore, it can be assumed that proteins related to cholesterol and/or lipid metabolism might have interaction with HGSNAT. 12 proteins were enriched in our dataset which is involved in lipid metabolism. Further investigation on them probably provides proper evidence of an interaction.

5. Abstract

The hypothesis underlying this thesis is that the function of lysosomal membrane proteins can be regulated by interconversion, specifically phosphorylation. This was investigated in a drug induced cellular model of the lysosomal storage disorder Niemann-Pick disease Type C (NPC). In this disease mutations in genes of two distinct proteins, NPC1 or NPC2, cause impaired cholesterol efflux from late endosomes/lysosomes to the endoplasmic reticulum. This results in the accumulation of unesterified cholesterol inside the lysosome and it was hypothesized that this causes changes in the lysosomal phosphoproteome. To that context, lysosomal cholesterol storage was induced in mouse embryonic fibroblasts by U18666A, an inhibitor of NPC1.

The phosphoproteome of lysosomes enriched from untreated and U18666A treated cells was compared to identify quantitative changes of protein phosphorylation. Mass spectrometry-based phosphoproteomic analysis of lysosomes identified and quantified a total of 7627 phosphosites with an upregulation of 123 phosphosites on U18666A treated samples. Of these, 11 phosphopeptides could be assigned to lysosomal proteins proving that in fact lysosomal proteins can be differentially phosphorylated. In an effort to demonstrate functional consequences of phosphorylation one of the 11 proteins - heparan α-glucosaminide N-acetyltransferase (HGSNAT) essential for heparan sulfate (HS) degradation - was chosen for more detailed examination. The 2.36-fold increased phosphorylation of serine211 within the cytosolic domain of this enzyme correlated well with a 2.8-fold increase of activity upon U18666A treatment suggesting that phosphorylation regulated HGSNAT activity. Moreover, blocking heparan sulfate degradation by suramin treatment decreased HGSNAT activity 4-fold, indicating that HS catabolism influences HGSNAT activity. However, phosphomimetic and phosphoresistant amino acid substitutions of serine211 did not show any significant change in HGSNAT activity ruling out that phosphorylation regulates activity. Moreover, immunohistochemistry showed that wild type and all HGSNAT mutants, localize to lysosomes, indicating that phosphorylation has no effect on subcellular localization of HGSNAT. When enriched lysosomes were treated with shrimp alkaline phosphatase, surprisingly dephosphorylation caused a ~2.5-fold increase of HGSNAT activity. This was only observed with endogenous HGSNAT in non-overexpressing cells. When lysosomes enriched from cells overexpressing wild type or serine211 phosphomimetic/resistant HGSNAT, dephosphorylation had no effect. Since this suggested that HGSNAT interacting proteins influencing its activity may exist, SILAC (Stable isotope labeling of amino acids in cell culture)-based co-immunoprecipitation experiments to identify HGSNAT interacting proteins were performed.

6. Appendix

6.1 Plasmid maps
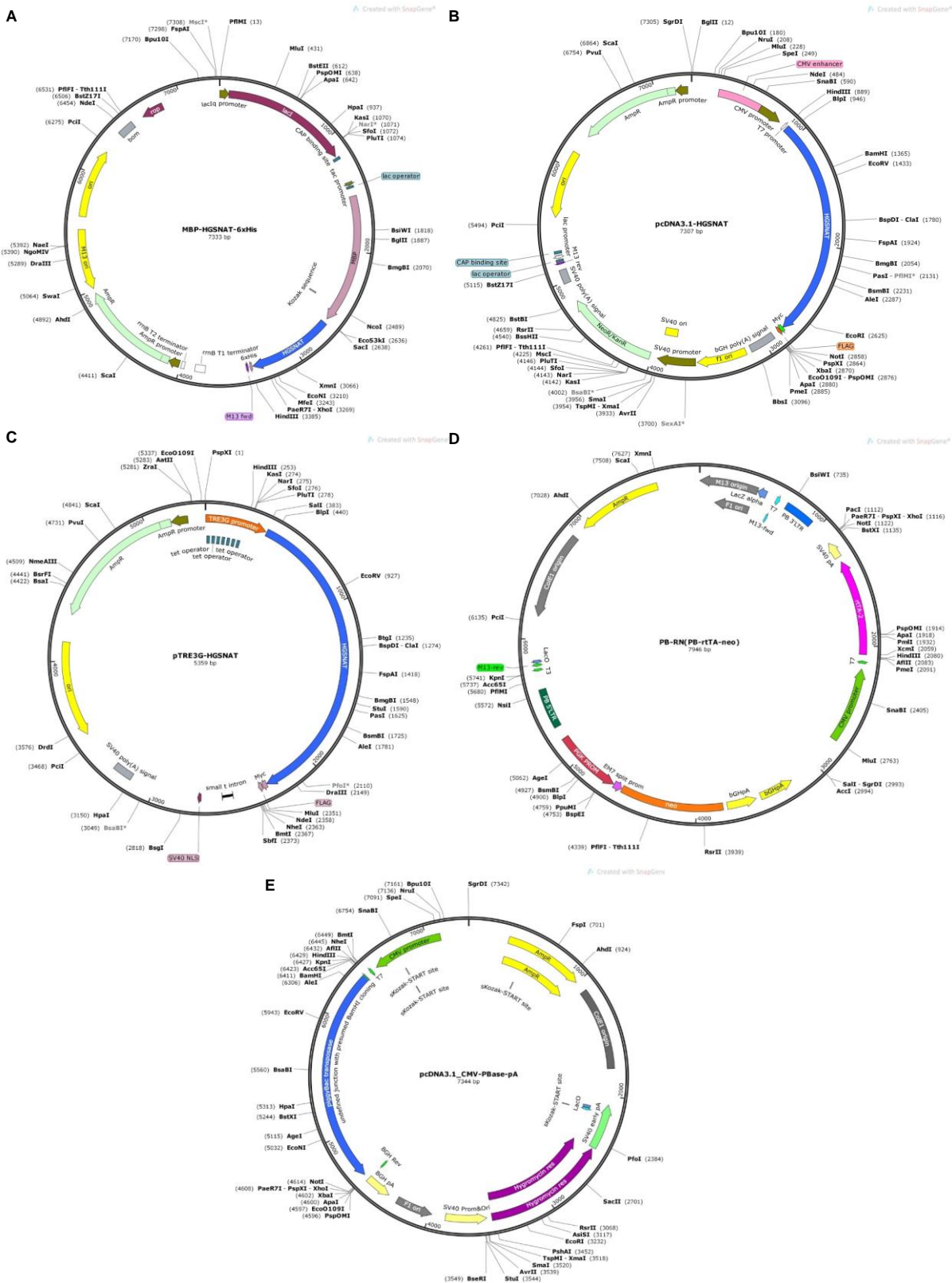
A. Plasmid map of MBP-HGSNAT-6xHis; B. Plasmid map of pcDNA3.1-HGSNAT-Myc-FLAG; C. Plasmid map of pTRE3G-HGSNAT-Myc-FLAG; D. Plasmid map of PB-rtTA-Neo; E. Plasmid map of pcDNA3.1-CMV-PBase-pA

6.2 Amino acid sequences

Mouse HGSNAT

```
          10         20         30         40         50         60         70
  MGAGPALAAL LLAGSVLSAT LLAPGRRAEP DLDEKRNVEL KMDQALLLIH NELLGTSLTV YWKSDDCYQC
          80         90        100        110        120        130        140
  TFQPLANVSH GGKPAKPSVA PVSVSTQHGS ILQVNSTSEE RAACRLEYKF GEFGNYSLLV QHASSGANKI
         150        160        170        180        190        200        210
  ACDIIVNENP VDSNLPVSIA FLVGLALIVA VSLLRLLLSL DDVNNWISKT IASRETDRLI NSELGSPSRA
         220        230        240        250        260        270        280
  DPLSADYQPE TRRSSANRLR CVDTFRGLAL VLMVFVNYGG GKYWYFKHSS WNGLTVADLV FPWFVFIMGT
         290        300        310        320        330        340        350
  SIFLSMTSIL QRGCSKLKLL GKIVWRSFLL ICIGVIIVNP NYCLGPLSWD KVRIPGVLQR LGVTYFVVAV
         360        370        380        390        400        410        420
  LEFFFWKPVP DSCTLESSCF SLRDITSSWP QWLTILTLES IWLALTFFLP VPGCPTGYLG PGGIGDLGKY
         430        440        450        460        470        480        490
  PHCTGGAAGY IDRLLLGDNH LYQHPSSTVL YHTEVAYDPE GVLGTINSIV MAFLGVQAGK ILVYYKDQTK
         500        510        520        530        540        550        560
  AILTRFAAWC CILGLISIVL TKVSANEGFI PINKNLWSIS YVTTLSCFAF FILLILYPVV DVKGLWTGTP
         570        580        590        600        610        620  624
  FFYPGMNSIL VYVGHEVLEN YFPFQWKLAD EQSHKEHLIQ NIVATALWVL IAYVLYKKKL FWKI
```

