Université de Montréal

# N-acétyltransférase lysosomale: organisation, fonctionnement et défauts moléculaires chez les patients atteints du syndrome de Sanfilippo type C.

Par

Matthew Feldhammer

Département de Biochimie

Faculté de Médecine

Mémoire présenté à la Faculté des études supérieures

en vue de l'obtention du grade de Maîtrise ès sciences

en Biochimie

Décembre, 2009

© Matthew Feldhammer, 2009

Université de Montréal

Faculté des études supérieures

Ce mémoire intitulé :

N-acétyltransférase lysosomale: organisation, fonctionnement et défauts moléculaires chez les patients atteints du syndrome de Sanfilippo type C.

Présenté par :

Matthew Feldhammer

a été évalué par un jury composé des personnes suivantes :

Dr. Jurgen Sygusch
**président-rapporteur**

Dr. Alexey Pshezhetsky
**directeur de recherche**

Dr. Andréa Richter
**membre du jury**

# Résumé

L’acétylation des résidus de glucosamine terminaux par la N-acétyltransférase lysosomale (HGSNAT) est une étape essentielle de la dégradation catabolique de l’héparan sulfate. Des défauts dans cette réaction causent une maladie de surcharge lysosomale autosomale récessive rare : le désordre de Sanfilippo type C (SFC). À ce jour, 54 mutations ont été rapportées chez des patients SFC, incluant 13 mutations des sites d’épissage, 11 insertions et délétions, 8 mutations non-sens, 18 mutations faux-sens et 4 polymorphismes, avec différentes manifestations phénotypiques. Nous avons identifié 10 d’entre elles et effectué une étude exhaustive portant sur l’éventail des mutations SFC, leur distribution dans la population de patients, ainsi que leur impact potentiel sur la structure de la HGSNAT. Les erreurs d’épissage, les mutations non-sens, les insertions et les délétions devraient toutes entraîner un ARN non fonctionnel qui est rapidement dégradé par des mécanismes de contrôle qualité cellulaire. Les 4 polymorphismes identifiés sont des changements d'acides aminés qui ne modifient pas l'activité enzymatique, la glycosylation ou la localisation et n'ont donc pas de signification au niveau clinique. Au niveau des enzymes, les polymorphismes sont des changements d’acides aminés qui n’affectent pas la fonction, mais dans un contexte d’acides nucléiques ils peuvent être considérés comme des mutations faux-sens. Les dix-huit mutations faux-sens qui ont été exprimées ont produit des protéines inactives, en raison d'erreurs dans leur repliement. Ceci expliquerait donc la progression sévère de la maladie chez les personnes porteuses de ces mutations. Les protéines mutantes mal repliées sont

anormalement glycosylées et conservées dans le réticulum endoplasmique. La thérapie par amélioration de l’activité enzymatique par des chaperonnes est une option thérapeutique potentielle, spécifiquement conçue pour exploiter l'activité enzymatique résiduelle de mutants mal repliés, afin d’éliminer les substrats stockés. Nous avons démontré que le traitement de plusieurs lignées de fibroblastes de patients SFC avec le chlorhydrate de glucosamine, un inhibiteur spécifique de la HGSNAT, a partiellement restauré l’activité de l'enzyme mutante, fournissant une preuve de l’utilité future de la thérapie par des chaperonnes dans le traitement de la maladie de SFC.

**MOTS-CLÉS :**

lysosomes, désordres métaboliques, mucopolysaccharidose IIIC, syndrome de Sanfilippo type C, analyse mutationnelle, repliement de protéines, thérapie par amélioration de l’activité enzymatique.

# ABSTRACT

The acetylation of terminal glucosamine residues by lysosomal N-acetyltransferase (HGSNAT) is an essential part of the catabolic breakdown of heparan sulfate. Defects in this reaction result in the rare autosomal recessive lysosomal storage disorder Sanfilippo syndrome type C (SFC). To date 54 mutations in SFC patients have been reported including 13 splice-site mutations, 11 insertions and deletions, 8 nonsense, 18 missense and 4 polymorphisms, with different phenotypic manifestations. We have identified 10 of them and conducted a comprehensive review discussing the spectrum of Sanfilippo C mutations, their distribution within the patient population as well as how the mutations could potentially affect the structure of HGSNAT. Splicing errors, nonsense mutations, insertions and deletions were all predicted to result in non-functional RNA which is rapidly degraded by cellular quality control mechanisms. The 4 identified polymorphisms resulted in amino acid changes which did not affect the enzyme activity, glycosylation or targeting and were therefore not clinically significant. Polymorphisms, in the context of enzymes are amino acid changes not affecting function, but in the context of nucleic acids can still be considered as missense mutations. Eighteen missense mutations were expressed and shown be inactive due to errors in protein folding providing an explanation for the severe disease progression seen in individuals with these mutations. Misfolded mutants were abnormally glycosylated and retained in the endoplasmic reticulum. Enzyme enhancement/chaperone therapy is a potential treatment option specifically designed to exploit the residual enzyme activity of

misfolded mutants in order to clear stored substrates. We demonstrated that treatment of several fibroblast lines of SFC patients with a specific inhibitor of HGSNAT; glucosamine-hydrochloride partially rescued mutant enzyme activity providing a proof of principle for the future use of chaperone therapeutics in the treatment of SFC.

**KEYWORDS:**

lysosomes, metabolic disorders, Mucopolysaccharidosis IIIC, Sanfilippo syndrome C, mutational analysis, protein folding, enzyme enhancement therapy.

# Table of Contents

**RÉSUMÉ……………………………………………………………………iii**
**ABSTRACT…………………………………………………………...............v**
**TABLE OF CONTENTS…………………………………………………vii**
**LIST OF FIGURES……………………………………………………..ix**
**LIST OF TABLES………………………………………………………x**
**ABBREVIATIONS……………………………………………….......xi**
**DEDICATION…………………………………………………………..xiii**
**ACKNOWLEDGEMENTS………………………………………………xiv**

**CHAPTER 1: LITERATURE REVIEW……………………………………1**

1.1.1 Lysosomes and lysosome biogenesis 2
1.1.2 Lysosomal membranes 3
1.1.3 Biogenesis of lysosomal enzymes 5
1.1.3.1 Sorting of soluble lysosomal proteins 5
1.1.3.2 Sorting of lysosomal membrane proteins 7
1.2.1 Lysosomal storage disorders 10
1.2.2 Lysosomal storage disease overview 12
1.2.3 Lysosomal storage disease therapies 17
1.2.3.1 Bone marrow transplantation 18
1.2.3.2 Enzyme replacement therapy 20
1.2.3.3 Enzyme enhancement/chaperone therapy 21
1.2.3.4 Substrate reduction therapy 23
1.3.1 Mucopolysacchridosis III type C 26
1.3.2 Heparan sulfate catabolism 27
1.3.3 Evidence for occurrence of a transmembrane acetylation reaction 29
1.3.4 Identification and cloning of the gene responsible for MPS IIIC 30
1.3.5 Biochemical properties of HGSNAT 33
1.3.6 Mutations in HGSNAT 35
1.3.7 **Hypotheses and objectives** **36**

**CHAPTER 2: Mutations in HGSNAT……………………………….......37**

2.1 Foreword 38
2.2 **Article 1** **39**
Sanfilippo syndrome type C: mutation spectrum in the heparan sulfate acetyl-CoA: alpha-glucosaminide N-acetyltransferase (HGSNAT) gene.

**CHAPTER 3: Characterization of molecular defects in *HGSNAT* gene…69**

3.1 Foreword 70
3.2 **Article 2** **71**
Protein misfolding as an underlying molecular defect in mucopolysaccharidosis III type C.

**CHAPTER 4: General discussion and conclusions………………………...108**
4.1 General discussion 109
4.2 Conclusions 113

**REFERENCES……………………………………………………………114**

**APPENDIX I……………………………………………………………….xv**
**APPENDIX II……………………………………………………………...xvii**
**APPENDIX III……………………………………………………………xxii**
**APPENDIX IV……………………………………………………………xxiv**

# List of Figures

**Chapter 1: Literature review**

Figure 1: Catabolic pathway for degradation of heparan sulfate………………28

Figure 2: Proposed catalytic mechanism for HGSNAT……………….............31

**Chapter 2: Mutations in HGSNAT**

Figure 1: Distribution of missense mutations in HGSNAT protein……………65

Figure 2: Location of MPS IIIC mutations in the *HGSNAT* gene……………...67

**Chapter 3: Characterization of molecular defects in *HGSNAT* gene**

Figure 1: Enzymatic activity and expression of HGSNAT mutants…………...96

Figure 2: Deglycosylation of HGSNAT by endoglycosidase H, PNGase F and tunicamycin…………………………………………………...98

Figure 3: Analysis of the wild-type recombinant HGSNAT by the anion-exchange FPLC……………………………………………………100

Figure 4: Localization of HGSNAT mutants expressed in cultured human skin fibroblasts by immunofluorescence microscopy………………..102

Figure 5: Partial refolding of HGSNAT mutants by glucosamine……………..104

Figure 6: Distribution of missense mutations in HGSNAT protein……………106

# List of Tables

**Chapter 1: Literature review**

Table 1: Lysosomal storage disorders, enzymes involved and stored substrates…………………………………………………………..14

**Chapter 2: Mutations in HGSNAT**

Table 1: Mutations in the *HGSNAT* gene………………………………………61

**ABBREVIATIONS**

| | |
|---|---|
| AcCoA | acetyl coenzyme A |
| AP | adaptor protein |
| ASSC | active-site-specific chaperones |
| BMT | bone marrow transplantation |
| CBP | calmodulin-binding peptide |
| CMA | chaperone mediated autophagy |
| CNS | central nervous system |
| EET | enzyme enhancement therapy |
| ER | endoplasmic reticulum |
| ERT | enzyme replacement therapy |
| GAG | glycosaminoglycans |
| GGA | Golgi-localized, γ-ear-containing, Arf-binding |
| $GM_2AP$ | $GM_2$ activator protein |
| GVHD | graph versus host disease |
| HGSNAT | heparan alpha-glucosaminide N-acetyltransferase |
| IV | intravenous |
| LAMP | lysosomal associated membrane proteins |
| LIMP | lysosomal integral membrane protein |
| LSD | lysosomal storage disorder |
| M6P | mannose 6-phosphate |
| M6PR | M6P receptor |

| | |
|---|---|
| MPS | mucopolysaccharidosis |
| MPS IIIC | mucopolysaccharidosis IIIC |
| NMD | nonsense-mediated mRNA decay |
| PTC | premature termination codon |
| RER | rough ER |
| SBP | streptavidin-binding peptide |
| SRT | substrate reduction therapy |
| TAP | tandem affinity purification |
| TGN | trans-Golgi network |
| UCH | ubiquitin C-terminal hydrolase |

## DEDICATION

This thesis is dedicated to the memory of Morris Kokin PhD. a mentor, friend and source of inspiration for me, and all those who were privileged to know him, you will forever be missed.

“Learn from yesterday, live for today, hope for tomorrow. The important thing is to not stop questioning.” Albert Einstein

**ACKNOWLEDGEMENTS**

This work was supported by research grants from March of Dimes Foundation, Sanfilippo Childrens Research Foundation and Canadian Institutes for Health Research (MOP84430) to Dr. Alexey V. Pshezhetsky and a salary award to Matthew Feldhammer from the faculty of graduate studies at the University of Montreal. First and foremost I have to thank my supervisor Dr. Pshezhetsky for providing me with the opportunity to work in his laboratory, for his support and guidance during the course of my studies. I am incredibly thankful to Madame Stéphanie Durand MSc. for the incredible contribution to this project that she has provided, specifically in the areas of cloning and mutagenesis, this project would not have progressed to the level it did without her support and assistance. I have to thank Dr. Volkan Seyrantepe for introducing me to the world of cell culture and providing guidance and opinion in the drafting of my manuscripts. I also have to thank Dr. Maryssa Canuel for advice on microscopy techniques which proved invaluable to these studies and my advisor Dr. Nikolaus Heveker for his support. Additionally, I have to mention the assistance of Ms Sarah Pickles MSc. who helped to edit this manuscript and has been a constant source of support for me. Finally, I have to thank my parents, my brother my aunt and my grandmother for their unwavering support, and love without whom none of this would be possible.

# CHAPTER 1

# INTRODUCTION

LITERATURE REVIEW

### 1.1.1 Lysosomes and lysosome biogenesis

Lysosomes, discovered in 1949 by Belgian researcher Christian de Duve are cytoplasmic membrane-bound organelles carrying over 100 hydrolytic enzymes, including: proteases, nucleases, lipases, glycosidases, phosphatases, sulfatases and phospholipases. All lysosomal enzymes have acidic pH optimum in the range of 5.0 and are responsible for degradation and recycling of biological macromolecules including lipids, proteins, nucleic acids, sugars, as well as bacteria and non-functioning or impaired organelles such as mitochondria. There is only one presently known lysosomal enzyme that is not a hydrolase: heparan alpha-glucosaminide N-acetyltransferase (HGSNAT) which is deficient in patients with mucopolysaccharidosis IIIC (MPS IIIC or Sanfilippo syndrome C) and will be the focus of this thesis.

There are three main pathways involved in lysosomal catabolism. The first pathway involves macromolecules taken up by receptor-mediated endocytosis. The endocytosed material, including receptors, their ligands and associated membrane is initially delivered into early endosomes. Part of the cargo is selectively retrieved and sorted to the plasma membrane though "recycling" endosomes. The rest becomes part of the late endosome with a concomitant drop in pH and fusion with other late endosomes containing acid hydrolases from the Golgi apparatus to gradually form the mature lysosome. The second pathway involves the degradation of impaired organelles such as the mitochondria in a process called autophagy. In this process, the organelle is enclosed by membranes derived from the endoplasmic reticulum (ER), thus creating an autophagosome

which fuses with Golgi-derived vesicles carrying lysosomal enzymes and becomes a mature lysosome. The last pathway involves the uptake and degradation of large particles and microorganisms like bacteria, parasites and dead or dying cells by white blood cells (monocytes, macrophages and dendritic cells) collectively known as phagocytes forming phagosomes upon engulfment of the object, which then like the autophagosomes are acidified and filled with lysosomal enzymes.

The biogenesis of lysosomes is a complex process involving the integration of the endocytic and biosynthetic pathways of the cell. Lysosomal targeting of acid hydrolases and other lysosomal proteins can be direct, from the trans-Golgi network (TGN) to the endosomal system, or indirect involving transport to the plasma membrane and subsequent endocytosis [1, 2]. Furthermore, biogenesis requires that proteins destined for the lysosome are segregated from other proteins with different subcellular destinations and transferred to developing or mature lysosomes.

### 1.1.2 Lysosomal Membrane

The lysosomal membrane is a semi-permeable barrier between the lysosomal lumen and the cytoplasm that serves to protect the cytoplasmic proteins from degradation by the acidic hydrolases of the lysosome. The membrane is equipped with multiple transport and carrier systems responsible for controlling access into and out of the lumen in addition to containing structural proteins. One of the most important and well-studied lysosomal membrane

proteins is the proton pump that is responsible for maintaining the acidity of the lysosomal lumen. The proton v-ATPase pumps protons into the lumen reducing the pH to around 5. The membrane also facilitates transportation of degradation products from the lumen to the cytosol through transporters such as cystinosin (cystein transporter) and sialin (sialic acid transporter). Over 270 proteins have been identified in purified lysosomal membrane fractions through proteomic analysis (215 in the integral membrane and 55 in the membrane associated fractions), many with yet unknown functions [3].

Similar to proteins on other biological membranes, lysosomal membrane proteins are classified as associated with or integral to the membrane. The 2 most abundant lysosomal membrane integral proteins, lysosomal associated membrane proteins (LAMPs) 1 and 2, are important both structural and functional constituents of the membrane contributing to membrane stability, in addition to participating in chaperone mediated autophagy (CMA) [4, 5]. Both LAMPs serve as canonical lysosomal membrane markers due to their ubiquitous expression pattern.

Proper functioning of lysosomal membrane proteins is essential to maintain cellular homeostasis. There are a number of disorders that can arise from defects in membrane proteins including but not limited to Nieman-Pick type C, sialic acid storage disease, cystinosis and MPS IIIC.

### 1.1.3 Biogenesis of lysosomal enzymes

#### 1.1.3.1 Sorting of soluble lysosomal proteins.

All known lysosomal proteins are synthesized in polyribosomes bound to rough ER (RER) membranes. Similar to secreted proteins, they contain an N-terminal signal peptide consisting of domains rich in positively-charged and hydrophobic amino acids which directs the ribosome to the endoplasmic reticulum membrane and initiates transport of the growing polypeptide across the ER membrane, [6] where they are modified with N-linked oligosaccharide chains (glycans).

The initial stages of glycoprotein synthesis involve a large precursor oligosaccharide $Glc_3Man_9GlcNAc_2$, assembled with a pyrophosphate link to a ligand carrier, dolichol, and then transferred to the target asparagine residue of the precursor glycoprotein. While still in the RER, the signal peptide is cleaved and the processing of the asparagine-linked (N-linked) sugar takes place. Vesicular transport then shuttles proteins to the Golgi stacks where they undergo post-translational modifications and are sorted for targeting to the lysosome, secretory granules or the plasma membrane. While traveling through the Golgi, the oligosaccharide chain residues of secretory and membrane glycoproteins are processed in a stepwise fashion to a complex type unit via the addition of N-acetylglucosamine, galactose and sialic acid to mannose residues.

In the Golgi, all soluble lysosomal glycoproteins acquire a phosphomannosyl recognition marker: mannose 6-phosphate (M6P) that mediates their future translocation to the endocytic compartment. This tag is recognized by

one of the two M6P receptors (M6PR), ubiquitously expressed in the TGN; the 300-kDa cation-independent M6PR or the 46 kDa cation-dependent M6PR [7]. Both calcium dependent and independent receptors bind M6P with approximately the same affinity (7-8x$10^{-6}$M), however the calcium-independent receptor binds diphosphorylated oligosaccharides with a much higher affinity (2x$10^{-9}$M) than the calcium dependent receptor [8, 9].

The M6P marker is added in two steps. N-acetylglucosamine-1-phosphotransferase transfers N-acetyltglucosamine-1-phosphate from UDP-GlcNAc to one or more mannose residues on the lysosomal protein to give rise to a phosphodiester intermediate [10, 11]. Then the N-acetylglucosamine-1-phosphodiester α-N-acetylglucosaminidase removes the N-acetylglucosamine residue to generate an active phosphomonoester [12-14]. The M6P-tagged glycoprotein is then ready to interact with its receptor to facilitate transport to the lysosome.

The interaction between the receptors and the M6P moieties is pH-dependent. The complex is formed at neutral or slightly acidic pH in the Golgi, whereas the low pH of the late endosome and lysosome favours dissociation of the M6P tagged enzyme from the receptor. As a result, the enzyme remains in the lysosome to carry out its function while the receptor is recycled back to either the Golgi complex or the plasma membrane where it can aid in internalizing exogenous ligands [15].

Much of the evidence of a M6PR-independent sorting mechanism comes from studying B-lymphocytes from patients with I-cell disease. I-cell disease is

characterized by a deficiency in N-acetylglucosamine-phosphotransferase activity, and results in the lack of the ability to add M6P to newly synthesized hydrolases. Despite this, B-lymphocytes from patients with I-cell disease maintain near normal levels of lysosomal enzymes, suggesting alternative targeting motifs [16]

One of these alternative routes involves the lysosomal sorting receptor: sortilin, that has been shown to be essential for the lysosomal trafficking of $GM_2$ activator protein ($GM_2AP$), a cofactor of β-hexosaminidase A and prosaposin, a precursor of small proteins, saposins A-D that facilitate the catabolism of glycolipids in the lysosome. Intracellular trafficking experiments of prosaposin and $GM_2AP$ in cells transfected with a dominant negative truncated human sortilin demonstrated that both $GM_2AP$ and prosaposin use sortilin instead of the M6PR to reach the lysosome [17, 18]. Furthermore, siRNA-mediated knockdown of sortilin blocked the trafficking of prosaposin and $GM_2AP$ to the lysosome and resulted in their retention in the Golgi [19].

#### 1.1.3.2 Sorting of lysosomal membrane proteins.

Similar to soluble lysosomal proteins, lysosomal membrane proteins are glycoproteins rich in complex N-linked oligosaccharides. They are synthesized and glycosylated in the RER and their sugars are modified in the Golgi before transport to the lysosome. However, unlike the majority of soluble lysosomal proteins, lysosomal membrane proteins are sorted independently of M6PRs [20, 21]. Interestingly, soluble lysosomal protein, acid phosphatase is trafficked to the

lysosome as a membrane bound precursor, and is cleaved from the membrane upon arrival in the lysosome where it becomes soluble [22]. The lysosomal targeting of membrane proteins involves either the internalization of proteins from the plasma membrane via clathrin-coated pits or direct transport from the Golgi.

In both cases targeting to the lysosomes involves a complex mechanism requiring association of specific amino acid motifs in the cytoplasmic domains of the protein with a μ subunit of adaptor protein (AP) complexes [23, 24]. Targeting occurs via either tyrosine or dileucine based sorting signals. Tyrosine-based sorting motifs conform to the NPXY or YXXØ (where Ø represents a bulky hydrophobic residue) consensus motifs. Dileucine-based signals consist of [DE]XXXL[LI] or DXXLL consensus motifs. YXXØ and [DE]XXXL[LI] signals are recognized by the AP-1, AP-2, AP-3 and AP-4 complexes, essential components of the protein coats associated with the cytosolic face of membranes. DXXLL signals are recognized by the Golgi-localized, γ-ear-containing, Arf-binding (GGA) family of adaptor proteins via their VHS domains. Signal recognition events are tightly controlled via phosphorylation [2].

Specifically, lysosomal integral memebrane protein (LIMP)2 is targeted to the lysosome via a dileucine-based signal within the C-terminal cytoplasmic tail which interacts with AP-3 [2]. Furthermore LIMP2 was found to be the M6P-independent trafficking receptor for beta-glucocerebrosidase [25]. LAMP-1, LAMP-2 and acid phosphatase are trafficked via the plasma membrane where their YXXØ targeting signal results in rapid internalization and lysosomal sorting

via interaction with AP-1 and AP-2. Sortilin and both calcium dependant and independent-M6PRs themselves contain the DXXLL sorting signal and are recognized by the VHS domains of GGAs which facilitates their transport to the lysosomes. Recent evidence indicates that M6PRs require both AP-1 and GGA for proper sorting [2, 26].

### 1.2.1 Lysosomal Storage Disorders

Intracellular homeostasis requires a delicate balance between degradation and synthesis of cellular components be maintained for proper cellular functioning. Defects in lysosomal enzymes or their cofactors, membrane transporters, or in biogenesis of the lysosome itself disrupt this balance and cause build-up and sequestering of the undegraded or unprocessed substrates within the lumen of the lysosome. In 1965 Belgian scientist H.G. Hers introduced the concept of lysosomal storage to describe how the deficiency of a lysosomal enzyme alpha-glucosidase could cause lethal Pompe disease. The undegraded substrate would gradually accumulate within the lysosomes, increasing their size and number, leading to cellular death and malfunctioning of the affected organ [27].

The storage of undegraded substrates usually does not occur in all cells but rather in distinct organs and tissues where most of degradation usually occurs [28]. For example, the brain is rich in glycosphingolipids because of their presence in neuronal membranes. Defective hydrolysis of these molecules occurs in $G_{M1}$ and $G_{M2}$ gangliosidoses as well as in Krabbe, Niemann-Pick types A and B and Gaucher diseases. Storage of undegraded substrates, primarily within the neurons results in their death and causes diverse neurological symptoms. In contrast, muscles which utilize glycogen as a source of energy are severely affected and become weakened in Pompe disease because glycogen breakdown requires acid alpha-glucosidase, defective in this disease.

While it is widely agreed that storage of undegraded substrates leads to cell death and organ dysfunction, the mechanism(s) underlying processes from enzyme deficiency to cell death are not completely understood. In addition to specific toxicity of particular stored compounds such as galactocerebroside stored in Krabbe disease, recent studies identified several mechanisms that may be common for different lysosomal disorders. First, mRNA profiles from lysosomal storage disorder (LSD) patient fibroblasts identified decreased expression of the ubiquitin C-terminal hydrolase (UCH)-L1 in 8 distinct LSD. UCH-L1 is essential for proper functioning of the proteasomal degradation pathway and its downregulation correlated with an impairment of this pathway which could contribute to the increased cell death observed in LSD [29]. Second, Settembre et al. proposed a model which identifies a block of autophagy as a crucial component in the pathogenesis of all LSDs. As mentioned earlier autophagy is a catabolic process the cell utilizes to regulate its own composition, by targeting the vacuolated cytoplasmic substrates to the lysosomes. According to this model, lysosomal accumulation of undegraded substrates results in defective fusion between autophagosomes and lysosomes and blocks the autophagic pathway at this step. As a consequence of this block, toxic proteins and dysfuctional mitochondria accumulate, ultimately leading to apoptosis, either directly or through the induction of chronic inflammation and cytokine release [30].

### 1.2.2 Lysosomal Storage Diseases overview

There are currently at least 50 genetically distinct LSDs caused by the absence of one or more lysosomal enzymes [31]. Individually, these disorders are considered rare, and each one typically affects, fewer that 10,000 people worldwide, although high prevalence values have been reported in some populations. Taken together however, these disorders affect 1 in 5000 newborn children [32, 33]. Inheritance for LSDs is usually autosomal recessive or X-linked such as in case for MPS II and Fabry disease.

Currently almost all the genes responsible for LSDs have been cloned. The genetic information gained is being used primarily for establishing disease prognoses based on genotype-phenotype correlation and in genetic counselling for the affected families. In general, LSDs exhibit high allelic variation (different mutations on the same gene) and significant clinical heterogeneity. Clinical features vary between different disorders but the most common characteristics include hepatosplenomegaly (enlargement of spleen and liver), skeletal-muscular abnormalities, and neurodegeneration.

LSDs manifest in several clinical forms with varying onsets (infantile, juvenile or adult) and display a correlated severity (severe, moderate, and attenuated). The progression and severity of the disease can be correlated with the residual enzyme activity, which is itself dependent on the underlying molecular defect(s). Often, very low amounts of residual enzyme activity can be adequate to carry out normal substrate clearance. Therefore, heterozygotes, whose enzyme activity is about 50% of normal, generally are clinically unaffected [34] and

patients develop symptoms when their residual enzyme activity falls below a critical threshold of 15 to 20%. There have even been incidences of asymptomatic individuals with beta-hexosaminidase A enzyme activity as low as 10% of normal, whereas patients with 5% activity exhibit the chronic adult onset form [35]. Therefore perturbations such as frame-shifts and nonsense mutations which leave no residual enzyme activity cause more severe, early onset illness; while missense mutations not affecting the enzyme active site and resulting in partially active enzymes correlate with the delayed onset or milder forms of the disorder.

LSDs are grouped into several major classes based on both biochemical and clinical manifestation of the disease (Table 1). The largest group of LSDs, sphigolipidoses, is associated with deficiencies of glycosidases involved in the catabolism of sugar chains of glycolipids. This group includes 2 of the most frequent LSD; Gaucher disease and Tay-Sachs disease, which disproportionately affect Ashkenazi Jews, with a carrier frequency of 4.6% and 3% respectively [36, 37]. Prior to the introduction of genetic testing, this represented a significant public health risk to people within this community. Tay-Sachs and the clinically indistinguishable Sandhoff disease are caused by defects in the degradation of $G_{M2}$-gangliosides, secondary to mutations in the *HEXA* and *HEXB* genes, respectively, which encode the subunits of beta-hexosaminidase A. Interestingly, the AB variant of Sandhoff disease is caused by a defect in the $G_{M2}$ activator protein (previously described), a cofactor for beta-hexosaminidase A. Gaucher disease is an autosomal recessive disorder caused by a deficiency of

| Disease | Enzyme Defect | Stored Substance | Soluble or membrane protein |
|---|---|---|---|
| **1. Sphingolipidoses** | | | |
| **A. Gangliosidoses** | | | |
| $G_{M1}$-gangliosidosis | β-galactosidase | $G_{M1}$ gangliosides | soluble |
| Tay-Sachs | β-hexosaminidase A | $G_{M2}$ gangliosides | soluble |
| Sandhoff | β-hexosaminidase A&B | $G_{M2}$ gangliosides | soluble |
| AB variant | $G_{M2}$ ganglioside activator | $G_{M2}$ gangliosides | soluble |
| **Visceral storage disorders** | | | |
| Fabry | α-galactosidase A | globotriasylceramide | soluble |
| Gaucher | glucocerebrosidase | glucocerebroside | soluble |
| Niemann-Pick, SMPD1-associated | acid sphingomyelinase | sphingomyelin | soluble |
| Niemann-Pick C | NPC1 & NPC2 | cholesterol/glycolipids | membrane |
| **Leukodystrophies** | | | |
| Krabbe | galactosylceramidase | galactocerebroside | soluble |
| Metachromatic leucodsytrophy | arylsulfatase A | sulfatides | soluble |
| **2. Glycoprotein Storage Disorders** | | | |
| Galactosialidosis | cathepsin A | sialyloligosaccharides | soluble |
| Fucosidosis | α-fucosidase | fucose | soluble |
| α/β-Mannosidosis | α/β-mannosidase | mannose/oligosaccharides | soluble |
| Schindler | α-N-acetygalactosaminidase | oligosaccharides | soluble |
| **3. Mucopolysaccharidoses** | | | |
| I-Hurler | α-L-iduronidase | heparan/dermatan-sulfate | soluble |

| Disease | Enzyme Defect | Stored Substance | Soluble or membrane protein |
|---|---|---|---|
| II-Hunter | iduronate-sulfatase | heparan/dermatan-sulfate | soluble |
| IIIA-Sanfilippo A | heparan N-sulfamidase | heparan sulfate | soluble |
| IIIB-Sanfilippo B | α-N-acetylglucosaminidase | heparan sulfate | soluble |
| IIIC-Sanfilippo C | heparan-α-glucosaminide N-acetyltransferase | heparan sulfate | membrane |
| IIID-Sanfilippo D | N-acetylglucosamine 6-sulfatase | heparan sulfate | soluble |
| IVA-Morquio A | galactose-6-sulfate sulfatase | keratan/chondroitin 6-sulfate | soluble |
| IVA-Morquio B | β-galactosidase | keratin sulfate | soluble |
| VI-Maroteaux-Lamy | N-acetylgalactosamine-4-sulfatase | dermatan sulfate | soluble |
| VII-Sly | β-glucuronidase | heparan/dermatan/ Chondroitin 6-sulfate | soluble |
| IX-Natowicz | hyaluronidase | hyaluronic acid | soluble |

**4. Mucolipidoses (ML)**

| | | | |
|---|---|---|---|
| MLI-Sialidosis | α-neuraminidase 1 | sialyloligosaccharides | soluble |

| Disease | Enzyme Defect | Stored Substance | Soluble or membrane protein |
|---|---|---|---|
| MLII-I-cell disease | N-acetyl glucosamine-1-phosphotransferase | carbohydrates/lipids | soluble |
| MLIII-Pseudo-Hurler polydystrophy | N-acetyl glucosamine-1-phosphotransferase | carbohydrates/lipids | soluble |
| MLIV | mucolipin I | lipids | membrane |
| **5. Other disorders** | | | |
| Pompe | α-glycosidase | glycogen | soluble |
| Wolman disease | acid lipase | cholesterol esters | soluble |
| Batten disease | battenin | lipofuscines | membrane |
| Cystinosis | cystinosin | cystine | membrane |
| Salla disease | sialin | sialic acid | membrane |
| Danon disease | LAMP2 | N/A | membrane |
| Cobalamin F-type disease | LMBD1 | cobalamin | membrane |

beta-glucocerebrosidase, resulting in the lysosomal accumulation of glycosylceramide.

More than 30% of all lysosomal diseases are MPSs, disorders which affect the enzymes needed for the stepwise degradation of glycosaminoglycans (GAG, mucopolysaccharides). These are repeating carbohydrates generally found attached to proteins (proteoglycans) and include: heparan sulfate, dermatan sulfate, chondroitin 6-sulfate and hyaluronic acid. Because proteoglycans are primarily structural molecules produced by almost all cell types and are found mainly on the surface of cells and in the extracellular matrix, most MPSs present with hepatosplenomegaly, central nervous system (CNS) pathologies as well as skeletal and muscular abnormalities. Several of the enzymes deficient in MPS result in blocks of multiple GAG due to overlapping degradation pathways as indicated in Table 1.

### 1.2.3 LSD therapies

Forty years ago, experiments by Dr. E. Neufeld revealed that the metabolic defect in cultured fibroblasts from mucopolysaccharidosis patients can be compensated for by addition of enzymes which are secreted by cells not having the same defect [38]. This idea of complementation is the underlying principle in 2 of the 4 current therapeutic options for patients with LSD, namely bone marrow transplantation (BMT), and enzyme replacement therapy (ERT), the other 2 being enzyme enhancement/chaperone therapy (EET) and substrate reduction therapy (SRT). As mentioned previously in most of cases only a small

percentage of normal enzymatic activity is required to reduce storage to normal levels and to alleviate the symptoms in patients.

#### 1.2.3.1 Bone marrow transplantation

Bone marrow transplantation has historically been the treatment method of choice for patients with MPS, specifically for Sly syndrome, Maroteaux-Lamy syndrome and Hurler syndrome. In this case the hematopoietic cells from the healthy donor become the source of enzyme for the other tissues and organs, including the brain through infiltration of microglia cells. In general BMT results in a long term survival of the patients due to sustainable enzyme production. BMT however, comes with large risks as the procedure has high morbidity and mortality rates (currently up to 10% in the best clinics) mostly due to infections, rejection of the transplant and graft versus host disease (GVHD). The first BMT used for treatment of MPS was performed by Hobbs in 1981 in a patient with Hurler syndrome (MPS I). The procedure led to the reversal of many of the clinical features [39]. After the initial success observed in Hobbs' patient, it was thought that all MPSs could be treated with BMT therapy; however, despite several attempts further success was only realized in patients with Maroteau-Lamy (MPS VI) and Sly (MPS VII) syndromes.

Many of the immediate benefits realized for MPS patients treated with BMT occur in the organ systems and include restoration of the normal mass of liver and spleen [40]. CNS improvement is usually delayed and deterioration may only be stabilized 6-12 months post BMT due to the slow turnover of microglia and their replacement by donor derived cells [41]. Despite these benefits, little

improvement is seen in bone disease after BMT, most likely due to poor enzymatic penetration into the chondrocytes and a failure to correct or replace osteocytes [42].

BMT therapy has continued to be used as a therapeutic option in patients with Hurler syndrome (MPS I). The treatment has been shown to lead to rapid reduction in GAG substrate in liver, tonsils, conjunctiva, cerebrospinal fluid and urine. This has resulted in reduced obstruction in the airway, reduction of hepatosplenogmagaly and corneal clouding, as well as elimination of heart failure, and stabilization of myocardial muscle function [40]. As mentioned above, some features of the disease are not alleviated such as the skeletal deformities, and developmental disorders, especially in the first year post BMT.

Patients with Maroteaux-Lamy syndrome (MPS VI) have shown significant improvements including enzymatic and biochemical correction, resolution of hepatosplenomegaly, stabilization of cardiopulmonary function and improvement in visual acuity and joint mobility [40]. Treatment of Sly syndrome (MPS VII) with BMT has led to the amelioration of symptoms in only a few patients.

In summary, BMT has been shown to provide increases in patient quality and length of life. BMT therapy will thus continue to be a viable treatment for some MPS patients. At the same time scientists are exploring interventions that have less risk and are less invasive.