Human HGSNAT

```
          10         20         30         40         50         60         70
  MSGAGRALAA LLLAASVLSA ALLAPGGSSG RDAQAAPPRD LDKKRHAELK MDQALLLIHN ELLWTNLTVY
          80         90        100        110        120        130        140
  WKSECCYHCL FQVLVNVPQS PKAGKPSAAA ASVSTQHGSI LQLNDTLEEK EVCRLEYRFG EFGNYSLLVK
         150        160        170        180        190        200        210
  NIHNGVSEIA CDLAVNEDPV DSNLPVSIAF LIGLAVIIVI SFLRLLLSLD DFNNWISKAI SSRETDRLIN
         220        230        240        250        260        270        280
  SELGSPSRTD PLDGDVQPAT WRLSALPPRL RSVDTFRGIA LILMVFVNYG GGKYWYFKHA SWNGLTVADL
         290        300        310        320        330        340        350
  VFPWFVFIMG SSIFLSMTSI LQRGCSKFRL LGKIAWRSFL LICIGIIIVN PNYCLGPLSW DKVRIPGVLQ
         360        370        380        390        400        410        420
  RLGVTYFVVA VLELLFAKPV PEHCASERSC LSLRDITSSW PQWLLILVLE GLWLGLTFLL PVPGCPTGYL
         430        440        450        460        470        480        490
  GPGGIGDFGK YPNCTGGAAG YIDRLLLGDD HLYQHPSSAV LYHTEVAYDP EGILGTINSI VMAFLGVQAG
         500        510        520        530        540        550        560
  KILLYYKART KDILIRFTAW CCILGLISVA LTKVSENEGF IPVNKNLWSL SYVTTLSSFA FFILLVLYPV
         570        580        590        600        610        620        630
  VDVKGLWTGT PFFYPGMNSI LVYVGHEVFE NYFPFQWKLK DNQSHKEHLT QNIVATALWV LIAYILYRKK
         635
  IFWKI
```

Maltose Binding Protein (MBP)

```
          10         20         30         40         50         60         70
  MKIEEGKLVI WINGDKGYNG LAEVGKKFEK DTGIKVTVEH PDKLEEKFPQ VAATGDGPDI IFWAHDRFGG
          80         90        100        110        120        130        140
  YAQSGLLAEI TPDKAFQDKL YPFTWDAVRY NGKLIAYPIA VEALSLIYNK DLLPNPPKTW EEIPALDKEL
         150        160        170        180        190        200        210
  KAKGKSALMF NLQEPYFTWP LIAADGGYAF KYENGKYDIK DVGVDNAGAK AGLTFLVDLI KNKHMNADTD
         220        230        240        250        260        270        280
  YSIAEAAFNK GETAMTINGP WAWSNIDTSK VNYGVTVLPT FKGQPSKPFV GVLSAGINAA SPNKELAKEF
         290        300        310        320        330        340        350
  LENYLLTDEG LEAVNKDKPL GAVALKSYEE ELAKDPRIAA TMENAQKGEI MPNIPQMSAF WYAVRTAVIN
         360        370        379
  AASGRQTVDE ALKDAQTNSS SGGGGGGGS
```

6.3 List of known lysosomal proteins

In-house curated list of lysosomal and lysosome-associated proteins (189 proteins):

Uniprot ID	Gene	Uniprot ID	Gene	Uniprot ID	Gene	Uniprot ID	Gene
P41234	ABCA2	P16675	CTSA	Q9JHS3	LAMTOR2	Q921J2	RHEB
Q9DC29	ABCB6	P10605	CTSB	O88653	LAMTOR3	Q9CQ01	RNASET2
Q9JJ59	ABCB9	P97821	CTSC	Q8CF66	LAMTOR4	Q8K4Q0	RPTOR
O89016	ABCD4	P18242	CTSD	Q9D1L9	LAMTOR5	Q80X95	RRAGA
P24638	ACP2	Q9R013	CTSF	Q60961	LAPTM4A	Q6NTA4	RRAGB
Q05117	ACP5	P49935	CTSH	Q91XQ6	LAPTM4B	Q99K70	RRAGC
Q64191	AGA	P55097	CTSK	Q61168	LAPTM5	Q7TT45	RRAGD
Q9D1F4	AKT1S1	P06797	CTSL1	P16110	LGALS3	O35114	SCARB2
O35298	AOAH	Q8BM88	CTSO	Q9JL15	LGALS8	Q920A5	SCPEP1
P12246	APCS	O70370	CTSS	O89017	LGMN	Q8K0E7	SERINC2
Q8VEH3	ARL8A	Q9WUU7	CTSZ	Q9Z0M5	LIPA	Q9EQ08	SGSH
Q9CQW2	ARL8B	Q570Y9	DEPTOR	Q9JLJ0	LITAF	P70665	SIAE
P50428	ARSA	Q8BFQ6	DIRC2	Q8K0B2	LMBRD1	Q8CIF6	SIDT2
P50429	ARSB	P56542	DNASE2	Q9D920	LOH12CR1	Q8BN82	SLC17A5
Q3TYD4	ARSG	Q9ET22	DPP7	O09159	MAN2B1	Q80ZD3	SLC26A11
Q9D2L1	ARSK	Q9DBT4	ENTPD4	O54782	MAN2B2	Q8VE96	SLC35F6
Q9WV54	ASAH1	Q99M71	EPDR1	Q8K2I4	MANBA	Q8K4D3	SLC36A1
Q9R1Q9	ATP6AP1	Q99LJ1	FUCA1	Q9DBD2	MARCH8	Q9WU81	SLC37A2
Q9Z1G4	ATP6V0A1	Q8VDC1	FYCO1	Q99J21	MCOLN1	Q3TIT8	SLC37A3
P63082	ATP6V0C	P70699	GAA	Q99L90	MCRS1	Q8BWH0	SLC38A7
P51863	ATP6V0D1	P54818	GALC	Q9D6Y4	MEF2BNB	Q8BGD6	SLC38A9
Q9CQD8	ATP6V0E1	Q571E4	GALNS	Q9DC37	MFSD1	Q9DC26	SLC46A3
P50516	ATP6V1A	P17439	GBA	Q8BH31	MFSD8	Q04519	SMPD1
P62814	ATP6V1B2	Q9Z0L8	GGH	Q9DCJ1	MLST8	Q9Z266	SNAPIN
Q9Z1G3	ATP6V1C1	P51569	GLA	P11247	MPO	Q9JJF9	SPPL2A
P57746	ATP6V1D	P23780	GLB1	Q9JLN9	MTOR	P58749	TM6SF1
P50518	ATP6V1E1	Q9JHJ3	GLMP	Q9D7V9	NAAA	Q8BNQ3	TM7SF1
Q9D1K2	ATP6V1F	Q60648	GM2A	Q9QWR8	NAGA	Q3TB48	TMEM104
Q9CR51	ATP6V1G1	Q8BFR4	GNS	O88325	NAGLU	Q80X71	TMEM106B
Q8BVE3	ATP6V1H	P12265	GUSB	P57716	NCSTN	Q8BGP5	TMEM127
Q6PGD0	ATRAID	P11032	GZMA	O35657	NEU1	Q9CXY1	TMEM175
Q07813	BAX	P04187	GZMB	Q8BZL1	NEU4	Q9CXT7	TMEM192
O54918	BCL2L11	P29416	HEXA	O35604	NPC1	Q9CZX7	TMEM55A
O55102	BLOC1S1	P20060	HEXB	Q9Z0J0	NPC2	Q3TWL2	TMEM55B
Q9CWG9	BLOC1S2	Q3UDW8	HGSNAT	Q8BGT0	OSTM1	Q8BQU7	TMEM74
Q9CRC6	BORCS7	Q6YGZ1	HPSE	Q3V3Q7	PACS2	Q9ESN3	TMEM8A
Q9D6W8	C17ORF59	Q91ZJ9	HYAL1	Q9CQF9	PCYOX1	Q9CR23	TMEM9
Q8C1Y8	CCZ1	O35632	HYAL2	Q9WVJ3	PGCP	Q9EQJ0	TPCN1
P41731	CD63	Q08890	IDS	Q8VEB4	PLA2G15	Q8BWC0	TPCN2
P31996	CD68	P48441	IDUA	Q3TCN2	PLBD2	O89023	TPP1
P31996	CD68	Q9ESY9	IFI30	O35405	PLD3	Q61037	TSC2
O70496	CLCN7	O09046	IL4I1	O88531	PPT1	Q99J59	TSPAN1
Q61124	CLN3	Q80XH1	KXD1	O35448	PPT2	Q3TH73	TTYH2
Q3UMW8	CLN5	P11438	LAMP1	Q8C4N4	PQLC2	Q6P5F7	TTYH3
Q9D3S9	CPVL	P17047	LAMP2	Q7TMR0	PRCP	O70404	VAMP8
O88668	CREG1	Q7TST5	LAMP3	Q61207	PSAP	Q78T54	VMA21
Q8R242	CTBS	Q9D387	LAMP5	Q8R143	PTTG1IP		
P57757	CTNS	Q9CQ22	LAMTOR1	P06281	REN1		

Panther annotated list of lysosomal and lysosome-associated proteins (411 proteins):