**1.2.3.2 Enzyme Replacement Therapy**

The findings of Dr. Neufeld allowed the hypothesis that exogenous delivery of the missing enzyme *in vivo* could lead to clearance of the stored substrate and alleviation of the disease symptoms. This approach was further advanced by discovery that mannose 6-phosphate receptors on the cell surface are able to bind to and deliver exogenously supplied enzyme to the lysosomes of cultured cells [43]. These results paved the way for ERT intervention in patients with LSDs.

Currently an intravenous (IV) infusion of purified recombinant enzyme produced in cultured Chinese hamster ovary cells or human fibroblasts is the standard therapy for patients with the non-neuronopathic form of Gaucher disease, Fabry disease, MPS I, MPS II, and Pompe disease [44, 45]. ERT for patients with Gaucher type 1 has been approved by the FDA since 1991. There are currently over 4300 patients receiving treatment and the effectiveness of the treament is well confirmed [44]. Patients displayed decreases in spleen and liver size, increase in haemoglobin and blood platelet counts as well as some improvement of bone structure. It is important however that ERT must begin early in order to prevent progressive bone disease [46].

In Fabry disease the major site of pathology is the vascular endothelium which is readily accessed by exogenous alpha-galactosidase administered via bi-monthly IV infusion. This treatment has been shown to stabilize deteriorating renal function and led to improved cardiac functioning in hundreds of patients.

ERT for some other MPS are now undergoing clinical trials including that for MPS VI and MPS IVA. Morquio A (MPS IVA) syndrome, relatively frequent in Quebec, is caused by the deficiency of galactosamine-6-sulfatase and characterized by accumulation of both keratan sulfate and chondroitin 6-sulfate in the lysosomes of ligaments, connective tissue, bone and cartilage. This results in systemic skeletal displasia, mild hepatosplenomegaly and heart valvular disease in addition to other skeletal deformities. Pre-clinical data from a mouse model of Morquio A showed reduction of storage in affected organs and tissues, providing a proof of concept for an ERT option for patients this syndrome [47].

Despite its many therapeutic benefits ERT is not suitable for treatment of LSD with CNS involvement since the recombinant enzyme are is not able to cross the blood brain barrier and therefore cannot alleviate the CNS pathology. Additionally the treatment comes with a very high cost which may limit its access to patients.

#### 1.2.3.3 Enzyme enhancement/chaperone therapy

Chaperone therapy is an emerging method for therapy with the potential to treat many genetic diseases which are a result of misfolded or improperly assembled enzymes. Many lysosomal disorders are therefore candidates for such treatment. As previously mentioned many of these disorders have the additional advantage of only requiring very minimal amounts of residual enzyme activity to prevent or even reduce substrate accumulation.

Lysosomal enzymes fold into their native conformation in the ER with the assistance of various molecular chaperones, such as BiP, heat shock proteins and calnexin/calreticulin [48, 49]. Since folding is a thermodynamically controlled process some mutant proteins containing altered amino acid sequences tend to misfold into an alternative conformation. These mutants are retained in the ER and degraded by the proteosome pathway [50].

It has been previously shown that competitive inhibitors that mimic substrate binding in the active site can work as active-site-specific chaperones (ASSC) of the mutant enzymes. This is accomplished by stabilizing the proper position of active site residues and shifting the equilibrium toward the correctly folded state of the enzyme [51]. As a result, the correctly folded mutant enzyme passes the quality-control system of the ER and undergoes further maturation and normal transport to the lysosome. Once a mutant enzyme-ASSC complex reaches the lysosome, the ASSC is replaced by the accumulated substrate which allows the enzyme to function thereby leading to an increase in residual enzyme activity in the cell and partial clearance of the stored substrate [51].

Previous studies identified effective ASSCs for several lysosomal enzymes, and showed that some of them restored the impaired lysosomal catabolism in cultured patient cells or had a therapeutic effect in mouse models of GM1-gangliosidosis, GM2-gangliosidosis, Gaucher and Fabry diseases [44]. In all above cases ASSCs were competitive inhibitors of the deficient enzymes, suggesting that this strategy can be extended for other lysosomal enzymes and

disorders. These experiments provide a further proof of principle that chaperone therapy can be a viable treatment option for patients with folding disorders.

The utilization of pharmalogical chaperones in the treatment of LSDs is still in its infancy. In particular, the pharmacodynamics, bioavailability and the ability of the drug to cross the blood brain barrier and treat CNS pathologies will have to be evaluated on a case by case basis before any clinical trials can even be proposed.

#### 1.2.3.4 Substrate Reduction Therapy

SRT aims to lower the availability of the defective enzyme substrate to prevent its storage. This can be accomplished by either blocking an earlier step in the catabolism of the substrate, or by rerouting the affected metabolic pathway via altered signalling or gene silencing [52]. If such a drug can then be designed to cross the blood brain barrier there is a potential to target every cell in the body and thus provide an opportunity to ameliorate some of the CNS pathologies present in many LSDs.

SRT based on N-butyldeoxynojirimycin, an inhibitor of ceramide glucosyltransferase which can block the synthesis of glycoshphingolipids has proven to be successful in glycosphingolipid storage disorders by reducing pools of GM1 and GM2 gangliosides in Sandhoffs, Tay Sachs and Gaucher patients. Currently it has been FDA approved for the treatment of non-neuronopathic type 1 Gaucher patients who are unwilling or unable to receive ERT [53].

Additionally, SRT clinical trials for patients with late-onset Tay Sachs, Sandhoff disease, and Niemann-Pick disease type C are ongoing.

Another study utilized siRNAs targeted towards *XYLT1, XYLT2, GALTI* and *GALTII* genes involved in the initial steps of GAG synthesis and showed that their inhibition was correlated with decreased synthesis of GAGs in MPS IIIa fibroblasts. More work is needed on animal models however; these initial results are very encouraging. Behavioural improvements were also seen in MPS IIIa mice treated with rhodamine B, a non-specific inhibitor of GAG synthesis, providing hope that the neurological pathologies can be improved in patients [54].

Interestingly, Sandhoff mice treated with combination therapy of SRT and non-steroidal anti-inflammatory drugs showed synergistic effects of the two treatments and had a drastic increase in life expectancy. These results confirm that targeting more than one aspect of the disease pathology can result in a better therapeutic outcome [55].

Therapeutic interventions for patients of LSDs have come a long way since inception, and can provide some patients with alleviation of symptoms and longer healthier lives. There are however certain obstacles that will need to be overcome in order to provide the best possible treatment outcomes for patient. One of the most pressing issues that will need to be addressed is the ability of drugs to cross the blood brain barrier thus allowing for the targeting of CNS pathologies. While it may be inconceivable currently to cure these diseases, the

advent of these and other newer therapies can help insure that patients of LSDs can be treated effectively.

### 1.3.1 Mucopolysaccharidosis III type C

Mucopolysacharidosis III (MPS III) or Sanfilippo syndrome is characterized by the storage of heparan sulfate in the lysosomes of all tissues and organs and its excretion in urine. The four subtypes of the disease are defined by their respective enzyme deficiencies: MPS III type A (heparan N-sulfatase); MPS III type B (α-N-acetylglucosaminidase); MPS III type C (N-acetyltransferase); and MPS III type D (N-acetylglucosamine 6-sulfatase) (Table 1). A clinical diagnosis is initially made by a visual evaluation then confirmed by analysis of GAG content in urine. Subtype determination is done via activity assay in fibroblasts obtained by a skin biopsy as all subtypes are clinically indistinguishable. All subtypes have onset in infancy or early childhood and show progressive and severe neurological deterioration causing hyperactivity, sleep disorders and loss of speech accompanied by behavioural abnormalities, neuropsychiatric problems, mental retardation, hearing loss, and visceral manifestations, such as mild hepatosplenomegaly, mild dwarfism with joint stiffness and biconvex dorsolumbar vertebral bodies, mild coarse faces, and hypertrichosis [56]. Many patients become demented and die before adulthood but some survive to the fourth decade with progressive dementia and retinitis pigmentosa [57, 58].

MPS IIIC was described by Kresse et al. who found that three patients with the phenotype of Sanfilippo syndrome had deficiency of an enzyme that transfers an acetyl group from cytoplasmically derived acetyl-CoA to terminal α-glucosamine residues of heparan sulfate within the lysosomes [59]. Klein et al.

reported a similar deficiency in 11 patients with the Sanfilippo syndrome, therefore suggesting that the disease is a relatively frequent subtype [60, 61]. Later the birth prevalence of the MPS IIIC in Australia, Portugal and the Netherlands was estimated at 0.07, 0.12 and 0.21 per 100,000, respectively [33, 62, 63].

### 1.3.2 Heparan sulfate catabolism.

Heparan sulfate is found in proteoglycans associated with the cell membrane in nearly all cells, but is most abundant in connective tissues [64]. It consists of repeating disaccharides comprised of units of sulfated L-iduronic or glucuronic acid linked to N-glucosamine. Heparan sulfate is degraded within the lysosomes by a collection of eight enzymes: four sulfatases, three exo-glycosydases and one N-acetyltransferase, that act to take apart the disaccharide pairs successively at the terminus of heparan sulfate chains, until the molecule is fully degraded, producing free sulfate and monosaccharides (Figure 1). The first enzyme, iduronate sulfatase releases sulfate from the terminal 2-sulfated iduronic acid of the heparan sulfate chain (Fig. 1, reaction 1). Next, α-iduronidase releases the L-iduronic acid (reaction 2). This is followed by heparan N-sulfatase which removes the sulfate linked to the amino group of glucosamine (reaction 3). The last step is the removal of glucosamine. However,

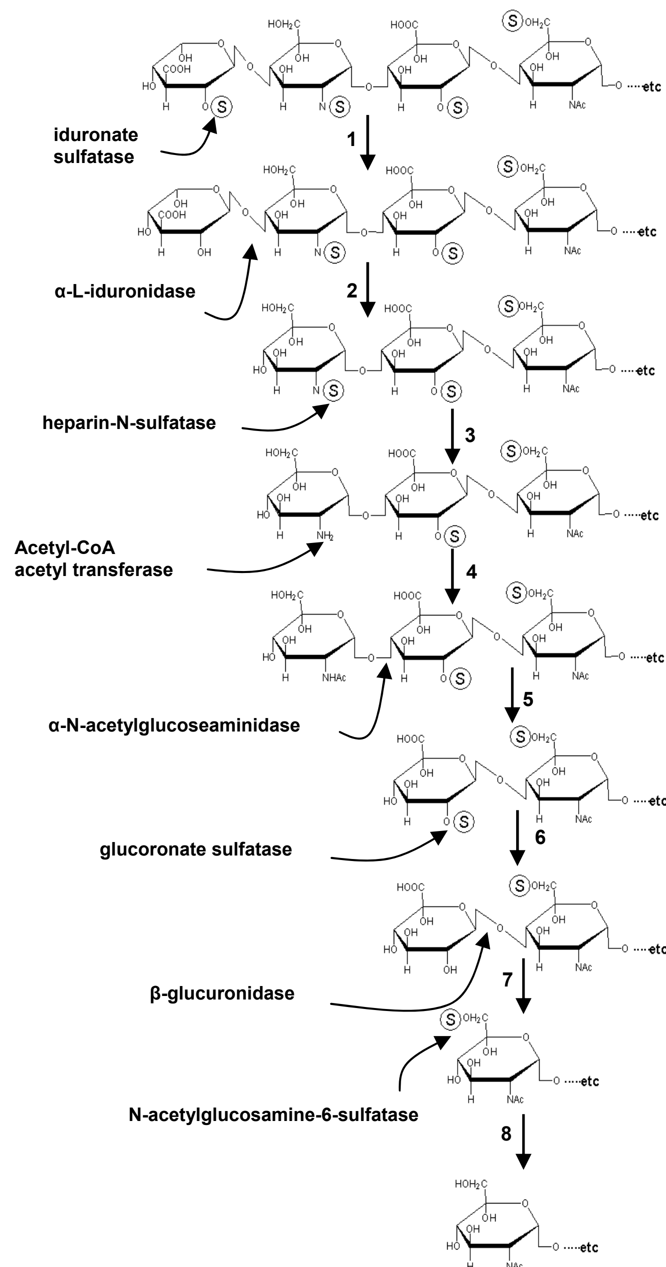



**Figure 1. Catabolic pathway for degradation of heparan sulfate**. (1) Iduronate sulfatase releases sulfate from the terminal 2-sulfated iduronic acid of the heparan sulfate chain; (2) α-iduronidase releases the L-iduronic acid; (3) heparan N-sulfatase removes the sulfate linked to the amino group of glucosamine; (4) glucosamine is acetylated by N-acetyltransferase to generate N-acetylglucosamine which is removed (5) by α-N-acetylgucosaminidase; (6) glucuronic sulfatase followed by ß-glucuronidase (7) cleave the sulfate and glucuronic acid; (8) N-acetylglucosamine 6-sulfatase cleaves sulfate esters from N-acetylglucosamine-sulfate. Genetic disorders are: MPS II (1); MPS I (2); MPS IIIA (3); MPS IIIC (4); MPS IIIB (5); MPS VII (7); MPS IIID (8).

it must be first acetylated by N-acetyltransferase to generate N-acetylglucosamine (reaction 4) which is then removed by α-N-acetylgucosaminidase (reaction 5). Glucuronic sulfatase (reaction 6) followed by ß-glucuronidase (reaction 7) cleave the sulfate and glucuronic acid, with subsequent glucosamine hydrolysis. Lastly, N-acetylglucosamine 6-sulfatase cleaves sulfated esters from N-acetylglucosamine-sulfate (reaction 8) reviewed by [65].

### 1.3.3 Evidence for occurrence of a transmembrane acetylation reaction.

The most intriguing of the enzymes of heparan sulfate catabolism is N-acetyltransferase. The enzyme is critical for generating the acetylated version of glucosamine since there is no enzyme that can act on the unacetylated molecule. The mechanism by which it is achieved has been the topic of considerable investigation, since the acetyl donor, acetyl-CoA, is not stable in the lysosomal environment [66].

Bame and Rome were first to report that AcCoA is not transported from the cytoplasm into the lysosome as a free molecule [67]. Instead, they proposed that the N-acetyltransferase catalyses the acetylation of heparan sulfate via a double displacement transmembrane reaction, with covalent linkage of the acetyl group to the enzyme as the intermediate [68, 69]. In this scheme AcCoA binds to the enzyme on the cytoplasmic side of the lysosomal membrane; the acetyl group is covalently transferred to a histidine residue on the enzyme followed by the release of CoA; then the acetylated enzyme undergoes a conformational change resulting in a transmembrane relocation of the bound acetyl group towards the

lysosomal interior. Once inside the lysosome, the acetylhistidine intermediate is protonated. The substrate glucosamine binds to a separate site on the enzyme and becomes acetylated by transfer from the acetylhistidine. The product is released and the enzyme returns to its original conformation (Figure 2). This mechanism was supported by genetic evidence which showed that lysates of 5 cell lines of MPS IIIC patients were able to catalyze acetylation of the lysosomal membrane and to carry out the acetyl-CoA/CoA exchange, but could not transfer the bound acetyl group to glucosamine [68].

### 1.3.4 The identification and cloning of the gene responsible for MPS IIIC.

To identify the gene responsible for MPS IIIC, a genome-wide association scan on MPS IIIC patients and unaffected family members was performed. This allowed Ausseil et al. to map the gene to an 8.3-cM interval of chromosome 8 [70]. The group then confirmed the chromosome 8 assignment by complementation of patient fibroblasts with human chromosome 8 [71]. In order to narrow down the size of the MPS IIIC locus, all available MPS IIIC families and individuals were genotyped for 22 microsatellite markers in the pericentromeric region of chromosome 8 [72] which further reduced the size of the loci.

Based on the previous studies that defined the molecular properties of the lysosomal N-acetyltransferase, they searched the candidate region for a gene

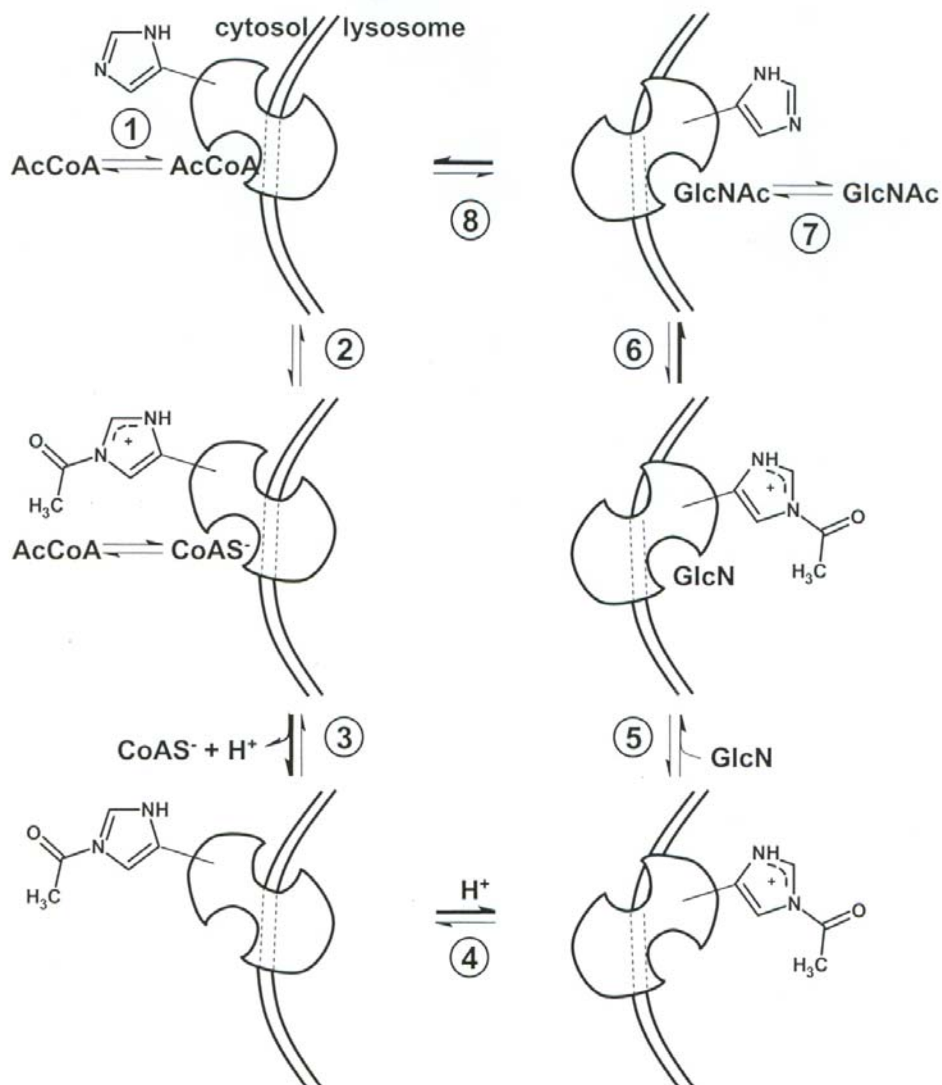



**Figure 2. Proposed catalytic mechanism for HGSNAT.** (1) AcCoA binds to the enzyme on the cytoplasmic side of the lysosomal membrane; (2) the acetyl group is covalently transferred to a histidine residue on the enzyme followed by the release of CoA (3); (4) the acetylated enzyme undergoes a conformational change resulting in a transmembrane relocation of the bound acetyl group towards the lysosomal interior and the acetylhistidine intermediate is protenated; (5) the substrate glucosamine binds to a separate site on the enzyme and (6) becomes acetylated by transfer from the acetylhistidine; the product is released (7) and the enzyme returns to its original conformation (8).

encoding a protein with multiple transmembrane domains and a molecular weight of approximately 100 kDa. This allowed for the exclusion of the majority of the genes in the region with an exception of the *TMEM76* gene. Furthermore, of all the genes present in the candidate interval only *TMEM76* showed a statistically significant reduction of the transcript level in the cells of two MPS IIIC patients in a custom oligonucleotide-based microarray assay. Both patients were later shown to have nonsense mutations presumably causing mRNA decay [72].

Northern blot analysis revealed that transcripts were ubiquitously expressed in various tissues. The highest expression was detected in leukocytes, heart, lung, placenta, and liver, whereas in the thymus, colon, and brain, the gene was expressed at a much lower level, which is consistent with the expression patterns of lysosomal proteins. The cloning of *TMEM76* as the gene responsible for MPS IIIC was performed independently by another lab through sequencing of genes identified by proteomic-based analysis of lysosomal proteins in mouse [73].

**1.3.5 Biochemical properties of HGSNAT.**

Early work on HGSNAT determined it to be an integral membrane protein, requiring high concentrations of detergent to achieve proper solubilization and that maximum enzyme activity only occurred in the presence of phospholipids and glycolipids [74]. In addition to acetylation of terminal glucosamine residues of heparan sulfate, HGSNAT was also shown to acetylate several heparin derived di- and trisaccharides containing α-glucosamine residues at their nonreducing positions, and the monosaccharide glucosamine [75-77]

Bame and Rome determined that the active site residue was a histidine by treating the lysosomal membrane with various amino acid modifying agents and then assaying for enzyme activity [69]. As previously mentioned Bame and Rome proposed that HGSNAT functions via a ping-pong (double displacement) mechanism involving two half-reactions whereby first HGSNAT is acetylated on its cytoplasmic face and then the acetyl group is transferred to glucosamine [67]. This mechanism was supported by genetic evidence which showed that lysates of 5 cell lines of MPS IIIC patients were able to catalyze acetylation of the lysosomal membrane and to carry out the acetyl-CoA/CoA exchange, but could not transfer the bound acetyl group to glucosamine, whereas a sixth cell line was devoid of both activities [68]. In contrast Meikle proposed an alternative mechanism whereby the transfer of the acetyl group operates by a random-order ternary-complex mechanism, however they were unable to demonstrate specific acetylation of the lysosomal membranes [78].

Studies by Ausseil et al. described partially purified HGSNAT as having a molecular mass of 240 kDa and an isoelectric focusing point (pI) of 7.4. Additionally they determined that the specific activity directly correlated with protein concentration, suggesting that HGSNAT may be part of a multiprotein complex that acts to sequester heparan sulfate within the lysosomes. Ion-exchange chromatography identified two peaks having HGSNAT activity, suggesting that the enzyme is composed of several subunits each with its own activity. Radiolabelling of the active site determined that the catalytic subunit is represented by a 120 kDa protein, which is supported by the absence of this protein from cells of MPS IIIC patients [79].

Identification of the *TMEM76* gene revealed that a 1,908-bp coding sequence transcribed from 18 exons codes for the HGSNAT enzyme. The enzyme contains two possible ATG start codons and the deduced amino acid sequence predicts 11 transmembrane domains and five potential N-glycosylation sites oriented toward the lysosomal lumen consistent with the molecular properties of lysosomal N-acetyltransferase [72]. The C-terminus of the enzyme exposed to the cytoplasm contains a highly conserved TALWVLIAYILYRKK domain which was predicted to carry out lysosomal targeting. Although no evidence has been put forth yet to confirm this localization signal; there is strong evidence that Leu-Leu sequence motifs involved in the interaction with the adaptor proteins may be responsible for the lysosomal targeting of membrane proteins as previously mentioned [2]. HGSNAT does not share homology with any other protein of

known function; however its sequence is highly conserved among all species tested [72].

### 1.3.6 Mutations in HGSNAT

Initially, Hrebicek *et al*. identified 4 nonsense mutations, 3 frameshift mutations due to deletions or a duplication, 6 splice-site mutations, and 14 missense mutations among 30 probands with MPS IIIC (Table 2). Whereas Fan *et al*. identified 1 splice-site mutation and 1 duplication.

Several studies have since been published examining the mutation spectrum in Dutch, Italian and Portuguese patient populations. Fedele *et al*. described nine alleles in patients of Italian origin including 3 splice-site mutations, 3 frameshift deletions resulting in premature stop codons, 1 nonsense mutation, and 2 missense mutations [80]. Ruijter et al. described 14 novel mutations: two splice site mutations, one frame shift mutation due to an insertion, three nonsense mutations and eight missense mutations. Interestingly, the two missense mutations p.R344C and p.S518F were very frequent among probands of Dutch origin and represented 22.0% and 29.3% respectively of the mutant alleles described. Additionally the authors described the clinical course of the disease, and identified delayed speech, motor development and behavioural problems in almost all the patients [81]. More recently a small study of MPS IIIC in the Portuguese patient population identified one splice-site mutation and one duplication mutation. Interestingly, the authors found the duplication c.525dupT

in 5 out of 6 patient alleles tested suggesting a possible founder effect [82]. All reported HGSNAT mutations are listed in Article 1 Table 1.

### 1.3.7 HYPOTHESES and OBJECTIVES

Our major hypothesis was that misfolding of mutant HGSNAT underlies the mechanism of the disease in MPS IIIC patients having missense mutations. Misfolding of the enzyme impairs its proper processing and targeting to the lysosome. It is retained in the ER and therefore cannot contribute any residual activity to clear stored heparan sulfate in lysosomes, causing the severe disease progression observed in patients. Consequently, if MPS IIIC is a folding disorder, then ASSCs can bind to and stabilize the active site of the misfolded protein thereby increasing HGSNAT activity in MPS IIIC patients' cells and suggesting a possible treatment option.

We therefore propose to study the expression and biogenesis of recombinant HGSNAT containing missense mutations in cellular systems. For each missense mutants, we will determine its activity, localization, and processing (including the extent of glycosylation) versus a wild-type control. Additionally, we will investigate several potential ASSCs of HGSNAT that we will test in patient fibroblasts to determine if they can rescue the enzyme activity.

# CHAPTER 2

# Mutations in HGSNAT

# Foreword

MPS IIIC exhibits high clinical heterogeneity resulting in a large number of reported mutations. To better understand the effects these mutations have on the enzyme and how they may correlate with patient phenotype we compiled all the reported mutations in a "Mutation Update". This update includes a comprehensive table of all MPS IIIC mutations, patients and the predicted changes or effect on the enzyme. The article also reports 10 novel mutations that we identified in specimens of 14 previously unreported MPS IIIC patients referred to us by our colleagues from around Quebec and the world. Additionally, we characterized 4 previously reported mutants as polymorphisms which presumably have no effect on the enzyme's function and are therefore not clinically relevant.

Mutant identification was done by sequencing patient PCR products and analyzing the results against control alleles, confirmation was carried out in parental DNA when possible.

Analyses of sequencing data, compiling parts of the mutation table as well as writing and editing the manuscript were conducted by Matthew Feldhammer.

*Article published in Human Mutation Journal following peer review

# SANFILIPPO SYNDROME TYPE C: MUTATION SPECTRUM IN THE HEPARAN SULFATE ACETYL-CoA: A-GLUCOSAMINIDE N-ACETYLTRANSFERASE (*HGSNAT*) GENE

Matthew Feldhammer[1,2†], Stéphanie Durand[1†], Lenka Mrázová[4], Renée-Myriam Boucher[6], Rachel Laframboise[6], Robert Steinfeld[7], James E. Wraith[8], Helen Michelakakis[9], Otto P. van Diggelen[10], Martin Hřebíček[4], Stanislav Kmoch[4,5], and Alexey V. Pshezhetsky[1,2,3,11]

[1]Department of Medical Genetics, University Hospital Center Sainte-Justine and Departments of [2]Biochemistry and [3]Pediatrics, University of Montreal, Montreal, Canada;
[4]Institute for Inherited Metabolic Disorders and [5]Center for Applied Genomics, Charles University in Prague, 1st Faculty of Medicine, Prague, Czech Republic;
[6]Department of Pediatric Neurology and Department of Pediatrics, University Hospital Centre, Laval University, Quebec, Canada;
[7]Department of Pediatrics, University of Goettingen, Germany;
[8]Willink Biochemical Genetics Unit, Royal Manchester Children's Hospital, Manchester, United Kingdom;
[9]Department of Enzymology and Cellular Function, Institute of Child Health, Agia Sophia Children's Hospital, Athens, Greece;
[10]Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, Netherlands;
[11]Department of Anatomy and Cell Biology, Faculty of Medicine, McGill University, Montreal, Canada.
Correspondence should be addressed to:

Alexey V. Pshezhetsky, Ph.D.

Service de génétique médicale, CHU Sainte-Justine

3175 Côte Ste-Catherine, Montréal (Qc) H3T 1C5 - Canada

Tel. (514) 345-4931/2736

Fax. (514) 345-4801

E-Mail: Alexei Pshezhetsky

[†] M.F. and S.D. equally contributed as first authors.

**ABSTRACT**

Mucopolysaccharidosis type IIIC or Sanfilippo syndrome type C (MPS IIIC, MIM #252930) is a rare autosomal recessive disorder caused by the deficiency of the lysosomal membrane enzyme, heparan sulfate acetyl-CoA: α-glucosaminide N-acetyltransferase (HGSNAT, EC 2.3.1.78), which catalyses transmembrane acetylation of the terminal glucosamine residues of heparan sulfate prior to their hydrolysis by α-N-acetylglucosaminidase. Lysosomal storage of undegraded heparan sulfate in the cells of affected patients leads to neuronal death causing neurodegeneration and severely impaired development accompanied by mild visceral and skeletal abnormalities, including mild dwarfism, coarse facies, and joint stiffness. To date, 50 HGSNAT mutations have been identified in MPS IIIC patients: 40 were previously published and 10 novel mutations are reported here. The mutations span the entire structure of the gene and include 13 splice-site mutations, 11 insertions and deletions, 8 nonsense, and 18 missense mutations (http://chromium.liacs.nl/LOVD2/home.php?select_db=HGSNAT). In addition 4 polymorphisms result in amino acid changes that do not affect activity of the enzyme. In this paper we discuss the spectrum of MPS IIIC mutations, their clinical presentation and distribution within the patient population, and speculate how the mutations may affect the structure and function of HGSNAT.

Key words: HGSNAT**,** N-acetyltransferase, lysosome, heparan sulfate, mucopolysaccharidosis, mutation analysis.

**BACKGROUND**

More than 50 metabolic storage diseases are currently known, caused by genetic blocks in the breakdown of macromolecules in the lysosomes. These diseases comprise ~14% of all inherited diseases of metabolism and, although individually rare, together affect one of every 7700 newborn children [Meikle et al., 1999]. More than 30% of all lysosomal diseases are mucopolysaccharidoses, disorders which affect the enzymes needed for the stepwise degradation of glycosaminoglycans (mucopolysaccharides), repeating carbohydrates generally found attached to proteoglycans. One of them, heparan sulfate, is found in proteoglycans associated with the cell membrane in nearly all cells, but is most abundant in connective tissues [Gallagher 2006]. It consists of a repeating disaccharide comprised of units of sulfated L-iduronic or glucuronic acid linked to N-glucosamine. Stepwise degradation of heparan sulfate occurs within the lysosomes by the concerted action of at least eight enzymes: four sulfatases (iduronate sulfatase, heparan N-sulfatase, glucuronate sulfatase, and N-acetylglucosamine 6-sulfatase) three exo-glycosidases (α-iduronidase; ß-glucuronidase and α-N-acetylglucosaminidase) and heparan sulfate acetyl-CoA: α-glucosaminide N-acetyltransferase (HGSNAT) that generates N-acetylglucosamine which is then removed by α-N-acetylglucosaminidase.

The most biochemically intriguing of the enzymes of heparan sulfate catabolism is HGSNAT, the only known lysosomal enzyme that is not a hydrolase. This enzyme is critical to generate the acetylated version of

glucosamine since none of the lysosomal glycosidases can act on the unacetylated molecule. The mechanism by which this is achieved has been the topic of considerable investigation, since the acetyl donor, acetyl-CoA (AcCoA), is not stable in the acidic lysosomal environment [Rome et al., 1983]. More than 20 years ago, Bame and Rome reported that AcCoA is not transported from the cytoplasm into the lysosome as a free molecule [Bame and Rome, 1985]. Instead, they proposed that the N-acetyltransferase catalyses the acetylation of heparan sulfate via a ping-pong (double displacement) transmembrane reaction, with covalent linkage of the acetyl group to the enzyme as the intermediate [Bame and Rome, 1985, 1986a, 1986b]. In this scheme, AcCoA binds to HGSNAT on the cytoplasmic side of the lysosomal membrane; the acetyl group is covalently transferred to its active site and the acetylated enzyme undergoes a conformational change resulting in a transmembrane relocation of the bound acetyl group towards the lysosomal interior. Once inside the lysosome, the substrate glucosamine binds to the enzyme and becomes acetylated. The product is then released and the enzyme returns to its original conformation.

Seven lysosomal storage diseases are caused by genetic deficiencies of the enzymes involved in heparan sulfate catabolism. Three of them: MPS II (Hunter syndrome, iduronate sulfatase deficiency), MPS I (Hurler-Scheie syndrome, α-iduronidase deficiency) and MPS VII (Sly syndrome, glucuronidase deficiency) also have blocks in dermatan sulfate degradation and, therefore, result in accumulation of both compounds. Defects in the remaining enzymes, α-N-acetylglucosaminidase, heparan N-sulfatase, N-acetyltransferase and N-

acetylglucosamine 6-sulfatase, uniquely cause blocks in the catabolism of heparan sulfate. They are classified as subtypes of a single disorder, MPS III or Sanfilippo syndrome [Sanfilippo, 1963]. No disorder has been found associated with glucuronate sulfatase deficiency.

Biochemically, MPS III is characterized by abnormal storage of heparan sulfate in the lysosomes of all tissues and organs and its excretion in urine. The four subtypes of the disease are defined by the enzyme deficiency: MPS III type A (heparan N-sulfatase, MIM #252900), MPS III type B (α-N-acetylglucosaminidase, MIM #252920), MPS III type C (HGSNAT, MIM #252930), and MPS III type D (N-acetylglucosamine 6-sulfatase, MIM #252940). All subtypes have similar clinical phenotypes with onset in infancy or early childhood: progressive and severe neurological deterioration causing hyperactivity, sleep disorders, loss of speech accompanied by behavioral abnormalities, neuropsychiatric problems, mental retardation, hearing loss, and visceral manifestations, such as mild hepatomegaly, mild dysostosis multiplex , mild coarse facies, and hypertrichosis [Bartsocas et al., 1979; Valstar et al. 2008]. Some patients experience severe mental retardation and die before adulthood but others survive to the fourth decade with progressive dementia and retinitis pigmentosa [Bartsocas et al., 1979]. Interestingly, in the very rare MPS IIIC patients with the onset of symptoms in adulthood the disease's progression was similar in severity and time course to the forms with onset in childhood [Berger-Plantinga et al. 2004].

MPS IIIC was first described by Kresse et al. [1976] who found that three patients with the phenotype of Sanfilippo syndrome had a deficiency of an enzyme that transfers an acetyl group from cytoplasmically derived acetyl-CoA to terminal N-glucosamine residues of heparan sulfate within the lysosomes. Klein et al. [1978, 1981] reported a similar deficiency in 11 patients with the Sanfilippo syndrome, therefore suggesting that the disease is a relatively frequent subtype. Later, the birth prevalence of MPS IIIC in Australia, Portugal and the Netherlands was estimated at 0.07, 0.12 and 0.21 per 100,000, respectively [Meikle et al., 1999; Pinto et al., 2004; Poorthuis et al., 1999].