Uniprot ID	Gene	Uniprot ID	Gene	Uniprot ID	Gene	Uniprot ID	Gene
P41234	ABCA2	P24270	CAT	P70658	CXCR4	Q31099	H2-DMB2
Q8K448	ABCA5	Q8VE99	CCDC115	Q60720	CYB561	P04230	H2-EB1
Q9JJ59	ABCB9	O08786	CCKAR	Q6P1H1	CYB561A3	O35424	H2-OB
P48410	ABCD1	Q8C1Y8	CCZ1	Q925G2	CYBRD1	O35668	HAP1
P09470	ACE	Q9R0L9	CD164	P61460	DEPDC5	P08103	HCK
P24638	ACP2	P11609	CD1D1	Q8BFQ6	DIRC2	P29416	HEXA
Q05117	ACP5	Q6SJQ5	CD300LD3	P56542	DNASE2	P20060	HEXB
Q8CE08	ACPP	Q64314	CD34	Q9QY48	DNASE2B	Q99LI8	HGS
P03958	ADA	P41731	CD63	Q7TNF0	DOC2A	Q3UDW8	HGSNAT
Q05910	ADAM8	P31996	CD68	Q9ET22	DPP7	Q99KG7	HPS4
Q64191	AGA	P04441	CD74	Q9DC58	DRAM1	Q8BLY7	HPS6
P45376	AKR1B1	Q9DB75	CDIP1	Q9CR48	DRAM2	Q6YGZ1	HPSE
Q91Y97	ALDOB	Q922Q9	CHID1	Q3UIR3	DTX3L	Q9ESB3	HRG
Q8C8R3	ANK2	Q9D7Q1	CHIT1	Q8R1Q8	DYNC1LI1	P11499	HSP90AB1
G5E8K5	ANK3	Q8BJF9	CHMP2B	Q6PDL0	DYNC1LI2	P63017	HSPA8
Q3UMR0	ANKRD27	O70496	CLCN7	Q99M71	EPDR1	Q91ZJ9	HYAL1
P10107	ANXA1	Q80U30	CLEC16A	Q91WM6	EVA1A	O35632	HYAL2
P97384	ANXA11	Q61124	CLN3	P41047	FASLG	Q8VEI3	HYAL3
P07356	ANXA2	Q3UMW8	CLN5	Q61851	FGFR3	Q08890	IDS
P14824	ANXA6	Q68FD5	CLTC	Q80TY0	FNBP1	P48441	IDUA
Q8BJ63	AP5M1	Q9WVJ3	CPQ	P09528	FTH1	Q9ESY9	IFI30
Q9D742	AP5S1	Q60571	CRHBP	P29391	FTL1	Q9CQW9	IFITM3
P08226	APOE	P21460	CST3	P49945	FTL2	P33896	IFNAR1
P56402	AQP2	Q8R242	CTBS	Q99LJ1	FUCA1	Q07113	IGF2R
Q8VEH3	ARL8A	P57757	CTNS	Q99KR8	FUCA2	P10749	IL1B
Q9CQW2	ARL8B	Q9CR35	CTRB1	Q8VDC1	FYCO1	O09046	IL4I1
Q7TPQ9	ARRDC3	Q91ZD5	CTS3	P70699	GAA	Q60766	IRGM1
P50428	ARSA	Q9ET52	CTS6	Q9DCD6	GABARAP	Q91WI7	ITFG2
P50429	ARSB	Q91ZF2	CTS7	P54818	GALC	Q91VK4	ITM2C
Q3TYD4	ARSG	Q9JI81	CTS8	Q571E4	GALNS	P97414	KCNQ1
Q9WV54	ASAH1	P16675	CTSA	P17439	GBA	A2AFS3	KIAA1324
P16460	ASS1	P10605	CTSB	P03995	GFAP	P28740	KIF2A
Q9CTG6	ATP13A2	P97821	CTSC	Q9Z0L8	GGH	Q99JN2	KLHL22
P63082	ATP6V0C	P18242	CTSD	Q8BWF2	GIMAP5	Q8VCX6	KPTN
Q9CR51	ATP6V1G1	Q9R013	CTSF	P23242	GJA1	Q80XH1	KXD1
Q6PGD0	ATRAID	P49935	CTSH	P51569	GLA	P11438	LAMP1
Q99ML1	BBC3	Q9R014	CTSJ	P23780	GLB1	P17047	LAMP2
E9Q623	BC051665	P55097	CTSK	Q9JHJ3	GLMP	Q7TST5	LAMP3
Q9Z0H7	BCL10	P06797	CTSL	Q60648	GM2A	Q9D387	LAMP5
O55102	BLOC1S1	Q3ULP7	CTSLL3	Q8BFR4	GNS	Q9CQ22	LAMTOR1
Q9CWG9	BLOC1S2	Q9JL96	CTSM	P05201	GOT1	Q9JHS3	LAMTOR2
Q9D920	BORCS5	Q8BM88	CTSO	Q8CFZ4	GPC3	Q8CF66	LAMTOR4
Q9D6W8	BORCS6	Q91ZF4	CTSQ	P70259	GPR143	Q9D1L9	LAMTOR5
Q9CRC6	BORCS7	Q9JIA9	CTSR	P28798	GRN	Q60961	LAPTM4A
Q9D6Y4	BORCS8	O70370	CTSS	P12265	GUSB	Q91XQ6	LAPTM4B
P28662	BRI3	P56203	CTSW	P04187	GZMB	Q61168	LAPTM5
Q9R1W5	CALCRL	Q9WUU7	CTSZ	P01910	H2-AA	Q8BMJ2	LARS
O35350	CAPN1	Q9JLB4	CUBN	P28078	H2-DMA	P35951	LDLR
O08529	CAPN2	P35343	CXCR2	P35737	H2-DMB1	O89017	LGMN
P30730	LHCGR	Q9JLJ0	LITAF	Q9ESE1	LRBA	P24668	M6PR
Q9Z0M5	LIPA	Q8K0B2	LMBRD1	Q5S006	LRRK2	O09159	MAN2B1
O54782	MAN2B2	Q8K2I4	MANBA	Q91VR7	MAP1LC3A	Q9CQV6	MAP1LC3B

Panther annotated list of lysosomal and lysosome-associated proteins (continued):

Uniprot ID	Gene	Uniprot ID	Gene	Uniprot ID	Gene	Uniprot ID	Gene
Q6NZQ8	MARCH1	Q62086	PON2	Q99P65	SLC29A3	Q8BQU7	TMEM74
Q99M02	MARCH2	O88531	PPT1	P97441	SLC30A3	Q9D709	TMEM79
Q8BRX9	MARCH3	O35448	PPT2	O35149	SLC30A4	Q9ESN3	TMEM8A
Q9DBD2	MARCH8	Q8C4N4	PQLC2	Q8VE96	SLC35F6	Q9CR23	TMEM9
Q3TZ87	MARCH9	Q7TMR0	PRCP	Q8K4D3	SLC36A1	Q8VD00	TMEM97
Q99J21	MCOLN1	O08709	PRDX6	Q8BGD6	SLC38A9	Q9JJR8	TMEM9B
Q8K595	MCOLN2	P10820	PRF1	Q9D8M3	SLC48A1	Q60769	TNFAIP3
Q8R4F0	MCOLN3	Q9QXE5	PRSS16	Q8BXR1	SLC7A14	Q9EQJ0	TPCN1
Q8BH31	MFSD8	Q14B24	PRSS57	Q04519	SMPD1	Q8BWC0	TPCN2
Q8VC42	MIC1	Q61096	PRTN3	O09044	SNAP23	O89023	TPP1
Q3TB92	MILR1	Q61207	PSAP	Q9Z266	SNAPIN	Q02844	TPSAB1
Q8VE19	MIOS	Q8C1C1	PSAPL1	Q9WV80	SNX1	Q3LRV9	TREML4
Q8VHK5	MLC1	P49769	PSEN1	Q8BHY8	SNX14	Q8BGX0	TRIM23
P33435	MMP13	P35283	RAB12	Q8C080	SNX16	Q8R2Q0	TRIM29
P11247	MPO	Q91V41	RAB14	Q9CWK8	SNX2	Q8CJ53	TRIP10
P27573	MPZ	Q9ERI2	RAB27A	Q6P8X1	SNX6	Q91YD4	TRPM2
P02802	MT1	Q8QZZ8	RAB38	P08228	SOD1	Q61037	TSC2
Q9JLN9	MTOR	Q8BHD0	RAB39A	Q6PHU5	SORT1	Q99J59	TSPAN1
Q99104	MYO5A	P51150	RAB7A	A0A1B0GSZ0	SPAR	Q02053	UBA1
P97479	MYO7A	Q8VEA8	RAB7B	Q9JIA7	SPHK2	Q99PL6	UBXN6
Q9D7V9	NAAA	Q9R0M6	RAB9A	Q9JJF9	SPPL2A	B2RUP2	UNC13D
Q9QWR8	NAGA	Q9WUP1	RAMP3	Q3TD49	SPPL2B	Q8VCW4	UNC93B1
O09043	NAPSA	Q5ND29	RILP	A2A6C4	SPPL2C	Q8K245	UVRAG
P97432	NBR1	O35291	RNASE2	Q64337	SQSTM1	P70280	VAMP7
Q5U4H9	NCOA4	Q9D244	RNASE6	P05480	SRC	O70404	VAMP8
P57716	NCSTN	Q9CQ01	RNASET2	P13609	SRGN	Q91W86	VPS11
O35657	NEU1	O54965	RNF13	Q64704	STX3	Q920Q4	VPS16
Q9Z0J4	NOS1	Q8BG47	RNF152	O70439	STX7	Q8R307	VPS18
O35604	NPC1	A2A7Q9	RNF19B	O88983	STX8	P40336	VPS26A
Q9Z0J0	NPC2	Q8K4Q0	RPTOR	Q64324	STXBP2	Q9D2N9	VPS33A
P05125	NPPA	Q80X95	RRAGA	Q9D939	SULT1C2	P59016	VPS33B
Q9WUE4	NPRL2	Q6NTA4	RRAGB	Q9R0N3	SYT11	Q9EQH3	VPS35
Q8VIJ8	NPRL3	Q99K70	RRAGC	Q9R0N7	SYT7	Q91XD6	VPS36
Q62092	NSG1	Q7TT45	RRAGD	A2A9C3	SZT2	Q8R5L3	VPS39
P47759	NSG2	Q80U62	RUBCN	Q80VP0	TECPR1	Q5KU39	VPS41
Q62052	OCA2	O35114	SCARB2	P58681	TLR7	Q8VEJ9	VPS4A
Q8BGT0	OSTM1	Q9D1M0	SEC13	Q9EQU3	TLR9	O88384	VTI1B
P35383	P2RY2	Q8R2U0	SEH1L	P58749	TM6SF1	Q8CFJ9	WDR24
P32114	PAX2	P50405	SFTPB	Q9DBU0	TM9SF1	Q8BH57	WDR48
Q80W65	PCSK9	Q9EQ08	SGSH	Q8BJZ3	TMBIM1	Q8C0M0	WDR59
Q9CQF9	PCYOX1	Q6P7W2	SHKBP1	Q80X71	TMEM106B	Q5ND34	WDR81
P05622	PDGFRB	P70665	SIAE	Q91WN2	TMEM150A	Q3UPF5	ZC3HAV1
Q9D9G2	PEBP4	Q8CIF6	SIDT2	Q8R218	TMEM150B	Q91V17	ZNRF1
Q6PF93	PIK3C3	P41251	SLC11A1	Q8C8S3	TMEM150C	Q71FD5	ZNRF2
Q8VEB4	PLA2G15	P49282	SLC11A2	P52875	TMEM165	Q91YT8	TMEM63A
Q50L42	PLA2G4E	Q8BPX9	SLC15A3	Q9CXY1	TMEM175	Q80ZD3	SLC26A11
Q50L41	PLA2G4F	Q61983	SLC17A1	Q9CXT7	TMEM192	A2AJ88	PNPLA7
Q8VCI0	PLBD1	Q5SZA1	SLC17A2	Q5SYH2	TMEM199	Q9QY73	TMEM59
Q3TCN2	PLBD2	G5E894	SLC17A3	Q9CZX7	TMEM55A	Q8BN82	SLC17A5
Q3TB82	PLEKHF1	Q5NCM1	SLC17A4	Q3TWL2	TMEM55B	Q7TSI1	PLEKHM1