While the genes responsible for the other MPS's have been known for a long time, the identification of *HGSNAT* has been hampered for almost three decades by the low tissue content and instability of this enzyme. The gene was cloned only recently by two independent laboratories [Fan et al., 2006; Hrebicek et al., 2006]. Based on our previous studies that defined the chromosomal localization of the MPS IIIC allele [Ausseil et al., 2004] and molecular properties of HGSNAT [Hopwood et al., 1983; Pohlmann et al., 1981; Ausseil et al., 2006], we searched the centromeric region of human chromosome 8 for a gene encoding a protein with multiple transmembrane domains and a molecular weight of ~100 kDa which allowed us to exclude the majority of the genes with an exception of the gene initially called *TMEM76* [Hrebicek et al., 2006]. Furthermore, of all the genes present in the candidate interval only *TMEM76* showed a statistically significant reduction of the transcript level in the cells of two MPS IIIC patients. Both patients were later shown to have nonsense mutations presumably causing

mRNA decay [Hrebicek et al., 2006]. Northern blot analysis identified two major transcripts of 4.5 and 2.1 kb ubiquitously expressed in various human tissues. The highest expression was detected in leukocytes, heart, lung, placenta, and liver, whereas in the thymus, colon, spleen and brain, the gene was expressed at a much lower level, which is consistent with the expression patterns of other lysosomal proteins. A 1908 bp cDNA is transcribed from 18 exons of the gene. The deduced amino acid sequence predicts 11 transmembrane domains and five potential N-glycosylation sites (Fig. 1) oriented towards the lysosomal lumen. The first 39 amino acids are predicted to comprise the signal peptide with length and composition resembling those of lysosomal proteins. According to predictions made by empirical computer algorithms, the C-terminus of the enzyme is exposed to the cytoplasm and contains conserved Tyr-X-X-Θ (where Θ is a hydrophobic bulk residue) and Leu-Leu sequence motifs involved in the interaction with the adaptor proteins responsible for the lysosomal targeting of membrane proteins [Bonifacino and Traub, 2003]. Significantly, Fan et al. [2006] who independently identified the *HGSNAT* gene by large-scale proteomic study of the lysosomal compartment have come to the same structure of the transmembrane domains.

## MUTATIONS AND THEIR BIOLOGICAL RELEVANCE

A total of 50 mutations affecting almost all 18 exons and many introns of the *HGSNAT* gene are identified in MPS IIIC patients from 83 families (Table 1 and Fig. 2). Among them are 18 missense mutations affecting 8 of the 11

transmembrane segments of the enzyme as well as its luminal and cytosolic domains (Fig. 1).

Below we summarize all previously identified mutations and present them with a unified numbering system based on the sequence of GenBank entries NM_152419.2 and NG_009552.1 as reported in Fan et al. [2006], Ruijter at al. [2008] and Fedele et al. [2007]. We also report 10 novel mutations, including those identified in patient families from Pakistan, the USA, the UK, Germany, Greece, Turkey and Belarus. These novel mutations were identified by sequencing the PCR products of patients genomic DNA as previously described [Hrebicek et al., 2006] and were absent in 210 control alleles. In addition, confirmation of each mutation was made based on sequencing of parental DNA and/or of cDNA from patients cultured skin fibroblasts when available. All the published variants have been entered in a locus specific database (LSDB) available online at http://chromium.liacs.nl/LOVD2/home.php?select_db=HGSNAT. This database also contains updated information about the genotype of the MPS IIIC patients previously reported by Hrebicek et al. [2006] as having no mutations in the *HGSNAT* gene or having only one heterozygous mutation. Two mutations in heterozygosis have now been identified for all of them.

**Splice site mutations**

Ten splice site mutations potentially resulting in abnormal splicing and rapid degradation of mRNA have been previously identified in MPS IIIC patients

(Table 1); six of them (c.118+1G>A, c.234+1G>A, c.493+1G>A, c.851+5G>A, c.1250+1G>A and c.1726+1G>A) were reported in our publication originally describing the *HGSNAT* gene [Hrebicek et al., 2006]. Fedele et al. [2007] reported the mutation c.852-1G>A present in homozygosis in patients from three families of Italian ancestry. The same report also described another splice mutation, c.1464+1G>A also found in homozygosis in a patient from the same region in Italy. Ruijter et al. [2007] reported one additional splice site mutation, c.743+1dupG in a Dutch patient and, most recently, Coutinho et al. [2008] reported the novel splice junction mutation c.372-2A>G which is predicted to result in the skipping of exon 4.

In this study we report 3 additional splice site mutations. c.744-2A>G, was found in homozygosis in a patient of Pakistani origin from the UK. The second, c.1250+2T>C which occurs just after exon 12 was found in a patient from the USA in heterozygosis with the missense mutation p.A489E. The third novel mutation, c.1542+4dupA, was found in heterozygosis in one patient from the UK along with the nonsense mutation p.Y558X. Parental DNA or patient's cDNA were not available for confirmation of these mutations or of their effect on the mRNA.

**Insertions and deletions**

Insertions and deletions can result in frame shifts and premature termination of translation causing rapid degradation of the resulting transcripts. Eight insertions and deletions affecting exons 5, 7, 11, 13, 17 and 18 of the

*HGSNAT* gene (Table 1) have been already reported. In particular: mutations c.1034_1049del, c.1336_1372dup and c.1750delG in patients from Czech Republic, Poland and Finland, respectively, were reported in our original publication, whereas Fedele et al. [2007] reported mutations c.682_740del, c.739delA, and c.1669_1674delinsACAT present in patients of Italian origin and Fan [2006] found c.1345dupG in a Canadian patient. Interestingly, the single-base deletion, c.739delA, predicted to result in a frameshift and the appearance of a premature termination codon (PTC) 87 bp downstream from the deletion was also found in homozygous fashion in a patient from North Africa and in combination with p.R506X in a Belgian patient. The mother of this Belgian patient who was found to be a carrier for this deletion is of Italian origin suggesting common ancestry for the two families. More recently Coutinho et al. [2008] reported the mutation c.525dupT in Portuguese patients. In two patients, the mutation was found in homozygosis and in the third, in heterozygosis with the splice site mutation c.372-2A>G.

We report here the large deletion c.1-1925_118+296del found in a homozygous fashion in two unrelated patients of Turkish origin. This large 2339 base pair deletion includes the entire sequence of exon 1 and should subsequently result in complete loss of protein expression. In both patients the presence of the mutation was confirmed based on sequencing the deletion endpoints after PCR amplification of genomic DNA (data not shown). Additionally, we report a novel single-base deletion c.641delG found in a previously reported patient from Belarus together with the previously identified missense mutation p.C76F

[Hrebicek et al., 2006] and a single base insertion, c.1271dupG. This latter mutation was found in two patients from Greece, in combination with the missense mutation p.S541L in one, and with p.E471K in the other.

**Nonsense mutations**

The introduction of a PTC via a single nucleotide base change was previously thought to result in the synthesis of a truncated protein, however recent research has shown that PTC-carrying transcripts are rapidly degraded via the nonsense-mediated mRNA decay (NMD) pathway protecting the cell from potentially harmful effects of truncated proteins [Muhlemann et al., 2008; McGlincy and Smith, 2008]. Six nonsense mutations in the *HGSNAT* gene have thus far been described in the literature. Our first report included 4 nonsense mutations: p.W316X, p.L321X, p.R384X and p.R506X. The latter mutation was found in 10 patients, 5 of which were Polish, the others being from the rest of Europe and North America. Ruijter et al. [2007] reported two nonsense mutations, p.R203X and p.W510X, present in heterozygosis in patients of Dutch ancestry. We report here 2 novel nonsense mutations: p.S296X was found in homozygosis in one UK patient of Pakistani ancestry and p.Y558X was found in heterozygosis with c.1542+4dupA in another UK patient. Parental DNA or patient cDNA were not available for sequencing confirmation.

**Missense mutations**

Missense mutations, causing amino acid changes comprise the majority of the MPS IIIC mutations identified to date, with 16 mutations already reported in the literature. Here we present 2 novel mutations: p.G486E found in homozygosis in a German patient and p.A489E found in a patient from the USA. The p.A489E mutation was confirmed in maternal allele, but no parental DNA or patient cDNA was available to confirm the p.G486E change.

Phenotypic pathogenicity of missense mutations is difficult to predict based strictly on genotype, so expression of the mutant proteins is normally required to understand the effect of the amino acid change on folding, targeting, stability and activity of the enzyme. Since no such studies have been reported so far, we evaluated here (Table 1) potential changes of the native protein structure caused by the mutations based on the nature of the amino acid, its evolutionary conservation and position within the protein. Eight missense mutations, p.G262R, p.P283L, p.W403C, p.M482K, p.G486E, p.A489E, p.S518F, and p.P571L are predicted to reside within the highly hydrophobic transmembrane domains of the protein. Five of these mutations introduce hydrophilic or charged residues inside the transmembrane domains potentially affecting folding of the protein [reviewed in Gregersen et al., 2001] and drastically reducing its activity. Four other changes result in replacement of hydrophilic residues for hydrophobic ones that also could destabilize the transmembrane helix. Six mutations are found adjacent to the predicted transmembrane domains either on the cytoplasmic (p.D562V) or on the lumenal (p.R344C/H, p.E471K, p.S539C, p.S541L) side and 4 mutations

(p.C76F, p.L137P, p.N273K, p.G424S) reside inside the hydrophilic lumenal domains of the enzyme. These mutations may affect areas important for binding of AcCoA, its transport through the lysosomal membrane or for the catalytic activity of the enzyme, and therefore knowledge of the protein structure is required for understanding their effects.

**Polymorphisms**

Previously [Hrebicek et al. 2006] we have described so-called complex mutations in the *HGSNAT* gene, consisting either of a splice site mutation found in *cis* with a missense mutation or of two missense mutations in *cis*. In particular, the combination of mutations c.234+1G>A and p.P237Q was identified in three patients of Moroccan origin and in one patient from Spain, the combination of c.493+1G>A and p.K523Q was identified in a patient of Algerian origin and the combination of p.W403C and pA615T was identified in a patient from the UK and a patient from France. Of note, the mutation c.493+1G>A was also identified in four other patients, one from Canada and one from Singapore, but without the presence of the accompanying p.K523Q change, whereas the mutation c.234+1G>A without accompanying p.P237Q was found in patients from France, Italy and Turkey. The missense mutation p.V481L was identified in a patient from the UK together with a splice site mutation c.1542+4dupA and the nonsense mutation p.Y558X.

At the time of publication it was difficult to deduce which of these missense mutations identified as a part of complex mutations could be classified

as polymorphisms and which represented a clinically relevant genotype, although later we have identified a c.1843G>A (p.A615T) change present in heterozygosis in one normal control DNA sample. In an effort to resolve this issue we have evaluated effects of the above amino acid changes on the activity of HGSNAT by expressing the enzyme carrying these changes in COS-7 cells. While p.P237Q, p.V481L, p.K523Q and p.A615T mutants had enzymatic activity similar to that of the wild type HGSNAT, the p.W403C mutant completely lost its activity (data not shown). These data likely suggest that amino acid changes p.P237Q, p.V481L, p.K523Q and p.A615T are rare polymorphisms and do not have clinical significance.

**CLINICAL AND DIAGNOSTIC RELEVANCE**

Although the spectrum of mutations in MPS IIIC patients shows high heterogeneity, some identified mutations have high frequency within a population or even across populations suggesting founder effects. First, a comprehensive study in MPS IIIC patients [Ruijter et al., 2007] has shown founder effects in the Dutch population where the two mutations p.R344C and p.S518F account for 22.0% and 29.3% respectively, of the alleles among the probands of Dutch origin. p.S518F mutation has been identified so far only in Dutch patients and in one patient from Germany, while the p.R344C change was also found in families from France, UK, Germany and Singapore. Fedele et al. [2007] showed that probands from 3 of the 10 studied families were homozygous for the mutation c.852-1G>A, and that all these patients had origins in southern Italy [Fedele et al.,

2007]. The same mutation was also found in a family from Turkey. Finally, the mutation c.525dupT reported by Countinho et al. [2008] was found in five out of six alleles of MPS IIIC patients from Portugal, also suggesting a possible founder effect. In support of this we report here 2 different Portuguese patients each possessing the c.525dupT mutation in homozygosis. The high frequencies of the above mutations suggest that molecular analysis can be a valid diagnostic option for MPS IIIC patients of Dutch, Portuguese and Italian origin.

Several other mutations showing a relatively high frequency among the MPS IIIC families are also characterized by wide distribution which doesn't suggest a founder effect in any particular population but rather a global founder effect of an ancient mutation. For example, p.R384X mutation was found in both homo- and heterozygous fashion in 10 patients from 6 different countries (Poland, Czech Republic, Italy, The Netherlands, Canada and Turkey). Similarly, the most common splice site mutation, c.234+1G>A, is present in 7 patient families, one from France, one from Italy, one from Spain, one from Turkey, two from Morocco and one from Canada where both parents are of Moroccan origin. The second most common splice site mutation c.493+1G>A was reported in 3 Canadian families, one family from Algeria and one from Singapore. The point mutation p.R344H was found in 4 families from Eastern and Northern Europe (2 from Poland, one from Czech Republic and one from Finland). This mutation and the Dutch founder mutation p.R344C are caused by changes of the neighbouring nucleotides, c.1031 and c.1030, respectively, and affect the same amino acid residue. Finally, missense mutation p.S541L was reported in 4

families from France, Ireland, Poland and Portugal. As all mutations discussed in this paragraph change C to T or G to A in a CpG dinucleotide, it is also possible that they arose independently in different populations due to deamination of methylcytosine. All other mutations have only been found in one to three families.

The majority of MPS IIIC patients described so far show severe clinical phenotype accompanied by a complete or almost complete loss of N-acetyltransferase activity in their leukocytes or cultured skin fibroblasts. Therefore, no genotype-phenotype correlations or predictions of clinical outcome of the mutations can be reported at this point with an exception of the mutations p.G262R and p.S539C that probably correlate with an attenuated MPS III phenotype as they were only found in two sisters with late-onset form of the disease [Ruijter et al., 2007; Berger-Plantinga, 2004]. In addition, we have identified a splice site mutation c.118+1G>A and one missense mutation p.S541L in 3 sisters of Irish ancestry affected with a late-onset MPS IIIC, but it is impossible to confirm at this point whether p.S541L results in partially functional protein or whether a part of mRNA produced by the c.118+1G>A allele is correctly spliced.

Presently no studies have been reported addressing the effect of detected mutations in the *HGSNAT* gene locus on the enzyme activity, stability, localization, processing and ability to interact with other proteins. Such studies especially important in the case of missense mutations will allow to clarify the pathogenic mechanism of MPS IIIC, establish the correlations between the type

of mutation and severity of the biochemical and clinical phenotype useful for clinical prognosis and may also suggest novel therapeutic approaches. Such studies may also shed light on the molecular structure and function of the enzyme since naturally occurring mutations are known to cause the clinically significant deficiency of the enzyme and therefore appear in functionally important areas.

In the vast majority of patients at least one missense mutation affects an amino acid residue in or adjacent to the transmembrane domains of HGSNAT, probably interfering with the proper folding of the enzyme. If the expression studies confirm this hypothesis, active site-specific inhibitors of the enzyme could be tested for their ability to stabilize the mutants and modify their conformation closer to that of the wild type enzyme increasing the level of their residual activity as it has been demonstrated for other lysosomal enzymes [reviewed in Fan, 2003; Desnick, 2004]. Together with the inhibitors of heparan sulfate synthesis [Piotrowska et al., 2006; Jakóbkiewicz-Banecka et al., 2007], such "pharmacological chaperones" could probably be an interesting subject of future research in this area.

**Acknowledgements**

The authors thank the patients, and their families for participating in our study. This work was supported by the research grants from March of Dimes Foundation, Sanfilippo Children's Research Foundation and Canadian Institutes for Health Research (MOP84430) to A.V.P, and institutional projects from the Ministry of Education of the Czech Republic (projects MSM0021620806, and 1M6837805002)

**REFERENCES**

Ausseil J, Loredo-Osti JC, Verner A, Darmond-Zwaig C, Maire I, Poorthuis B, van Diggelen OP, Hudson TJ, Fujiwara TM, Morgan K, Pshezhetsky AV. 2004. Localisation of a gene for mucopolysaccharidosis IIIC to the pericentromeric region of chromosome 8. J Med Genet. 41:941-945.

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

Bame KJ, Rome LH. 1985. Acetyl-coenzyme A: α-glucosaminide N-acetyltransferase; evidence for a transmembrane acetylation mechanism. J Biol Chem 260:11293-1299.

Bame KJ, Rome LH. 1986. Acetyl-coenzyme A: α-glucosaminideN-acetyltransferase: evidence for an active site histidine residue. J Biol Chem 261:10127-10132.

Bame KJ, Rome LH. 1986. Genetic evidence for transmembrane acetylation by lysosomes. Science 233:1087-1089.

Bartsocas C, Grobe H, van de Kamp JJP, von Figura K, Kresse H, Klein U, Giesberts MAH. 1979. Sanfilippo type C disease: clinical findings in four patients with a new variant of mucopolysaccharidosis III. Eur J Pediatr 130:251-258.

Berger-Plantinga EG, Vanneste JA, Groener JE, van Schooneveld MJ. 2004. Adult-onset dementia and retinitis pigmentosa due to mucopolysaccharidosis III-C in two sisters. J Neurol. 251:479-481.

Bonifacino JS, Traub LM. 2003. Signals for sorting of transmembrane proteins to endosomes and lysosomes. Annu Rev Biochem. 72:395-447.

Coutinho MF, Lacerda L, Prata MJ, Ribeiro H, Lopes L, Ferreira C, Alves S. 2008. Molecular characterization of Portuguese patients with mucopolysaccharidosis IIIC: two novel mutations in the HGSNAT gene. Clin Genet 74:194-195.

Desnick RJ. 2004. Enzyme replacement and enhancement therapies for lysosomal diseases. J Inherit Metab Dis 27(3):385-410.

Fan J-Q. 2003. A contradictory treatment for lysosomal storage disorders: inhibitors enhance mutant enzyme activity. Trends Pharmacol Sci 24:355-360.

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

Fedele AO, Filocamo M, Di Rocco M, Sersale G, Lübke T, di Natale P, Cosma MP, Ballabio A. 2007. Mutational analysis of the HGSNAT gene in Italian patients with mucopolysaccharidosis IIIC (Sanfilippo C syndrome). Hum Mutat 28:523.

Gallagher JT. 2006. Multiprotein signalling complexes: regional assembly on heparan sulphate. Biochem Soc Trans 34(Pt 3):438-441.

Gregersen N, Bross P, Andrese BS, Pedersen CB, Corydon TJ, Bolund L. 2001. The role of chaperone-assisted folding and quality control in inborn errors of metabolism: protein folding disorders. J Inherit Metab Dis. 24(2):189-212.

Hopwood JJ, Freeman C, Clements PR, Stein R, Miller AL. 1983. Cellular location of N-acetyltransferase activities toward glucosamine and glucosamine-6-phosphate in cultured human skin fibroblasts. Biochem Int 6:823-830.

Hřebíček M, Mrázová L, Seyrantepe V, Durand S, Roslin NM, Nosková L, Hartmannová H, Ivánek R, Čížková A, Poupětová H, Sikora J, Uřinovská J, Stránecký V, Zeman J, Lepage P, Roquis D, Verner A, Ausseil J, Beesley CE, Maire I, Poorthuis B, van de Kamp J, van Diggelen OP, Wevers RA, Hudson TJ, Fujiwara TM, Majewski J, Morgan K, Kmoch S, Pshezhetsky AV. 2006. Mutations in TMEM76 cause mucopolysaccharidosis IIIC (Sanfilippo C syndrome). Am J Hum Genet 79:807-819.

Jakóbkiewicz-Banecka J, Wegrzyn A, Wegrzyn G. 2007. Substrate deprivation therapy: a new hope for patients suffering from neuronopathic forms of inherited lysosomal storage diseases. J Appl Genet. 48:383-388.

Klein U, Kresse H, von Figura K. 1978. Sanfilippo syndrome type C: deficiency of acetyl-CoA: α-glucosaminide N-acetyltransferase in skin fibroblasts. Proc Natl Acad Sci USA 75:5185-5189.

Klein U, van de Kamp JJP, von Figura K, Pohlmann R. 1981. Sanfilippo syndrome type C: assay for acetyl-CoA: α-glucosaminide N-acetyltransferase in leukocytes for detection of homozygous and heterozygous individuals. Clin Genet 20:55-59.

Kresse H, Von Figura K, Bartsocas C. 1976. Clinical and biochemical findings in a family with Sanfilippo disease, type C. (Abstract) Clin Genet 10:364.

McGlincy NJ, Smith CWJ. 2008. Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense? Trends Biochem Sci 33(8):385-393.

Meikle PJ, Hopwood JJ, Clague AE, Carey WF. 1999. Prevalence of lysosomal storage disorders. JAMA 281:249-254.

Mühlemann O, Eberle AB, Stalder L, Orozco RZ. 2008. Recognition and elimination of nonsense mRNA. Biochim Biophys Acta 1779(9):538-549.

Pinto R, Caseiro C, Lemos M, Lopes L, Fontes A, Ribeiro H, Pinto E, Silva E, Rocha S, Marcao A, Ribeiro I, Lacerda L, Ribeiro G, Amaral O, Sa Miranda MC. 2004. Prevalence of lysosomal storage diseases in Portugal. Eur J Hum Genet 12:87-92.

Piotrowska E, Jakóbkiewicz-Banecka J, Barańska S, Tylki-Szymańska A, Czartoryska B, Wegrzyn A, Wegrzyn G. 2006. Genistein-mediated inhibition of glycosaminoglycan synthesis as a basis for gene expression-

targeted isoflavone therapy for mucopolysaccharidoses. Eur J Hum Genet 14:846-852.

Pohlmann R, Klein U, Fromme HG, von Figura K. 1981. Localisation of acetyl-CoA: α-glucosaminide N-acetyltransferase in microsomes and lysosomes of rat liver. Hoppe Seylers Z Physiol Chem 362:1199-1207.

Poorthuis BJ, Wevers RA, Kleijer WJ, Groener JE, de Jong JG, van Weely S, Niezen-Koning KE, van Diggelen OP. 1999. The frequency of lysosomal storage diseases in The Netherlands. Hum Genet 105:151-156.

Rome LH, Hill DF, Bame KJ, Crain LR. 1983. Utilization of exogenously added acetyl coenzyme A by intact isolated lysosomes. J Biol Chem 258:3006-3011.

Ruijter GJ, Valstar MJ, van de Kamp JM, van der Helm RM, Durand S, van Diggelen OP, Wevers RA, Poorthuis BJ, Pshezhetsky AV, Wijburg FA. 2008. Clinical and genetic spectrum of Sanfilippo type C (MPS IIIC) disease in The Netherlands. Mol Genet Metab 93:104–111.

Sanfilippo SJ, Podosin R, Langer LO Jr, Good RA. 1963. Mental retardation associated with acid mucopolysaccharidūria (heparitin sulfate type). J Pediatr 63:837-838.

Spyropoulos IC, Liakopoulos TD, Bagos PG, Hamodrakas SJ. 2004. TMRPres2D: high quality visual representation of transmembrane protein models. Bioinformatics 20:3258-3260.

Valstar MJ, Ruijter, GJG, van Diggelen OP, Poorthuis BJ, Wijburg FA. 2008. Sanfilippo syndrome: A mini-review. J Inherit Metab Dis. 31:240-252.

Table I. Mutations in the *HGSNAT* Gene*

| Nucleotide change[a] | Predicted effect on protein | Exon | Effect | Origin | Reference |
|---|---|---|---|---|---|
| *Splicing site mutations* | | | | | |
| c.118+1G>A | | Intron 1 | | Ireland[b] | Hřebíček et al. [2006] |
| c.234+1G>A | p.D40VfsX19[c] | Intron 2 | | France, Morocco, Spain<br>Italy<br>Morocco, Turkey | Hřebíček et al. [2006]<br>Fedele et al. [2007]<br>Ruijter et al. [2008] |
| c.372-2A>G | | Intron 3 | Skipping of exon 4 | Portugal | Coutinho et al. [2008] |
| c.493+1G>A | | Intron 4 | | Canada<br>Algeria[d], Canada<br>**Singapore** | Fan et al. [2006]<br>Hřebíček et al. [2006]<br>This report |
| c.743+1dupG | p.I249DfsX31 | Intron 7 | | Netherlands | Ruijter et al. [2008] |
| **c.744-2A>G** | | Intron 7 | | Pakistan | This report |
| c.851+5G>A | p.F285X[c] | Intron 9 | Inclusion of 89 bases from 5'end of intron 9, skipping of exon 10. PTC | Czech Republic | Hřebíček et al. [2006] |
| c.852-1G>A | | Intron 9 | | Italy<br>**Turkey** | Fedele et al. [2007]<br>This report |
| c.1250+1G>A | | Intron 12 | | Poland<br>**Turkey**, **UK** | Hřebíček et al. [2006]<br>This report |
| **c.1250+2T>C** | | Intron 12 | | USA | This report |
| c.1464+1G>A | | Intron 14 | | Italy | Fedele et al. [2007] |
| **c.1542+4dupA** | | Intron 15 | | UK | This report |
| c.1726+1G>A | p.S539NfsX14[c] | Intron 17 | Skipping of exon 17, frameshift | Turkey | Hřebíček et al. [2006]<br>Ruijter et al. [2008] |
| *Insertions and deletions* | | | | | |
| **c.1-1925_118+296del** | | 1 | 2339 bp deletion including exon 1 | Turkey | This report |
| c.525dupT | p.V176CfsX16 | 5 | | Portugal | Coutinho et al. [2008] |
| **c.641delG** | p.Gly214AspfsX62 | 7 | | Belarus | This report |
| c.682_740del | p.P228GfsX32 | 7 | | Italy | Fedele et al. [2007] |
| c.739delA | p.R247GfsX29 | 7 | | Italy<br>**North Africa** | Fedele et al. [2007]<br>This report |
| c.1034_1049del | p.I345SfsX5 | 11 | | Czech Republic | Hřebíček et al. [2006] |
| **c.1271dupG** | **p.I425HfsX45** | 13 | | Greece | This report |
| c.1336_1372dup | p.V460GfsX22 | 13 | | Poland | Hřebíček et al. [2006] |
| c.1345dupG | p.D449GfsX21 | 13 | | Canada | Fan et al. [2006] |
| c.1669_1674delinsACAT | p.L557TfsX33 | 17 | | Italy | Fedele et al. [2007] |
| c.1750delG | p.V584SfsX16 | 18 | | Finland | Hřebíček et al. [2006] |
| *Nonsense mutations* | | | | | |
| c.607C>T | p.R203X | 6 | PTC | Netherlands | Ruijter et al. [2008] |
| **c.887C>A** | **p.S296X** | 10 | PTC | Pakistan | This report |
| c.947G>A | p.W316X | 10 | PTC | France | Hřebíček et al. [2006] |
| c.962T>G | p.L321X | 10 | PTC | Czech Republic | Hřebíček et al. [2006] |
| c.1150C>T | p.R384X | 12 | PTC | Canada, Czech Republic, Poland, Turkey | Hřebíček et al. [2006] |

| | | | | | |
|---|---|---|---|---|---|
| | | | | Italy<br>Netherlands, Turkey | Fedele et al. [2007]<br>Ruijter et al. [2008] |
| c.1516C>T | p.R506X | 15 | PTC | Czech Republic<br>**Belgium** | Hřebíček et al. [2006]<br>This report |
| c.1530G>A | p.W510X | 15 | PTC | Netherlands | Ruijter et al. [2008] |
| **c.1674C>G** | **p.Y558X** | 17 | PTC | UK | This report |
| *Missense mutations* | | | | | |
| c.227G>T | p.C76F | 2 | Changes a polar residue for a non-polar in the lysosomal lumen | Belarus | Hřebíček et al. [2006] |
| c.410T>C | p.L137P | 4 | Changes a highly hydrophobic residue for a hydrophilic one in the lysosomal lumen | Netherlands<br>**Germany** | Ruijter et al. [2008]; Fedele et al. [2007]<br>This report |
| c.784G>A | p.G262R | 8 | Changes a neutral non-polar hydrophobic residue for a strongly basic polar hydrophilic one in the membrane | Netherlands | Ruijter et al. [2008] |
| c.819T>G | p.N273K | 8 | Changes a neutral residue for a basic in the lysosomal lumen | Netherlands | Ruijter et al. [2008] |
| c.848C>T | p.P283L | 9 | Changes a hydrophilic residue for a hydrophobic in the membrane | France, UK<br>**Turkey** | Hřebíček et al. [2006]<br>This report |
| c.1030C>T | p.R344C | 11 | Changes a strongly basic and highly hydrophilic residue for a neutral and highly hydrophobic one just inside the lysosomal lumen | France, UK<br>Netherlands<br>**Germany, Singapore** | Hřebíček et al. [2006]<br>Ruijter et al. [2008]<br>This report |
| c.1031G>A | p.R344H | 11 | Changes a strongly basic residue for a weakly basic just inside the lysosomal lumen | Czech Republic, Finland, Poland | Hřebíček et al. [2006] |
| c.1209G>T | p.W403C | 12 | Changes a non-polar hydrophilic residue to a neutral hydrophobic one in membrane | France<br>**UK** | Hřebíček et al. [2006]<br>This report |
| c.1270G>A | p.G424S | 13 | Changes a non-polar residue for a polar in the lysosomal lumen | Canada | Hřebíček et al. [2006] |
| c.1411G>A | p.E471K | 14 | Changes an acidic residue for a basic just inside the lysosomal lumen | Canada, France<br>Netherlands<br>**Greece** | Hřebíček et al. [2006]<br>Ruijter et al. [2008]<br>This report |

| | | | | | |
|---|---|---|---|---|---|
| c.1445T>A | p.M482K | 14 | Changes a non-polar neutral hydrophobic residue for a polar basic hydrophilic one inside the membrane | Czech Republic | Hřebíček et al. [2006] |
| **c.1457G>A** | **p.G486E** | 14 | Changes neutral non-polar hydrophobic residue for a a polar acidic hydrophilic one inside the membrane | Germany | This report |
| **c.1466C>A** | **p.A489E** | 15 | Changes a non-polar neutral hydrophobic residue for a polar acidic hydrophilic one inside the membrane | USA | This report |
| c.1553C>T | p.S518F | 16 | Changes a polar hydrophilic residue for a non-polar hydrophobic one inside the membrane | Netherlands<br><br>**Germany** | Fedele et al. [2007]; Ruijter et al. [2008]<br><br>This report |
| c.1616C>G | p.S539C | 17 | Changes a hydrophilic residue for a hydrophobic inside the lysosomal lumen | Netherlands | Ruijter et al. [2008] |
| c.1622C>T | p.S541L | 17 | Changes a polar hydrophilic residue for a non-polar hydrophobic one just inside the lysosomal lumen | France, Ireland[b], Poland, Portugal,<br><br>**Greece** | Hřebíček et al. [2006]<br><br>This report |
| c.1685A>T | p.D562V | 17 | Changes a polar acidic highly hydrophilic residue for a non-polar neutral highly hydrophobic one in the cytoplasm | France | Hřebíček et al. [2006] |
| c.1712C>T | p.P571L | 17 | Changes a hydrophilic residue for a hydrophobic in the membrane | Czech Republic | Hřebíček et al. [2006] |
| *Polymorphisms* | | | | | |
| c.710C>A | p.P237Q | 7 | Changes a non-polar for a polar residue in the cytoplasm | Morocco, Spain<br><br>Morocco | Hřebíček et al. [2006]<br><br>Ruijter et al. [2008] |
| **c.1441G>T** | **p.V481L** | 14 | Conservative residue change | UK | This report |
| c.1567A>C | p.K523Q | 16 | Changes a basic residue for a neutral just inside the lysosomal lumen | Algeria[d] | Hřebíček et al. [2006] |
| c.1843G>A | p.A615T | 18 | Changes a non-polar hydrophobic residue for a polar hydrophilic one inside the membrane | France<br><br>**UK** | Hřebíček et al. [2006]<br><br>This report |

Table 1.*Novel mutations and patients are shown in bold. The following mutation was renamed according to the Human Genome Variation Society: c.743+1insG to c.743+1dupG

[a] The numbering for the nucleotide changes are based on cDNA sequence in accordance with GenBank entries NM_152419.2 and NG_009552.1, with +1 corresponding to the A of the ATG translation initiation codon in the reference sequence.

[b]Patient carrying this mutation originally published as from USA in Hřebíček et al. [2006]

[c]Effect confirmed by cDNA sequencing.

[d]Patient carrying this mutation originally published as from France in Hřebíček et al. [2006]. PTC, premature termination codon

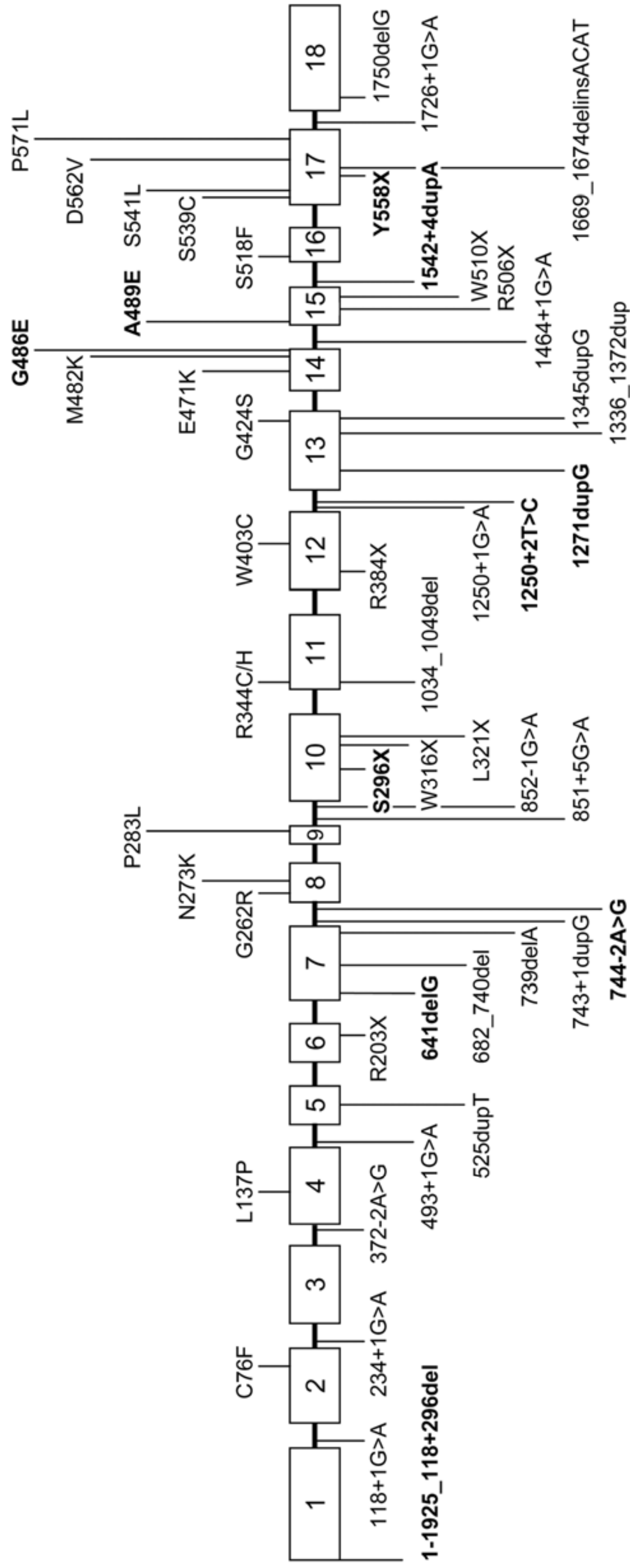
**Figure 1. Distribution of missense mutations in HGSNAT protein.** Visual representation of HGSNAT membrane topology was created using the TMRPres2D software [Spyropoulos et al. 2004]. Mutated residues are shown in white letters on black background and predicted N-glycosylation sites are boxed.

CYTOPLASM

LYSOSOME

L137P
C76F
G262R
P283L
N273K
R344C/H
W403C
G424S
A489E
G486E
M482K
E471K
S518F
S541L
S539C
D562V
P571L

**Figure 2. Location of MPS IIIC mutations in the *HGSNAT* gene.** The exons are represented by boxes. The positions of the missense mutations are shown above the gene. Other mutations (nonsense, insertions, deletions and splice-site mutations) are shown below the gene. Nonsense and missense mutations are presented based on amino acid changes, while splice-site mutations, insertions and deletions are presented based on nucleotide changes. Novel mutations are shown in bold.