7. Bibliography

Abu-Remaileh M, Wyant GA, Kim C, Laqtom NN, Abbasi M, Chan SH, et al. Lysosomal metabolomics reveals V-ATPase- and mTOR-dependent regulation of amino acid efflux from lysosomes. Science. 2017 10;358(6364):807–813

Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature. 2003 Mar;422(6928):198

Aerts JMFG, Hollak CEM, Boot RG, Groener JEM, Maas M. Substrate reduction therapy of glycosphingolipid storage disorders. J Inherit Metab Dis. 2006 Jun;29(2–3):449–456

Alpert AJ. Electrostatic repulsion hydrophilic interaction chromatography for isocratic separation of charged solutes and selective isolation of phosphopeptides. Anal Chem. 2008 Jan 1;80(1):62–76

Anand S, Samuel M, Ang C-S, Keerthikumar S, Mathivanan S. Label-Based and Label-Free Strategies for Protein Quantitation. Methods Mol Biol Clifton NJ. 2017;1549:31–43

Andersson L, Porath J. Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography. Anal Biochem. 1986 Apr;154(1):250–254

Andrade F, Aldámiz-Echevarría L, Llarena M, Couce ML. Sanfilippo syndrome: Overall review. Pediatr Int Off J Jpn Pediatr Soc. 2015 Jun;57(3):331–338

Arai K, Kanaseki T, Ohkuma S. Isolation of highly purified lysosomes from rat liver: identification of electron carrier components on lysosomal membranes. J Biochem (Tokyo). 1991 Oct;110(4):541–547

Arrington JV, Hsu C-C, Elder SG, Andy Tao W. Recent advances in phosphoproteomics and application to neurological diseases. The Analyst. 2017 Nov 20;142(23):4373–4387

Arthur JR, Heinecke KA, Seyfried TN. Filipin recognizes both GM1 and cholesterol in GM1 gangliosidosis mouse brain. J Lipid Res. 2011 Jul 1;52(7):1345–1351

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. Nat Genet. 2000 May 1;25:25–29

Ausseil J, Landry K, Seyrantepe V, Trudel S, Mazur A, Lapointe F, et al. An acetylated 120-kDa lysosomal transmembrane protein is absent from mucopolysaccharidosis IIIC fibroblasts: a candidate molecule for MPS IIIC. Mol Genet Metab. 2006 Jan;87(1):22–31

Awad H, Khamis MM, El-Aneed A. Mass Spectrometry, Review of the Basics: Ionization. Appl Spectrosc Rev. 2015 Feb 7;50(2):158–175

Baker M. Mass spectrometry for biologists. Nat Methods. 2010 Feb;7(2):157–161

Ballabio A, Gieselmann V. Lysosomal disorders: From storage to cellular damage. Biochim Biophys Acta BBA - Mol Cell Res. 2009 Apr 1;1793(4):684–696

Bame KJ, Rome LH. Genetic evidence for transmembrane acetylation by lysosomes. Science. 1986 Sep 5;233(4768):1087–1089

Barillot E, Calzone L, Hupe P, Vert J-P, Zinovyev A, Calzone L, et al. Computational Systems Biology of Cancer. 2012

Beausoleil SA, Jedrychowski M, Schwartz D, Elias JE, Villén J, Li J, et al. Large-scale characterization of HeLa cell nuclear phosphoproteins. Proc Natl Acad Sci U S A. 2004 Aug 17;101(33):12130–12135

Beg ZH, Stonik JA, Brewer HB. 3-Hydroxy-3-methylglutaryl coenzyme A reductase: regulation of enzymatic activity by phosphorylation and dephosphorylation. Proc Natl Acad Sci. 1978 Aug 1;75(8):3678–3682

Blenis J, Resh MD. Subcellular localization specified by protein acylation and phosphorylation. Curr Opin Cell Biol. 1993 Dec 1;5(6):984–989

Boersema PJ, Raijmakers R, Lemeer S, Mohammed S, Heck AJR. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. Nat Protoc. 2009;4(4):484–494

Cadigan KM, Spillane DM, Chang TY. Isolation and characterization of Chinese hamster ovary cell mutants defective in intracellular low density lipoprotein-cholesterol trafficking. J Cell Biol. 1990 Feb;110(2):295–308

Cang C, Aranda K, Seo Y, Gasnier B, Ren D. TMEM175 Is an Organelle K(+) Channel Regulating Lysosomal Function. Cell. 2015 Aug 27;162(5):1101–1112

Cenedella RJ. Cholesterol Synthesis Inhibitor U18666A and the Role of Sterol Metabolism and Trafficking in Numerous Pathophysiological Processes. Lipids. 2009 Jun 1;44(6):477–487

Chait BT. Mass Spectrometry: Bottom-Up or Top-Down? Science. 2006 Oct 6;314(5796):65–66

Chapel A, Kieffer-Jaquinod S, Sagné C, Verdon Q, Ivaldi C, Mellal M, et al. An extended proteome map of the lysosomal membrane reveals novel potential transporters. Mol Cell Proteomics. 2013 Jun;12(6):1572–1588

Cheng F, Mani K, van den Born J, Ding K, Belting M, Fransson L-A. Nitric oxide-dependent processing of heparan sulfate in recycling S-nitrosylated glypican-1 takes place in caveolin-1-containing endosomes. J Biol Chem. 2002 Nov 15;277(46):44431–44439

Chiba T, Inoue D, Mizuno A, Komatsu T, Fujita S, Kubota H, et al. Identification and characterization of an insulin receptor substrate 4-interacting protein in rat brain: implications for longevity. Neurobiol Aging. 2009 Mar;30(3):474–482

Choudhury A, Marks DL, Pagano RE. Use of Rab GTPases to study lipid trafficking in normal and sphingolipid storage disease fibroblasts. Methods Enzymol. 2005;403:166–182

Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol Cell Proteomics. 2014 Sep;13(9):2513–2526

Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. 2008 Dec;26(12):1367–1372

Crocker AC. The cerebral defect in Tay-Sachs disease and Niemann-Pick disease. J Neurochem. 1961 Apr;7:69–80

Crocker AC, Farber S. Niemann-Pick disease: a review of eighteen patients. Medicine (Baltimore). 1958 Feb;37(1):1–95

Das AT, Tenenbaum L, Berkhout B. Tet-On Systems For Doxycycline-inducible Gene Expression. Curr Gene Ther. 2016;16(3):156–167

Davies JP, Ioannou YA. Topological analysis of Niemann-Pick C1 protein reveals that the membrane orientation of the putative sterol-sensing domain is identical to those of 3-hydroxy-3-methylglutaryl-CoA reductase and sterol regulatory element binding protein cleavage-activating protein. J Biol Chem. 2000 Aug 11;275(32):24367–24374

Dehghani A. Application and development of mass spectrometry-based approaches to investigate the regulation of lysosomal-associated proteins by phosphorylation [Dissertation]. [Bonn]: University of Bonn; 2019

Dehghani A, Gödderz M, Winter D. Tip-Based Fractionation of Batch-Enriched Phosphopeptides Facilitates Easy and Robust Phosphoproteome Analysis. J Proteome Res. 2018 Jan 5;17(1):46–54

Dema A, Perets E, Schulz MS, Deák VA, Klussmann E. Pharmacological targeting of AKAP-directed compartmentalized cAMP signalling. Cell Signal. 2015 Dec;27(12):2474–2487

Diettrich O, Mills K, Johnson AW, Hasilik A, Winchester BG. Application of magnetic chromatography to the isolation of lysosomes from fibroblasts of patients with lysosomal storage disorders. FEBS Lett. 1998 Dec 28;441(3):369–372

Dorfman A, Matalon R. The mucopolysaccharidoses (a review). Proc Natl Acad Sci U S A. 1976 Feb;73(2):630–637

Durand S, Feldhammer M, Bonneil E, Thibault P, Pshezhetsky AV. Analysis of the biogenesis of heparan sulfate acetyl-CoA:alpha-glucosaminide N-acetyltransferase provides insights into the mechanism underlying its complete deficiency in mucopolysaccharidosis IIIC. J Biol Chem. 2010 Oct 8;285(41):31233–31242

de Duve C. The lysosome turns fifty. Nat Cell Biol. 2005 Sep;7(9):847–849

de Duve C, Pressman BC, Gianetto R, Wattiaux R, Appelmans F. Tissue fractionation studies. 6. Intracellular distribution patterns of enzymes in rat-liver tissue. Biochem J. 1955 Aug;60(4):604–617

Edgren G, Havsmark B, Jönsson M, Fransson LA. Glypican (heparan sulfate proteoglycan) is palmitoylated, deglycanated and reglycanated during recycling in skin fibroblasts. Glycobiology. 1997 Feb;7(1):103–112

Ehehalt R, Keller P, Haass C, Thiele C, Simons K. Amyloidogenic processing of the Alzheimer beta-amyloid precursor protein depends on lipid rafts. J Cell Biol. 2003 Jan 6;160(1):113–123

Esko JD, Kimata K, Lindahl U. Proteoglycans and Sulfated Glycosaminoglycans. In: Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, et al., editors. Essentials of Glycobiology. 2nd ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009

Fan X, Tkachyova I, Sinha A, Rigat B, Mahuran D. Characterization of the biosynthesis, processing and kinetic mechanism of action of the enzyme deficient in mucopolysaccharidosis IIIC. PloS One. 2011;6(9):e24951

Fan X, Zhang H, Zhang S, Bagshaw RD, Tropak MB, Callahan JW, et al. Identification of the gene encoding the enzyme deficient in mucopolysaccharidosis IIIC (Sanfilippo disease type C). Am J Hum Genet. 2006 Oct;79(4):738–744

Fedele AO. Sanfilippo syndrome: causes, consequences, and treatments. Appl Clin Genet. 2015;8:269–281

Fedele AO, Hopwood JJ. Functional analysis of the HGSNAT gene in patients with mucopolysaccharidosis IIIC (Sanfilippo C Syndrome). Hum Mutat. 2010 Jul;31(7):1574-1586

Fedele AO, Isenmann S, Kamei M, Snel MF, Trim PJ, Proud CG, et al. Lysosomal N-acetyltransferase interacts with ALIX and is detected in extracellular vesicles. Biochim Biophys Acta Mol Cell Res. 2018 Jul 4;1865(10):1451–1464

Feldhammer M, Durand S, Mrázová L, Boucher R-M, Laframboise R, Steinfeld R, et al. Sanfilippo syndrome type C: mutation spectrum in the heparan sulfate acetyl-CoA: alpha-glucosaminide N-acetyltransferase (HGSNAT) gene. Hum Mutat. 2009a Jun;30(6):918–925

Feldhammer M, Durand S, Pshezhetsky AV. Protein misfolding as an underlying molecular defect in mucopolysaccharidosis III type C. PLoS ONE. 2009b Oct 13;4(10):e7434

Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass spectrometry of large biomolecules. Science. 1989 Oct 6;246(4926):64-71

Fransson L-A, Belting M, Cheng F, Jönsson M, Mani K, Sandgren S. Novel aspects of glypican glycobiology. Cell Mol Life Sci CMLS. 2004 May;61(9):1016–1024

Fuller M, Meikle PJ, Hopwood JJ. Epidemiology of lysosomal storage diseases: an overview. In: Mehta A, Beck M, Sunder-Plassmann G, editors. Fabry Disease: Perspectives from 5 Years of FOS. Oxford: Oxford PharmaGenesis; 2006

Garver WS, Hossain GS, Winscott MM, Heidenreich RA. The Npc1 mutation causes an altered expression of caveolin-1, annexin II and protein kinases and phosphorylation of caveolin-1 and annexin II in murine livers. Biochim Biophys Acta. 1999 Feb 24;1453(2):193-206

Gelsthorpe ME, Baumann N, Millard E, Gale SE, Langmade SJ, Schaffer JE, et al. Niemann-Pick type C1 I1061T mutant encodes a functional protein that is selected for endoplasmic reticulum-associated degradation due to protein misfolding. J Biol Chem. 2008 Mar 28;283(13):8229-8236

Gibson TJ, Seiler M, Veitia RA. The transience of transient overexpression. Nat Methods. 2013 Aug;10(8):715-721

Gonzalez A, Valeiras M, Sidransky E, Tayebi N. Lysosomal integral membrane protein-2: a new player in lysosome-related pathology. Mol Genet Metab. 2014 Feb;111(2):84-91

Gossen M, Bujard H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc Natl Acad Sci U S A. 1992 Jun 15;89(12):5547-5551

Guo J, Zeng W, Chen Q, Lee C, Chen L, Yang Y, et al. Structure of the voltage-gated two-pore channel TPC1 from *Arabidopsis thaliana*. Nature. 2016 Mar;531(7593):196-201

Han G, Ye M, Zhou H, Jiang X, Feng S, Jiang X, et al. Large-scale phosphoproteome analysis of human liver tissue by enrichment and fractionation of phosphopeptides with strong anion exchange chromatography. Proteomics. 2008 Apr;8(7):1346-1361

Higgins ME, Davies JP, Chen FW, Ioannou YA. Niemann-Pick C1 is a late endosome-resident protein that transiently associates with lysosomes and the trans-Golgi network. Mol Genet Metab. 1999 Sep;68(1):1-13

Hirst W, Cox RA. The construction and analysis of sucrose gradients for use with zonal rotors. Biochem J. 1976 Nov;159(2):259-265

Hollak CEM, de Sonnaville ESV, Cassiman D, Linthorst GE, Groener JE, Morava E, et al. Acid sphingomyelinase (Asm) deficiency patients in The Netherlands and Belgium: disease spectrum and natural course in attenuated patients. Mol Genet Metab. 2012 Nov;107(3):526-533

Hornbeck PV, Zhang B, Murray B, Kornhauser JM, Latham V, Skrzypek E. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Res. 2015 Jan;43(Database issue):512-520

Hrebícek M, Mrázová L, Seyrantepe V, Durand S, Roslin NM, Nosková L, et al. Mutations in TMEM76* cause mucopolysaccharidosis IIIC (Sanfilippo C syndrome). Am J Hum Genet. 2006 Nov;79(5):807-819

Huang X, Suyama K, Buchanan J, Zhu AJ, Scott MP. A Drosophila model of the Niemann-Pick type C lysosome storage disease: dnpc1a is required for molting and sterol homeostasis. Dev Camb Engl. 2005 Nov;132(22):5115-5124

Hughes DA, Nicholls K, Shankar SP, Sunder-Plassmann G, Koeller D, Nedd K, et al. Oral pharmacological chaperone migalastat compared with enzyme replacement therapy in Fabry disease: 18-month results from the randomised phase III ATTRACT study. J Med Genet. 2017;54(4):288-296

Hurley JH, Dean AM, Sohl JL, Koshland DE, Stroud RM. Regulation of an enzyme by phosphorylation at the active site. Science. 1990 Aug 31;249(4972):1012-1016

Itakura E, Kishi-Itakura C, Mizushima N. The hairpin-type tail-anchored SNARE syntaxin 17 targets to autophagosomes for fusion with endosomes/lysosomes. Cell. 2012 Dec 7;151(6):1256–1269

Jinn S, Blauwendraat C, Toolan D, Gretzula CA, Drolet RE, Smith S, et al. Functionalization of the TMEM175 p.M393T Variant as a risk factor for Parkinson Disease. Hum Mol Genet. 2019 Jul 1:ddz136

Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem. 1988 Oct 15;60(20):2299-2301

Karten B, Vance DE, Campenot RB, Vance JE. Trafficking of cholesterol from cell bodies to distal axons in Niemann Pick C1-deficient neurons. J Biol Chem. 2003 Feb 7;278(6):4168-4175