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
118+1G>A
1-1925_118+296del
C76F
234+1G>A
372-2A>G
L137P
493+1G>A
525dupT
R203X
641delG
682_740del
739delA
743+1dupG
744-2A>G
G262R
N273K
P283L
S296X
W316X
L321X
852-1G>A
851+5G>A
R344C/H
1034_1049del
R384X
W403C
1250+1G>A
1250+2T>C
1271dupG
G424S
E471K
M482K
G486E
1345dupG
1336_1372dup
A489E
1464+1G>A
W510X
R506X
1542+4dupA
S518F
S539C
S541L
D562V
P571L
Y558X
1669_1674delinsACAT
1726+1G>A
1750delG

# CHAPTER 3

# Characterization of molecular defects in HGSNAT gene

# Foreword

MPS IIIC patients with missense mutations experience a severe disease progression similar to patients with mutations predicted to result in the complete absence of protein. In order to understand the effects these mutations have on the enzyme we expressed HGSNAT mutants containing each of the 17 reported missense mutations and 4 polymorphisms and examined their enzymatic activity, targeting and folding.

We determined that all the missense mutations caused HGSNAT to be misfolded. Furthermore we showed the HGSNAT missense mutants to be inactive, abnormally glycosylated and retained in the ER. We then identified glucosamine as a competitive inhibitor of HGSNAT, and showed that it was able to rescue activity in 8 of 9 patient fibroblast lines.

These studies utilized transient expression, enzyme assay, immunoblot and immunocytochemistry analyses.

All experiments as well as preparation of the manuscript was conducted by Matthew Feldhammer, with the exception of the mutants' generation and cloning, determination of the substrate kinetics and the glucosamine inhibition constant.

*Article published in PLOS One Journal following peer review

**Protein Misfolding as an Underlying Molecular Defect in Mucopolysaccharidosis III Type C**

**Running head: HGSNAT misfolding in MPS IIIC**

Matthew Feldhammer[1,2†], Stéphanie Durand[1†], and Alexey V. Pshezhetsky[1,2,3,4]

[1]Department of Medical Genetics, CHU Sainte-Justine and Departments of [2]Biochemistry and [3]Pediatrics, University of Montreal, Montreal, Canada; [4]Department of Anatomy and Cell Biology, Faculty of Medicine, McGill University, Montreal, Canada.

Correspondence should be addressed to:

Alexey V. Pshezhetsky, Ph.D.

Service de génétique médicale, CHU Sainte-Justine

3175 Côte Ste-Catherine, Montréal (Qc) H3T 1C5 Canada

Tel. (514) 345-4931/2736

Fax. (514) 345-4801

E-mail: Alexei Pshezhetsky

[†] S.D. and M.F. equally contributed as first authors.

**ABSTRACT**

Mucopolysaccharidosis type IIIC or Sanfilippo syndrome type C (MPS IIIC, MIM #252930) is an autosomal recessive disorder caused by deficiency of the lysosomal membrane enzyme, heparan sulfate acetyl-CoA: α-glucosaminide N-acetyltransferase (HGSNAT, EC 2.3.1.78), which catalyses transmembrane acetylation of the terminal glucosamine residues of heparan sulfate prior to their hydrolysis by α-N-acetylglucosaminidase. Lysosomal storage of undegraded heparan sulfate in the cells of affected patients leads to neuronal death causing neurodegeneration and is accompanied by mild visceral and skeletal abnormalities, including coarse facies and joint stiffness. Surprisingly, the majority of MPS IIIC patients carrying missense mutations are as severely affected as those with splicing errors, frame shifts or nonsense mutations resulting in the complete absence of HGSNAT protein.

In order to understand the effects of the missense mutations in HGSNAT on its enzymatic activity and biogenesis we have expressed 21 mutant proteins in cultured human fibroblasts and COS-7 cells and studied their folding, targeting and activity. We found that 17 of the 21 missense mutations in HGSNAT caused misfolding of the enzyme, which is anbormally glycosylated and not targeted to the lysosome, but retained in the endoplasmic reticulum. The other 4 mutants represented rare polymorphisms which had no effect on the activity, processing and targeting of the enzyme. Treatment of patient cells with a competitive

HGSNAT inhibitor, glucosamine, partially rescued several of the expressed mutants.

Altogether our data provide an explanation for the severity of MPS IIIC and suggest that search for pharmaceutical chaperones can in the future result in therapeutic options for this disease.

**AUTHOR SUMMARY**

Mucopolysaccharidosis III type C (MPS IIIC) belongs to a group of genetic metabolic diseases affecting the lysosome, a cellular compartment responsible for degradation and recycling of biological macromolecules. Lysosomal diseases comprise ~14% of all inherited diseases of metabolism and, although individually rare, together they affect one of every 7700 newborn children [1]. MPS IIIC causes rapid progressive neurodegeneration in infants and children, accompanied by variable problems of the eyes, bones and viscera. Our current study provides an explanation for the severity of the disease. We showed that in the patient cells the lysosomal enzyme called N-acetyltransferase contains amino acid substitutions that prevent its normal folding, processing and trafficking. We also showed that a small competitive inhibitor of the N-acetyltransferase can partially restore the folding, trafficking and activity of several mutants which may provide future therapeutic solutions for this devastating untreatable disease.

## INTRODUCTION

Mucopolysaccharidosis III (also called Sanfilippo syndrome) is an autosomal recessive disease caused by lysosomal accumulation of heparan sulfate [reviewed in 2] and includes four allelic subtypes caused by the genetic deficiencies of heparan N-sulfatase (MPS III type A; MIM #252900), α-N-acetylglucosaminidase (MPS III type B; MIM #252920), heparan sulfate acetyl-CoA: α-glucosaminide N-acetyltransferase or HGSNAT (MPS III type C; MIM #252930), and N-acetylglucosamine 6-sulfatase (MPS III type D; MIM #252940). The majority of MPS IIIC patients have severe clinical manifestations with onset in infancy or early childhood. They rapidly develop progressive and severe neurological deterioration causing hyperactivity, sleep disorders, loss of speech accompanied by behavioral abnormalities, neuropsychiatric problems, mental retardation, hearing loss, and visceral manifestations, such as mild hepatomegaly, mild dysostosis multiplex, mild coarse facies, and hypertrichosis [3,4]. A majority of patients experience severe mental retardation and die before adulthood but some survive to the fourth decade with progressive dementia and retinitis pigmentosa [2,4]. In the very rare MPS IIIC patients with the onset of symptoms in adulthood the disease progression was similar in severity and time course to the forms with onset in childhood [5]. The birth prevalence of MPS IIIC in Australia, Portugal and the Netherlands was estimated at 0.07, 0.12 and 0.21 per 100,000, respectively [1,6,7].

Although from the moment of discovery MPS IIIC was recognized as a deficiency of an enzyme that transfers an acetyl group from cytoplasmically derived acetyl-CoA to terminal N-glucosamine residues of heparan sulfate within the lysosomes [8-10], the molecular defects causing the disease have not been characterized for almost three decades because the identification and cloning of HGSNAT has been hampered by low tissue content and instability of the enzyme. Recently our group and others cloned the gene coding for HGSNAT [11,12] which paved the way for characterization of the molecular defects in MPS IIIC patients. To date, 54 HGSNAT sequence variants have been identified including 13 splice-site mutations, 11 insertions and deletions causing frame shifts and premature termination of translation, 8 nonsense, 18 missense mutations and 4 polymorphisms [reviewed in 13, http://chromium.liacs.nl/LOVD2/home.php?select_db=HGSNAT]. While the HGSNAT transcripts with abnormal splicing, frame shifts and premature stop codons are rapidly degraded via the nonsense-mediated mRNA decay pathway [11], the phenotypic pathogenicity of missense mutations was impossible to predict. Here, in order to understand the biochemical effects of the missense mutations we have expressed mutant proteins in cultured human fibroblasts and COS-7 cells and found that all mutations result in misfolding of the enzyme. In the consequence, it is abnormally glycosylated, and is not targeted to the lysosome, but retained in the endoplasmic reticulum (ER).

# RESULTS

## Expression, processing and enzymatic activity of HGSNAT mutants

The effect of HGSNAT mutations on the enzyme biogenesis and catalytic activity was studied by the transient expression of the mutant cDNA. Mutations were generated by site directed mutagenesis in the pCTAP-HGSNAT construct that expresses human HGSNAT with a C-terminal tandem affinity purification (TAP) tag consisting of a high affinity streptavidin-binding peptide (SBP) and a calmodulin-binding peptide (CBP) to allow purification of the recombinant protein using successively-applied streptavidin-resin and calmodulin-resin affinity purification steps or its detection with anti-CBP antibodies [14,15]. In preliminary experiments we have ensured that a human wild-type HGSNAT carrying a TAP tag on its C-terminus had a lysosomal localization and showed catalytic activity similar to that of the untagged enzyme (not shown). The sequences of the constructs were verified to ensure the correct introduction of mutations. In total, 21 constructs bearing nucleotide changes identified in patients were transfected in COS-7 cells and assayed for the N-acetyltransferase activity (Figure 1). All tested cells showed similar endogenous lysosomal β-hexosaminidase activity (not shown) but drastically different levels of the N-acetyltransferase activity. The four polymorphisms (P237Q, V481L, K523Q and A615T) displayed activity similar to that of the wild-type enzyme, whereas the activity of the rest of the mutants was below the detection level. Kinetics studies were further conducted using partially purified enzyme to determine if the P237Q

and V481L mutants showing full enzymatic activity would have a different affinity for the substrate, however their $K_m$ and $V_{max}$ values were similar to those of the wild-type (Supplementary figure 1, Appendix I).

The expression of HGSNAT mutants was studied by Western blot analysis of cellular homogenates using anti-CBP antibodies (Figure 1B). All mutants were expressed at a level approximately similar to that of the wild-type HGSNAT, but showed a difference in their molecular mass. The four polymorphisms (P237Q, V481L, K523Q and A615T) showed a molecular mass of ~77 kDa, similar to that of the wild-type enzyme, whereas all missense mutants (C76F, L137P, G262R, N273K, P283L, R344C, R344H, W403C, G424S, E471K, M482K, A489E, S518F, S539C, S541L, D562V, P571L) had a smaller molecular mass of ~67 kDa. Since HGSNAT is predicted to have 5 potential N-linked glycosylation sites each potentially contributing ~2 kDa to the size of the mature enzyme [11], this molecular mass difference was consistent with the hypothesis that inactive mutants lacked proper glycosylation. To confirm this, the wild-type enzyme, 2 polymorphic variants and 2 inactive mutants were treated with endoglycosidase H which cleaves immature mannose-rich oligosaccharide chains added to glycoproteins in the ER [16] Analysis by Western blot (Figure 2) showed that upon deglycosylation the molecular mass of the wild-type enzyme and polymorphic mutants was reduced to 67 kDa, whereas the inactive mutants displayed only a subtle shift in their position on the gel, indicating that they already lacked most of the glycans. The homogenates were also treated with the peptide: N-Glycosidase F (PNGase F) which fully removes

all sugars from mature glycoproteins (Figure 2B). In addition to performing enzymatic deglycosylation of cell homogenates we treated the COS-7 cells transiently expressing the wild-type enzyme, a polymorphic and an inactive mutant with tunicamycin, an inhibitor of N-acetylglucosamine transferase that blocks glycosylation of newly-synthesized proteins (Figure 2C). Results of both experiments were consistent with those obtained with Endoglycosidase H. Upon the treatment with tunicamycin or PNGase F the size of the wild type HGSNAT and the enzyme containing the polymorphism was reduced by approximately ~10 kDa whereas the inactive mutant displayed only a minor change in electrophoretic mobility. Western blot analysis also revealed that the homogenates of the cells transfected with both wild-type HGSNAT and polymorphic mutants, but not of untransfected cells or cells transfected with inactive mutants contained an immunoreactive protein with a molecular mass of ~48 kDa. Since this form of HGSNAT contains the C-terminal CBP cross-reacting with the antibodies we estimated that it corresponds to the HGSNAT fragment missing ~300 amino acids on the N-terminus. It seems that the appearance of this form is associated with the proper glycosylation and lysosomal targeting of HGSNAT, but further studies are necessary to conclude whether it represents a product of intralysosomal maturation (with the N-terminal chain not being detected by our antibodies) or a degradation product. On the other hand, the 77-kDa form of HGSNAT separated by FPLC anion-exchange chromatography showed full enzymatic activity (Figure 3).

**Subcellular localization of HGSNAT mutants**

The effects of missense mutations on the subcellular localization of the enzyme were studied by confocal immunofluorescence microscopy. Immortalized human skin fibroblasts from a normal control were transfected with constructs expressing each of the missense mutations. The cells were allowed to express the mutant and polymorphic enzymes for 42 hours and were then fixed by paraformaldehyde. To identify the lysosomal-late endosomal compartment, prior to fixation the cells were treated with LysoTracker Red DND-99 dye. After fixation the cells were permeabilized with Triton X-100 and probed with anti-CBP antibodies to localize the HGSNAT and with antibodies against the ER marker calnexin and lysosomal marker LAMP-2.

Distinct punctate staining characteristic of lysosomal targeting of the protein was evident for the recombinant wild-type enzyme and all active mutants. Accordingly, both wild-type recombinant HGSNAT and polymorphic mutants almost completely co-localized with the lysosomal markers LysoTracker Red or LAMP-2 (representative data are shown in Figure 4). In contrast, the inactive mutants exhibited a diffuse pattern throughout the cell and were co-localized not with the lysosomal markers but with the ER marker calnexin (representative data are shown in Figure 4, see Supplementary figure 2 for the data on all mutants, Appendix II). Partial ER localization (typically 2-5% of the total, not shown) was also observed for the wild-type recombinant and active mutants and most likely represented a pool of the enzyme processed in the ER on its way to the lysosomes.

**Glucosamine-mediated refolding of inactive HGSNAT mutants**

Since all the data were consistent with general folding defects and retention in the ER compartment of the HGSNAT mutants (C76F, L137P, G262R, N273K, P283L, R344C, R344H, W403C, G424S, E471K, M482K, A489E, S518F, S539C, S541L, D562V, P571L) we further tested whether competitive inhibitors of the enzyme that mimic the substrate binding in the active site would help to fold the enzyme in the ER so it can be properly modified and exported to the lysosome. We first tested several potential inhibitors of HGSNAT that were not expected to be highly toxic for the cultured cells and found that one of them, D-(+)-glucosamine hydrochloride, was a competitive inhibitor of the enzyme with a $K_I$ of 0.28 mM close to the $K_M$ value for the 4MU-βGlcN substrate (Supplementary figure 3, Appendix III). To establish the optimal concentration of the inhibitor in the culture medium and the length of the treatment, we have used the immortalized skin fibroblasts of previously reported MPS IIIC patient homozygous for the N273K mutation that causes misfolding of the enzyme (Figure 1 and Figure 4), since they were readily available in our lab [25]. Immortalized fibroblasts from a patient homozygous for a splice site mutation (c.1726+1G>A) [11] and normal immortalized fibroblasts were used as controls. We found that the treatment of fibroblasts from a patient homozygous for the N273K mutation with 14 mM glucosamine results in the progressive increase of the N-acetyltransferase activity in the cell homogenates which reaches 3 fold induction after 5 days of treatment (Figure 5A). The N-acetyltransferase

activity in the normal cells and in the cells of the MPS IIIC patient homozygous for the splice site mutation c1726+1G>A as well as the control β-hexosaminidase activity in all cell lines were not increased upon the treatment (not shown) suggesting that the observed induction of N-acetyltransferase activity was caused by rescuing the N273K mutant but not by the general induction of lysosomal enzymes. When the cells of the patient affected with the N273K mutation were treated for 5 days with increasing concentrations of glucosamine (Figure 5B) we found that the effect was proportional to the concentration of the inhibitor. At the glucosamine concentration of 14 mM (~50x $K_I$) the N-acetyltransferase activity in the treated cells was 0.3 nmol/h mg which corresponds to ~7% of the average activity in normal control fibroblasts, but further increase of glucosamine concentration resulted in inhibition of cell growth. To understand whether similar effects could be also observed for other missense mutants the available primary cultures of skin fibroblasts from 9 MPS IIIC patients carrying missense HGSNAT mutations or a missense mutation in combination with a splice site or nonsense mutation were treated with glucosamine and assayed for N-acetyltransferase activity. Nine missense mutations were studied altogether (Figure 5C; patient genotypes are listed in the figure legend). Although the maximal effect for different cell lines was achieved at different concentrations of glucosamine and at different treatment time (not shown), for 8 of 9 cell lines a statistically significant increase in the activity was observed (Figure 5C), suggesting that some of the mutants were partially stabilized by the glucosamine, correctly processed and targeted to the lysosomal compartment.

**DISCUSSION**

The current work provides an explanation for the severe phenotype of MPS IIIC patients carrying missense mutations in the *HGSNAT* gene. Altogether, we studied the activity and biogenesis of HGSNAT mutants having 21 amino acid substitutions previously identified in MPS IIIC families and affecting 8 of the 11 transmembrane segments of the enzyme as well as its luminal and cytosolic domains (Figure 6). The missense mutations studied in this paper represent all of the currently identified missense HGSNAT mutations. The polymorphisms (P237Q, V481L, K523Q and A615T, shown in green in Figure 6) identified in MPS IIIC only in *cis* either with a splice site mutation or with a missense mutation [11,13] were previously reported by us to result in catalytically active protein [13]. They were included in the current study because it was unclear whether they still could affect kinetic parameters or targeting of HGSNAT and therefore represent a clinically valuable phenotype. All variants were expressed as TAP-tagged proteins since no Western blotting or cellular immunostaining could be conducted using the patient fibroblasts due to a lack of specific antibodies against human HGSNAT. Our results show that the HGSNAT variants P237Q, V481L, K523Q and A615T are all correctly processed targeted to the lysosome and display full enzymatic activity. We conclude therefore that these four mutations represent rare polymorphisms in the *HGSNAT* gene and do not have clinical significance thus confirming our previous hypothesis [13].