Kirkin V, Lamark T, Sou Y-S, Bjørkøy G, Nunn JL, Bruun J-A, et al. A role for NBR1 in autophagosomal degradation of ubiquitinated substrates. Mol Cell. 2009 Feb 27;33(4):505–516

Klockenbusch C, Kast J. Optimization of formaldehyde cross-linking for protein interaction analysis of non-tagged integrin beta1. J Biomed Biotechnol. 2010;2010:927585

Ko DC, Gordon MD, Jin JY, Scott MP. Dynamic movements of organelles containing Niemann-Pick C1 protein: NPC1 involvement in late endocytic events. Mol Biol Cell. 2001 Mar;12(3):601–614

Kowalewski B, Lamanna WC, Lawrence R, Damme M, Stroobants S, Padva M, et al. Arylsulfatase G inactivation causes loss of heparan sulfate 3-O-sulfatase activity and mucopolysaccharidosis in mice. Proc Natl Acad Sci. 2012 Jun 26;109(26):10310–10315

Krebs EG. The Enzymology of Control by Phosphorylation. In: Boyer PD, Krebs EG, editors. The Enzymes 1986. p. 3–20

Kresse H, Von Figura K, Klein U. A new biochemical subtype of the Sanfilippo syndrome: characterization of the storage material in cultured fibroblasts of Sanfilippo C patients. Eur J Biochem. 1978 Dec;92(2):333–339

Kwon HJ, Abi-Mosleh L, Wang ML, Deisenhofer J, Goldstein JL, Brown MS, et al. Structure of N-Terminal Domain of NPC1 Reveals Distinct Subdomains for Binding and Transfer of Cholesterol. Cell. 2009 Jun 26;137(7):1213–1224

Laemmli UK. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Nature. 1970 Aug;227(5259):680

Lange Y, Ye J, Chin J. The Fate of Cholesterol Exiting Lysosomes. J Biol Chem. 1997 Jul 4;272(27):17018–17022

Lange Y, Ye J, Rigney M, Steck TL. Regulation of endoplasmic reticulum cholesterol by plasma membrane cholesterol. J Lipid Res. 1999 Dec;40(12):2264–2270

Lange Y, Ye J, Rigney M, Steck TL. Dynamics of lysosomal cholesterol in Niemann-Pick type C and normal human fibroblasts. J Lipid Res. 2002 Feb 1;43(2):198–204

Larsen MR, Thingholm TE, Jensen ON, Roepstorff P, Jørgensen TJD. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. Mol Cell Proteomics. 2005 Jul;4(7):873–886

Li Z, Michael CoIP, Zhou D, Nagy A, Rini JM. Simple piggyBac transposon-based mammalian cell expression system for inducible protein production. Proc Natl Acad Sci U S A. 2013 Mar 26;110(13):5004–5009

Liscum L, Klansek JJ. Niemann-Pick disease type C. Curr Opin Lipidol. 1998 Apr;9(2):131–135

Liscum L, Ruggiero RM, Faust JR. The intracellular transport of low density lipoprotein-derived cholesterol is defective in Niemann-Pick type C fibroblasts. J Cell Biol. 1989 May;108(5):1625–1636

Liu Y-W, Liu X-L, Kong L, Zhang M-Y, Chen Y-J, Zhu X, et al. Neuroprotection of quercetin on central neurons against chronic high glucose through enhancement of Nrf2/ARE/glyoxalase-1 pathway mediated by phosphorylation regulation. Biomed Pharmacother. 2019 Jan 1;109:2145–2154

Lloyd-Evans E, Morgan AJ, He X, Smith DA, Elliot-Smith E, Sillence DJ, et al. Niemann-Pick disease type C1 is a sphingosine storage disease that causes deregulation of lysosomal calcium. Nat Med. 2008 Nov;14(11):1247–1255

Lowry OH, Rosebrough NJ, Farr AL, Randall RJ. Protein measurement with the Folin phenol reagent. J Biol Chem. 1951 Nov;193(1):265–275

Malathi K, Higaki K, Tinkelenberg AH, Balderes DA, Almanzar-Paramio D, Wilcox LJ, et al. Mutagenesis of the putative sterol-sensing domain of yeast Niemann Pick C-related protein reveals a primordial role in subcellular sphingolipid distribution. J Cell Biol. 2004 Feb 16;164(4):547–556

Malmersjö S, Di Palma S, Diao J, Lai Y, Pfuetzner RA, Wang AL, et al. Phosphorylation of residues inside the SNARE complex suppresses secretory vesicle fusion. EMBO J. 2016 15;35(16):1810–1821

Mani K, Cheng F, Fransson L-Å. Defective nitric oxide-dependent, deaminative cleavage of glypican-1 heparan sulfate in Niemann–Pick C1 fibroblasts. Glycobiology. 2006 Aug 1;16(8):711–718

Marques ARA, Saftig P. Lysosomal storage disorders – challenges, concepts and avenues for therapy: beyond rare diseases. J Cell Sci. 2019 Jan 15;132(2):jcs221739

McNulty DE, Annan RS. Hydrophilic interaction chromatography reduces the complexity of the phosphoproteome and improves global phosphopeptide isolation and detection. Mol Cell Proteomics. 2008 May;7(5):971–980

Meikle PJ, Hopwood JJ, Clague AE, Carey WF. Prevalence of lysosomal storage disorders. JAMA. 1999 Jan 20;281(3):249–254

Meikle PJ, Whittle AM, Hopwood JJ. Human acetyl-coenzyme A:alpha-glucosaminide N-acetyltransferase. Kinetic characterization and mechanistic interpretation. Biochem J. 1995 May 15;308 (Pt 1):327–333

Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, et al. PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. Nucleic Acids Res. 2017 04;45(D1):D183–189

Mi H, Muruganujan A, Casagrande JT, Thomas PD. Large-scale gene function analysis with the PANTHER classification system. Nat Protoc. 2013 Aug;8(8):1551–1566

Miyawaki S, Mitsuoka S, Sakiyama T, Kitagawa T. Sphingomyelinosis, a new mutation in the mouse: a model of Niemann-Pick disease in humans. J Hered. 1982 Aug;73(4):257–263

Mumby M, Brekken D. Phosphoproteomics: new insights into cellular signaling. Genome Biol. 2005;6(9):230

Nakajima M, DeChavigny A, Johnson CE, Hamada J, Stein CA, Nicolson GL. Suramin. A potent inhibitor of melanoma heparanase and invasion. J Biol Chem. 1991 May 25;266(15):9661–9666

Nicot A-S, Lo Verso F, Ratti F, Pilot-Storck F, Streichenberger N, Sandri M, et al. Phosphorylation of NBR1 by GSK3 modulates protein aggregation. Autophagy. 2014 Jun;10(6):1036–1053

Nielsen GK, Dagnaes-Hansen F, Holm IE, Meaney S, Symula D, Andersen NT, et al. Protein replacement therapy partially corrects the cholesterol-storage phenotype in a mouse model of Niemann-Pick type C2 disease. PloS One. 2011;6(11):e27287

Obayashi T, Kagaya Y, Aoki Y, Tadaka S, Kinoshita K. COXPRESdb v7: a gene coexpression database for 11 animal species supported by 23 coexpression platforms for technical evaluation and evolutionary inference. Nucleic Acids Res. 2019 Jan 8;47(D1):D55–62

Oliveros JC. Venny. An interactive tool for comparing lists with Venn's diagrams. 2007. Available from: http://bioinfogp.cnb.csic.es/tools/venny/

Olsen JV, Vermeulen M, Santamaria A, Kumar C, Miller ML, Jensen LJ, et al. Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. Sci Signal. 2010 Jan 12;3(104):ra3

Omasits U, Ahrens CH, Müller S, Wollscheid B. Protter: interactive protein feature visualization and integration with experimental proteomic data. Bioinforma Oxf Engl. 2014 Mar 15;30(6):884–886

Ong Q-R, Lim M-L, Chua C-C, Cheung NS, Wong B-S. Impaired insulin signaling in an animal model of Niemann-Pick Type C disease. Biochem Biophys Res Commun. 2012a Aug 3;424(3):482–487

Ong S-E, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics MCP. 2002 May;1(5):376–386

Ong S-E, Mann M. Mass spectrometry-based proteomics turns quantitative. Nat Chem Biol. 2005 Oct;1(5):252–262

Ong S-E, Mann M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). Nat Protoc. 2006;1(6):2650–2660

Pacheco CD, Kunkel R, Lieberman AP. Autophagy in Niemann-Pick C disease is dependent upon Beclin-1 and responsive to lipid trafficking defects. Hum Mol Genet. 2007 Jun 15;16(12):1495–1503

Padyana AK, Gross S, Jin L, Cianchetta G, Narayanaswamy R, Wang F, et al. Structure and inhibition mechanism of the catalytic domain of human squalene epoxidase. Nat Commun. 2019 Jan 9;10(1):97

Parker BL, Yang G, Humphrey SJ, Chaudhuri R, Ma X, Peterman S, et al. Targeted phosphoproteomics of insulin signaling using data-independent acquisition mass spectrometry. Sci Signal. 2015 Jun 9;8(380):rs6

Peng L, Pan P, Chen J, Yu X, Wu J, Chen Y. A tetracycline-inducible CRISPR/Cas9 system, targeting two long non-coding RNAs, suppresses the malignant behavior of bladder cancer cells. Oncol Lett. 2018 Oct;16(4):4309–4316

Platt FM, d'Azzo A, Davidson BL, Neufeld EF, Tifft CJ. Lysosomal storage diseases. Nat Rev Dis Primer. 2018 Oct 1;4(1):27

Platt FM, Boland B, van der Spoel AC. Lysosomal storage disorders: The cellular impact of lysosomal dysfunction. J Cell Biol. 2012 Nov 26;199(5):723–734

Pu J, Guardia CM, Keren-Kaplan T, Bonifacino JS. Mechanisms and functions of lysosome positioning. J Cell Sci. 2016 Dec 1;129(23):4329–4339

Pu J, Schindler C, Jia R, Jarnik M, Backlund P, Bonifacino JS. BORC, a Multisubunit Complex that Regulates Lysosome Positioning. Dev Cell. 2015 Apr 20;33(2):176–188

Puertollano R, Ferguson SM, Brugarolas J, Ballabio A. The complex relationship between TFEB transcription factor phosphorylation and subcellular localization. EMBO J. 2018 01;37(11)

Puth M-T, Neuhäuser M, Ruxton GD. Effective use of Pearson's product–moment correlation coefficient. Anim Behav. 2014 Jul 1;93:183–189

Rappsilber J, Ishihama Y, Mann M. Stop and Go Extraction Tips for Matrix-Assisted Laser Desorption/Ionization, Nanoelectrospray, and LC/MS Sample Pretreatment in Proteomics. Anal Chem. 2003 Feb;75(3):663–670

Rauniyar N, Subramanian K, Lavallée-Adam M, Martínez-Bartolomé S, Balch WE, Yates JR. Quantitative Proteomics of Human Fibroblasts with I1061T Mutation in Niemann–Pick C1 (NPC1) Protein Provides Insights into the Disease Pathogenesis. Mol Cell Proteomics. 2015 Jul;14(7):1734–1749

Reinartz M, Raupach A, Kaisers W, Gödecke A. AKT1 and AKT2 induce distinct phosphorylation patterns in HL-1 cardiac myocytes. J Proteome Res. 2014 Oct 3;13(10):4232–4245

Robles MS, Humphrey SJ, Mann M. Phosphorylation Is a Central Mechanism for Circadian Control of Metabolism and Physiology. Cell Metab. 2017 10;25(1):118–127

van Rooden EJ, van Esbroeck ACM, Baggelaar MP, Deng H, Florea BI, Marques ARA, et al. Chemical Proteomic Analysis of Serine Hydrolase Activity in Niemann-Pick Type C Mouse Brain. Neurosci. 2018 Jul 3 (12)

Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, et al. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics. 2004 Dec;3(12):1154–1169

de Ruijter J, Maas M, Janssen A, Wijburg FA. High prevalence of femoral head necrosis in Mucopolysaccharidosis type III (Sanfilippo disease): a national, observational, cross-sectional study. Mol Genet Metab. 2013 May;109(1):49–53

Ruivo R, Anne C, Sagné C, Gasnier B. Molecular and cellular basis of lysosomal transmembrane protein dysfunction. Biochim Biophys Acta. 2009 Apr;1793(4):636–649

Rush J, Moritz A, Lee KA, Guo A, Goss VL, Spek EJ, et al. Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. Nat Biotechnol. 2005 Jan;23(1):94–101

Sacco F, Humphrey SJ, Cox J, Mischnik M, Schulte A, Klabunde T, et al. Glucose-regulated and drug-perturbed phosphoproteome reveals molecular mechanisms controlling insulin secretion. Nat Commun. 2016 14;7:13250

Saftig P, Klumperman J. Lysosome biogenesis and lysosomal membrane proteins: trafficking meets function. Nat Rev Mol Cell Biol. 2009 Sep;10(9):623–635

Saito T, Nah J, Oka S-I, Mukai R, Monden Y, Maejima Y, et al. An alternative mitophagy pathway mediated by Rab9 protects the heart against ischemia. J Clin Invest. 2019 Feb 1;129(2):802–819

Saito Y, Suzuki K, Hulette CM, Murayama S. Aberrant phosphorylation of alpha-synuclein in human Niemann-Pick type C1 disease. J Neuropathol Exp Neurol. 2004 Apr;63(4):323–328

Sano H, Liu SCH, Lane WS, Piletz JE, Lienhard GE. Insulin Receptor Substrate 4 Associates with the Protein IRAS. J Biol Chem. 2002 May 31;277(22):19439–19447

Sasisekharan R, Venkataraman G. Heparin and heparan sulfate: biosynthesis, structure and function. Curr Opin Chem Biol. 2000 Dec 1;4(6):626–631

Sawamura N, Gong JS, Garver WS, Heidenreich RA, Ninomiya H, Ohno K, et al. Site-specific phosphorylation of tau accompanied by activation of mitogen-

activated protein kinase (MAPK) in brains of Niemann-Pick type C mice. J Biol Chem. 2001 Mar 30;276(13):10314–10319

Schmid-Burgk JL, Schmidt T, Gaidt MM, Pelka K, Latz E, Ebert TS, et al. OutKnocker: a web tool for rapid and simple genotyping of designer nuclease edited cell lines. Genome Res. 2014 Oct 1;24(10):1719–1723

Schmidt A, Forne I, Imhof A. Bioinformatic analysis of proteomics data. BMC Syst Biol. 2014 Mar 13;8(2):S3

Schroeder MJ, Shabanowitz J, Schwartz JC, Hunt DF, Coon JJ. A Neutral Loss Activation Method for Improved Phosphopeptide Sequence Analysis by Quadrupole Ion Trap Mass Spectrometry. Anal Chem. 2004 Jul;76(13):3590–3598

Simons M, Keller P, De Strooper B, Beyreuther K, Dotti CG, Simons K. Cholesterol depletion inhibits the generation of beta-amyloid in hippocampal neurons. Proc Natl Acad Sci U S A. 1998 May 26;95(11):6460–6464

Slámová K, Bojarová P, Gerstorferová D, Fliedrová B, Hofmeisterová J, Fiala M, et al. Sequencing, cloning and high-yield expression of a fungal β-N-acetylhexosaminidase in Pichia pastoris. Protein Expr Purif. 2012 Mar;82(1):212–217

Sleat DE, Wiseman JA, El-Banna M, Price SM, Verot L, Shen MM, et al. Genetic evidence for nonredundant functional cooperativity between NPC1 and NPC2 in lipid transport. Proc Natl Acad Sci U S A. 2004 Apr 20;101(16):5886–5891

Sleat DE, Wiseman JA, Sohar I, El –Banna M, Zheng H, Moore DF, et al. Proteomic analysis of mouse models of Niemann-Pick C disease reveals alterations in the steady-state levels of lysosomal proteins within the brain. Proteomics. 2012 Dec;12(0):3499–3509

Steen H, Jebanathirajah JA, Rush J, Morrice N, Kirschner MW. Phosphorylation Analysis by Mass Spectrometry: Myths, Facts, and the Consequences for Qualitative and Quantitative Measurements. Mol Cell Proteomics. 2006 Jan 1;5(1):172–181

Steen H, Mann M. The ABC's (and XYZ's) of peptide sequencing. Nat Rev Mol Cell Biol. 2004 Sep;5(9):699–711

Strauss K, Goebel C, Runz H, Möbius W, Weiss S, Feussner I, et al. Exosome secretion ameliorates lysosomal storage of cholesterol in Niemann-Pick type C disease. J Biol Chem. 2010 Aug 20;285(34):26279–26288

Stringer SE, Gallagher JT. Heparan sulphate. Int J Biochem Cell Biol. 1997 May;29(5):709–714

Stuart SA, Houel S, Lee T, Wang N, Old WM, Ahn NG. A Phosphoproteomic Comparison of B-RAFV600E and MKK1/2 Inhibitors in Melanoma Cells. Mol Cell Proteomics MCP. 2015 Jun;14(6):1599–1615

Sugimoto Y, Ninomiya H, Ohsaki Y, Higaki K, Davies JP, Ioannou YA, et al. Accumulation of cholera toxin and GM1 ganglioside in the early endosome of

Niemann-Pick C1-deficient cells. Proc Natl Acad Sci U S A. 2001 Oct 23;98(22):12391–12396

Suzuki R, Lee K, Jing E, Biddinger SB, McDonald JG, Montine TJ, et al. Diabetes and insulin in regulation of brain cholesterol metabolism. Cell Metab. 2010 Dec 1;12(6):567–579

Sym M, Basson M, Johnson C. A model for Niemann–Pick type C disease in the nematode Caenorhabditis elegans. Curr Biol. 2000 May 1;10(9):527–530

Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019 Jan 8;47(D1):D607–613

Tamari F, Chen FW, Li C, Chaudhari J, Ioannou YA. PKC Activation in Niemann Pick C1 Cells Restores Subcellular Cholesterol Transport. Buratti E, editor. PLoS ONE. 2013 Aug 15;8(8):e74169

Tharkeshwar AK, Trekker J, Vermeire W, Pauwels J, Sannerud R, Priestman DA, et al. A novel approach to analyze lysosomal dysfunctions through subcellular proteomics and lipidomics: the case of NPC1 deficiency. Sci Rep. 2017 Jan 30;7:41408

Thelen M, Winter D, Braulke T, Gieselmann V. SILAC-Based Comparative Proteomic Analysis of Lysosomes from Mammalian Cells Using LC-MS/MS. Methods Mol Biol Clifton NJ. 2017;1594:1–18

Thompson A, Schäfer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, et al. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal Chem. 2003 Apr 15;75(8):1895–1904

Troyanovsky B, Levchenko T, Månsson G, Matvijenko O, Holmgren L. Angiomotin: an angiostatin binding protein that regulates endothelial cell migration and tube formation. J Cell Biol. 2001 Mar 19;152(6):1247–1254

Tumova S, Woods A, Couchman JR. Heparan sulfate proteoglycans on the cell surface: versatile coordinators of cellular functions. Int J Biochem Cell Biol. 2000 Mar;32(3):269–288

Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat Methods. 2016 Sep;13(9):731–740

Valstar MJ, Ruijter GJG, van Diggelen OP, Poorthuis BJ, Wijburg FA. Sanfilippo syndrome: a mini-review. J Inherit Metab Dis. 2008 Apr;31(2):240–252

Vance JE, Karten B. Niemann-Pick C disease and mobilization of lysosomal cholesterol by cyclodextrin. J Lipid Res. 2014 Aug;55(8):1609–1621

Vanier MT. Niemann-Pick disease type C. Orphanet J Rare Dis. 2010 Jun 3;5:16

Vanier MT. Complex lipid trafficking in Niemann-Pick disease type C. J Inherit Metab Dis. 2015 Jan;38(1):187–199

Vilar RE, Ghael D, Li M, Bhagat DD, Arrigo LM, Cowman MK, et al. Nitric oxide degradation of heparin and heparan sulphate. Biochem J. 1997 Jun 1;324 (Pt 2):473–479

Villén J, Gygi SP. The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. Nat Protoc. 2008;3(10):1630–1638

Vitner EB, Platt FM, Futerman AH. Common and uncommon pathogenic cascades in lysosomal storage diseases. J Biol Chem. 2010 Jul 2;285(27):20423–20427

Voznyi YV, Karpova EA, Dudukina TV, Tsvetkova IV, Boer AM, Janse HC, et al. A fluorimetric enzyme assay for the diagnosis of Sanfilippo disease C (MPS III C). J Inherit Metab Dis. 1993 Mar;16(2):465–472

Wahrle S, Das P, Nyborg AC, McLendon C, Shoji M, Kawarabayashi T, et al. Cholesterol-Dependent γ-Secretase Activity in Buoyant Cholesterol-Rich Membrane Microdomains. Neurobiol Dis. 2002 Feb 1;9(1):11–23

Walker MW, Lloyd-Evans E. A rapid method for the preparation of ultrapure, functional lysosomes using functionalized superparamagnetic iron oxide nanoparticles. Methods Cell Biol. 2015;126:21–43

Watari H, Blanchette-Mackie EJ, Dwyer NK, Sun G, Glick JM, Patel S, et al. NPC1-containing compartment of human granulosa-lutein cells: a role in the intracellular trafficking of cholesterol supporting steroidogenesis. Exp Cell Res. 2000 Feb 25;255(1):56–66

Wattiaux R, Wattiaux-De Coninck S, Ronveaux-dupal MF, Dubois F. Isolation of rat liver lysosomes by isopycnic centrifugation in a metrizamide gradient. J Cell Biol. 1978 Aug;78(2):349–368

Wattiaux R, Wibo M, Baudhuin P. Effect of the injection of Triton WR 1339 on the hepatic lysosomes of the rat. Arch Int Physiol Biochim. 1963 Jan;71:140–142

Wendeler M, Sandhoff K. Hexosaminidase assays. Glycoconj J. 2009 Nov;26(8):945–952

Whitfield GB, Brock TD, Ammann A, Gottlieb D, Carter HE. Filipin, an Antifungal Antibiotic: Isolation and Properties. J Am Chem Soc. 1955 Sep 1;77(18):4799–4801

Wijburg FA, Węgrzyn G, Burton BK, Tylki-Szymańska A. Mucopolysaccharidosis type III (Sanfilippo syndrome) and misdiagnosis of idiopathic developmental delay, attention deficit/hyperactivity disorder or autism spectrum disorder. Acta Paediatr Oslo Nor 1992. 2013 May;102(5):462–470

Wilson-Grady JT, Haas W, Gygi SP. Quantitative comparison of the fasted and re-fed mouse liver phosphoproteomes using lower pH reductive dimethylation. Methods San Diego Calif. 2013 Jun 15;61(3):277–286

Winter D, Seidler J, Ziv Y, Shiloh Y, Lehmann WD. Citrate Boosts the Performance of Phosphopeptide Analysis by UPLC-ESI-MS/MS. J Proteome Res. 2009 Jan 2;8(1):418–424

Wu X, Tian L, Li J, Zhang Y, Han V, Li Y, et al. Investigation of receptor interacting protein (RIP3)-dependent protein phosphorylation by quantitative phosphoproteomics. Mol Cell Proteomics MCP. 2012 Dec;11(12):1640–1651

Xu H, Ren D. Lysosomal Physiology. Annu Rev Physiol. 2015 Feb 10;77(1):57–80

Xu J, Dang Y, Ren YR, Liu JO. Cholesterol trafficking is required for mTOR activation in endothelial cells. Proc Natl Acad Sci U S A. 2010 Mar 9;107(10):4764–4769

Yogalingam G, Hopwood JJ. Molecular genetics of mucopolysaccharidosis type IIIA and IIIB: Diagnostic, clinical, and biological implications. Hum Mutat. 2001 Oct;18(4):264–281

Zerenturk EJ, Sharpe LJ, Ikonen E, Brown AJ. Desmosterol and DHCR24: unexpected new directions for a terminal step in cholesterol synthesis. Prog Lipid Res. 2013 Oct;52(4):666–680

Zhang Y, Fonslow BR, Shan B, Baek M-C, Yates JR. Protein analysis by shotgun/bottom-up proteomics. Chem Rev. 2013 Apr 10;113(4):2343–2394

Zimran A. How I treat Gaucher disease. Blood. 2011 Aug 11;118(6):1463–1471

Zoncu R, Bar-Peled L, Efeyan A, Wang S, Sancak Y, Sabatini DM. mTORC1 Senses Lysosomal Amino Acids Through an Inside-Out Mechanism That Requires the Vacuolar H+-ATPase. Science. 2011 Nov 4;334(6056):678–683

8. List of figures

9. List of tables

10. Acknowledgement

All praises go to Almighty who empowered us with knowledge. To fulfill my dream to be a doctorate, I have been supported and inspired by many people without whom it would not be possible to finish. Therefore, with the completion of my dissertation, I would like to take the opportunity to thank those people.

First of all, I would like to thank Prof. Dr. Volkmar Gieselmann, for accepting me as a Ph.D. student and being the first referee of my dissertation. His valuable advice and ideas regarding my project enlightened me as a researcher and allowed me to troubleshoot any kind of research-oriented issues. In addition, I would like to thank Prof. Dr. Walter Witke for being my second referee as well as Prof. Dr. Wolfram S. Kunz and Prof. Dr. Christian Steinhäuser for being the members of the doctoral committee.

Especially, I am obliged to my supervisor Dr. Dominic Winter. It would not be possible to complete my doctoral study without his scholastic supervision. Throughout this four year, since applying to DAAD scholarship until finishing my doctoral program, he consistently trained me how to deal with scientific research perfectly. I owe him for his infinite advice, ideas, suggestions, constructive criticisms and most importantly his valuable time.

Furthermore, I would like to thank Deutscher Akademischer Austauschdienst (DAAD) for generous funding of my doctoral project.

I am thankful to all my colleagues in the AG Winter group: Dr. Carmen Schoor, Alireza Dehghani, Shiva Ahmadi, Peter Mosen, Elham Pourbarkhordari, Jasjot Singh, Asisa Muchamedin, Inga Titova and Dr. Robert Hardt for their support, fruitful discussions, and all the fun we had in the last four years. Special thanks go to Srigayatri Ponnaiyan for being my continuous mental support since day one in Bonn. I would like to thank Andreas Verhülsdonk for assisting me in fluorescence microscopy experiments. I am also thankful to Lisa Sofie Banning for proofreading my thesis. Moreover, I am grateful to Norbert Rösel for his unlimited support and care. Norbert was not just a technical assistant for me; he was like a family member to me here in Germany. My gratitude also goes to Simone Arndt, Claudia Yaghootfam, Heidi Simonis, Jutta Mülich, Ralph Mahlberg, Karola Ragut, and Ellen Reininger for their technical and administrative support. I feel thankful to all my other colleagues of the AG Eckhardt, AG Matzner, AG Sylvester, AG Thelen and also to Dr. Christof Völker for their support and the pleasant time we spent together.

Last but most importantly, I would like to thank my parents and siblings for being the biggest support in my life. Today I deeply recall my father, who believed in me and let me be what I wanted to be. Finally, my gratefulness goes to my husband, Md. Nur A Alam Siddique, for becoming a part of my life. I adore him for his unconditional love, support, and inspiration which allow me to move forward and restrain from giving up.