Seventeen mutations (C76F, L137P, G262R, N273K, P283L, R344C, R344H, W403C, G424S, E471K, M482K, A489E, S518F, S539C, S541L, D562V and P571L, shown in red in Figure 6) result in production of misfolded HGSNAT protein that is abnormally glycosylated and not targeted to the lysosome. Seven of these mutations (G262R, P283L, W403C, M482K, A489E, S518F and P571L) are predicted to reside within the highly hydrophobic transmembrane domains of the protein. Four of them (G262R, W403C, M482K, A489E) introduce hydrophilic or charged residues inside the transmembrane domains which usually have a dramatic effect on the folding of the protein [reviewed in 18]. Three other changes (P283L, S518F and P571L) result in replacement of hydrophilic residues for hydrophobic ones that also could destabilize the transmembrane helix.

Six mutations (R344C, R344H, E471K, S539C, S541L and D562V) are found adjacent to the predicted transmembrane domains either on the cytoplasmic (D562V) or on the lumenal (R344C, R344H, E471K, S539C, and S541L) side and 4 mutations (C76F, L137P, N273K, and G424S) reside inside the hydrophilic lumenal domains of the enzyme. In most cases these mutations are predicted to have a drastic effect on the protein folding since they involve replacements with amino acids significantly different in hydrophobicity (C76F, D562V, S541L), charge (R344C/H, N273K, E471K) or size (C76F, L137P, G424S). Thus, enzyme folding defects due to missense mutations, together with nonsense-mediated mRNA decay seem to be the major molecular mechanisms underlying MPS IIIC.

For at least 5 of the above changes (N273K, R344C, R344H, S518F and S541L) the active conformation can be stabilized by the competitive inhibitor of HGSNAT glucosamine resulting in part of the enzyme pool being properly processed and targeted to the lysosomes. L137P and P283L mutants may also be stabilized by the glucosamine treatment, however this could not be verified experimentally because in the available patient cell lines they were present together with the responsive mutations S518F and R344C, respectively. Only one cell line carrying E471K and D562V mutations did not show a significant increase in N-acetyltransferase activity in response to glucosamine. Further structural studies are needed to fully understand the difference in the effect of glucosamine on these mutants.

Although the spectrum of mutations in MPS IIIC patients shows substantial heterogeneity, some of the missense mutations have a high frequency within the patient population. Importantly, the two mutations, R344C and S518F, responsive to glucosamine-mediated refolding account for 22.0% and 29.3%, respectively, of the alleles among the probands of Dutch origin [19]. The S518F mutation has also been identified in a patient from Germany, while the R344C change was found in families from France, UK, Germany and Singapore. The responsive mutation R344H was found in 4 families from Eastern and Northern Europe (2 from Poland, one from Czech Republic and one from Finland) and the responsive mutation S541L was reported in 4 families from France, Ireland, Poland and Portugal. In general, the vast majority of patients is affected with at least one missense mutation interfering with the proper folding of the enzyme that

could be partially rescued by the treatment of the cells with the competitive inhibitor of HGSNAT, glucosamine. We believe this makes MPS IIIC a good candidate for enzyme enhancement therapy [reviewed in 20,21] where active site-specific inhibitors are used as pharmacological chaperones to modify the conformation of the mutant lysosomal enzymes usually retained and degraded in the ER in order to increase the level of the residual activity to a point sufficient to reverse the clinical phenotype. Together with inhibitors of heparan sulfate synthesis, pharmacological chaperones could potentially reduce storage of this polymer in the central nervous system to levels sufficient to stop neuronal death and reverse inflammation.

## MATERIALS AND METHODS

The current study was conducted with ethics approval from the review board of CHU Ste-Justine, University of Montreal.

### Generation of expression constructs and site-directed mutagenesis

The wild-type HGSNAT-TAP plasmid was obtained by subcloning the HGSNAT 1992 bp coding sequence into pCTAP vector (Stratagene). Briefly, a 3' part of human pCMV-Script construct [11] missing the stop codon was amplified by PCR using primers 76-Cla-F 5'-TTG CTC TTA TAC TCA TGG TCT TTG TCA-3' and TM76-R4 5'-ATA TGT CGA CGA GCC ATC CGA TTT TCC-3' and then used to replace Cla I - Sal I segment of the same construct prior to subcloning into pCTAP using Hind III – Sal I sites. Missense mutants were constructed using QuikChange Lightning kit (Stratagene) with HGSNAT-TAP as a matrix (see Supplementary Table 1 for primers, Appendix IV). All primers were designed using QuikChange Primer Design Program (http://www.stratagene.com/qcprimerdesign). For all constructs the coding sequence included an extra 84 base pairs on their 5' end (encoding 28 amino acids) as the sequence first described for HGSNAT gene [11], but for clarity the nomenclature used reflects a unified numbering system based on the sequence of GenBank entries NM_152419.2 and NG_009552.1 as in Fan et al. [12], Ruijter et al. [19], Fedele et al. [22] and Feldhammer et al. [13].

**Cell culture and transfection**

Skin fibroblast lines of MPS IIIC patients, obtained with a written informed consent [11], and COS-7 cells (ATCC) were cultured in Eagles's minimal essential medium supplemented with 10% (v/v) fetal calf serum (Wisent). Transfections were carried out using polyethyleneimine 25,000 (PEI, Polysciences inc.) or Lipofectamine LTX (Invitrogen). PEI was dissolved in water at 1 mg/ml and pH was adjusted to 6.8 with HCl. For transfection, a mixture of 2 μg DNA and 8 μg PEI in 0.4 ml of serum free medium was incubated for 15 min at room temperature and added to ~70% confluent cells growing in a 10 cm dish. The transfection media was replaced with growth media 18 h later. Lipofectamine LTX was used as described in the manufacturer's protocol. Patients (N273K and c.1726+1G>A) and control fibroblasts immortalized by transfection with retroviral vectors expressing the type 16 human papilloma virus E7 gene and the catalytic component of human telomerase were described before [25].

**Glucosamine-mediated refolding**

Patient and control skin fibroblasts were cultured as described above in growth media supplemented with various concentrations of D-(+)-glucosamine hydrochloride (Sigma G4875). Medium was replaced every day and at specified time cells were harvested and assayed for N-acetyltransferase activity and β-hexosaminidase activity as described below.

**Enzyme assays**

N-acetyltransferase enzymatic activity was measured using fluorogenic substrate, 4-methylumbelliferyl *β*-D-glucosaminide (4MU-βGlcN, Moscerdam, Rotterdam, The Netherlands) as previously described by He et al. [23]. The reaction mixture containing 5 µl of cell homogenate, 5 µl of 6 mM acetyl-CoA and 5 µl of 3 mM 4MU-βGlcN in McIlvain buffer (100 mM sodium citrate, 200 mM sodium phosphate, pH 5.7) was incubated at 37ºC for 3-18 h. The reaction was terminated by adding 1.98 ml of 0.5 M $Na_2CO_3$/$NaHCO_3$, pH 10.7, and fluorescence was measured and used to calculate the specific activity. Lysosomal β-hexosaminidase activity was measured as previously described [17]. Protein concentration was measured according to the method of Bradford [24] using a commercially available reagent (BioRad).

**Deglycosylation of HGSNAT.**

To remove N-linked glycans from HGSNAT cell homogenates were treated with recombinant endoglycosidase H, or peptide: N-Glycosidase F (Endo H; PNGase F New England Biolabs). Briefly, the mix consisted of 10 µg of cell homogenate in 25 mM sodium phosphate buffer, pH 7.5, to which 1 µl (500 U) of concentrated Endo H or PNGase F was added before incubating at 37ºC overnight and analysis by Western blot.

To inhibit glycosylation of newly-synthesized HGSNAT COS-7 cells transfected with plasmids coding for the wild-type enzyme and the C76F or P237Q variants were treated with 1 µg/ml of tunicamycin (Sigma) added 5 h after

transfection and allowed to express the recombinant protein for 48 h in the presence of drug. Media and drug were changed after 24 h and the homogenates were analyzed by Western blot as described below.

**Western blotting**

Cell homogenates were sonicated and boiled in LDS sample buffer (Invitrogen) in the presence of 25 mM DTT. Proteins were resolved by SDS-polyacrylamide gel electrophoresis using NuPAGE 4-12% Bis-Tris gels (Invitrogen) and electrotransferred to PVDF membrane. Detection of TAP-tagged N-acetyltransferase protein was performed using anti-calmodulin binding peptide epitope tag (CBP) rabbit antibodies (Immunology Consultants Laboratory, dilution 1:30,000) and the Amersham ECL Western Blotting Detection Reagents (GE Healthcare) in accordance with the manufacturer's protocol.

**Analysis of the wild-type recombinant HGSNAT by the anion-exchange FPLC.**

COS-7 cells were harvested 42 h after transfection with the wild-type HGSNAT plasmid and suspended in lysis buffer (10 mM Tris-HCl, pH 7.5, 0.1% NP-40, 1 mM PMSF and Sigma P8340 protease inhibitor cocktail at 10 µl per 1 ml of cell suspension). The homogenate was sonicated, gently shaked at 4 ºC for 2 h and centrifuged at 13,000 rpm for 30 min. One ml of the supernatant containing 8 mg of total protein was applied to an ion-exhange Mono Q HR 5/5 column equilibrated with 10 mM Tris buffer, pH 8.2. The column was washed

with 4 ml of the same buffer and then eluted using a 20 mL gradient of NaCl (0-0.5 M) at a flow rate of 0.5 ml/min. One ml fractions were collected and assayed for N-acetyltransferase activity. Thirty µl aliquots from each fraction were analyzed by SDS-PAGE and Western blot using anti-calmodulin binding peptide epitope tag rabbit antibodies as described above.

**Confocal immunofluorescence microscopy**

Immortalized control human skin fibroblasts were transfected with HGSNAT-TAP or plasmids coding for the HGSNAT mutants using Lipofectamine LTX (Invitrogen) as described in the manufacturer's protocol. Forty two hours post-transfection cells were incubated for 1 hour with 1 µM Lysotracker Red DND-99 (Invitrogen) and then washed with ice-cold PBS. Cells were fixed with 4% paraformaldehyde, 4% sucrose in PBS for 5 min, and then rinsed 3 times with PBS. Cells were permeabilized by 0.25% Triton X-100 for 10 min and blocked for 1 h in 3% horse serum and 0.1% Triton X-100. Cells were either co-stained with rabbit anti-CBP (Immunology Consultants Laboratory; 1:400) and mouse monoclonal anti-calnexin (Millipore; 1:250) antibodies in 3% horse serum, with anti-CBP antibodies and mouse monoclonal antibodies against human LAMP-2 (Developmental Studies Hybridoma Bank; 1:150), or with anti-CBP antibodies and Lysotracker Red DND-99. Cells were then counterstained with Oregon Green 488-conjugated anti-rabbit IgG antibodies or Texas-Red-conjugated goat anti-mouse antibodies (Molecular Probes; 1:1000). Slides were studied on a Zeiss LSM510 inverted confocal microscope (Zeiss). Images were

processed using the LSM image browser software (Zeiss) and Photoshop (Adobe).

**Accession numbers**

GenBank entries NM_152419.2 and NG_009552.1

## ACKNOWLEDGEMENTS

The authors thank the patients, and their families for participating in our study. We also thank Maryssa Canuel for the helpful advice, Mila Ashmarina for critical reading and Eva Lacroix for the help in preparation of the manuscript.

## REFERENCES


1. Meikle PJ, Hopwood JJ, Clague AE, Carey WF (1999) Prevalence of lysosomal storage disorders. JAMA 281: 249-254.
2. Neufeld EF, Muenzer J (2001) The mucopolysaccharidoses *in* Scriver CR, Beaudet AL, Sly WS, Valle D, Editors. The Metabolic and Molecular Basis of Inherited Disease. McGraw-Hill, New York. pp. 3421-3452.
3. Bartsocas C, Grobe H, van de Kamp JJP, von Figura K, Kresse H, et al. (1979) Sanfilippo type C disease: clinical findings in four patients with a new variant of mucopolysaccharidosis III. Eur J Pediatr 130: 251-258.
4. Valstar MJ, Ruijter, GJG, van Diggelen OP, Poorthuis BJ, Wijburg FA (2008) Sanfilippo syndrome: A mini-review. J Inherit Metab Dis. 31: 240-252.
5. Berger-Plantinga EG, Vanneste JA, Groener JE, van Schooneveld MJ (2004) Adult-onset dementia and retinitis pigmentosa due to mucopolysaccharidosis III-C in two sisters. J Neurol. 251: 479-481.
6. Pinto R, Caseiro C, Lemos M, Lopes L, Fontes A, et al. (2004) Prevalence of lysosomal storage diseases in Portugal. Eur J Hum Genet 12: 87-92.
7. Poorthuis BJ, Wevers RA, Kleijer WJ, Groener JE, de Jong JG, et al. (1999) The frequency of lysosomal storage diseases in The Netherlands. Hum Genet 105: 151-156.
8. Klein U, Kresse H, von Figura K (1978) Sanfilippo syndrome type C: deficiency of acetyl-CoA: α-glucosaminide N-acetyltransferase in skin fibroblasts. Proc Natl Acad Sci U S A 75: 5185-5189.
9. Klein U, van de Kamp JJP, von Figura K, Pohlmann R (1981) Sanfilippo syndrome type C: assay for acetyl-CoA: α-glucosaminide N-acetyltransferase in leukocytes for detection of homozygous and heterozygous individuals. Clin Genet 20: 55-59.
10. Bame KJ, Rome LH (1986b) Genetic evidence for transmembrane acetylation by lysosomes. Science 233: 1087-1089.

11. Hřebíček M, Mrázová L, Seyrantepe V, Durand S, Roslin NM, et al. (2006) Mutations in TMEM76 cause mucopolysaccharidosis IIIC (Sanfilippo C syndrome). Am J Hum Genet 79: 807-819.
12. Fan X, Zhang H, Zhang S, Bagshaw RD, Tropak MB, et al. (2006) Identification of the gene encoding the enzyme deficient in mucopolysaccharidosis IIIC (Sanfilippo disease type C). Am J Hum Genet 79: 738-744.
13. Feldhammer M, Durand S, Mrázová L, Boucher RM, Laframboise R, et al. (2009) Sanfilippo Syndrome Type C: Mutation spectrum in the heparan sulfate acetyl-CoA: A-glucosaminide N-acetyltransferase (HGSNAT) gene. Hum Mut 30(6): 918-925.
14. Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M et al. (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* 17(10): 1030-1032.
15. Puig O, Caspary F, Rigaut G, Rutz B, Bouveret E, et al. (2001) The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods* 24(3): 218-219.
16. Maley F, Trimble RB, Tarentine AL, Plummer TH Jr. (1989) Characterization of glycoproteins and their associated oligosaccharides throgh the use of endoglycosidases. Analytical biochemistry 180(2): 195-204.
17. Ausseil J, Landry K, Seyrantepe V, Trudel S, Mazur A, et al. (2006) An acetylated 120-kDa lysosomal transmembrane protein is absent from mucopolysaccharidosis IIIC fibroblasts: a candidate molecule for MPS IIIC. Mol Genet Metab 87: 22–31.
18. Gregersen N, Bross P, Andrese BS, Pedersen CB, Corydon TJ, et al.(2001) The role of chaperone-assisted folding and quality control in inborn errors of metabolism: protein folding disorders. J Inherit Metab Dis. 24(2): 189-212.
19. Ruijter GJ, Valstar MJ, van de Kamp JM, van der Helm RM, Durand S et al. (2008) Clinical and genetic spectrum of Sanfilippo type C (MPS IIIC) disease in The Netherlands. Mol Genet Metab 93: 104–111.

20. Fan J-Q (2003) A contradictory treatment for lysosomal storage disorders: inhibitors enhance mutant enzyme activity. Trends Pharmacol Sci 24: 355-360.
21. Desnick RJ (2004) Enzyme replacement and enhancement therapies for lysosomal diseases. J Inherit Metab Dis 27(3): 385-410.
22. Fedele AO, Filocamo M, Di Rocco M, Sersale G, Lübke T, et al. (2007) Mutational analysis of the HGSNAT gene in Italian patients with mucopolysaccharidosis IIIC (Sanfilippo C syndrome). Hum Mutat 28: 523.
23. He W, Voznyi YaV, Huijmans JG, Geilen GC, Karpova EA, et al. (1994) Prenatal diagnosis of Sanfilippo disease type C using a simple fluorometric enzyme assay. Prenat Diagn. 14(1): 17-22.
24. Bradford MM (1976) A rapid and sensitive method for the quatitation of microgram quantities of protein utilizing the principle of protein dye binding. Anal Biochem 72: 248-254
25. Seyrantepe V, Tihy F, Pshezhetsky AV (2006) The microcell-mediated transfer of human chromosome 8 restores the deficient N-acetylytransferase activity in skin fibroblasts of Mucopolysaccharidosis type IIIC patients. Hum Genet. 120(2): 293-296.
26. Spyropoulos IC, Liakopoulos TD, Bagos PG, Hamodrakas SJ (2004) TMRPres2D: high quality visual representation of transmembrane protein models. Bioinformatics 20: 3258-3260.

## FIGURE LEGENDS

**Figure 1. Enzymatic activity and expression of HGSNAT mutants.**

COS-7 cells were harvested 42 h after transfection with HGSNAT-TAP plasmids bearing missense mutations. Cell homogenates were (A) assayed for N-acetyltransferase activity and (B) analyzed by Western blot using anti-CBP antibody as described in Materials and Methods.

**A.** N-acetyltransferase activity is shown as a fraction of the activity measured in the cells transfected with the wild-type HGSNAT-TAP plasmid. Values represent means ± S.D. of three independent experiments.

**B.** Blot shows a representative image of triplicate experiments.

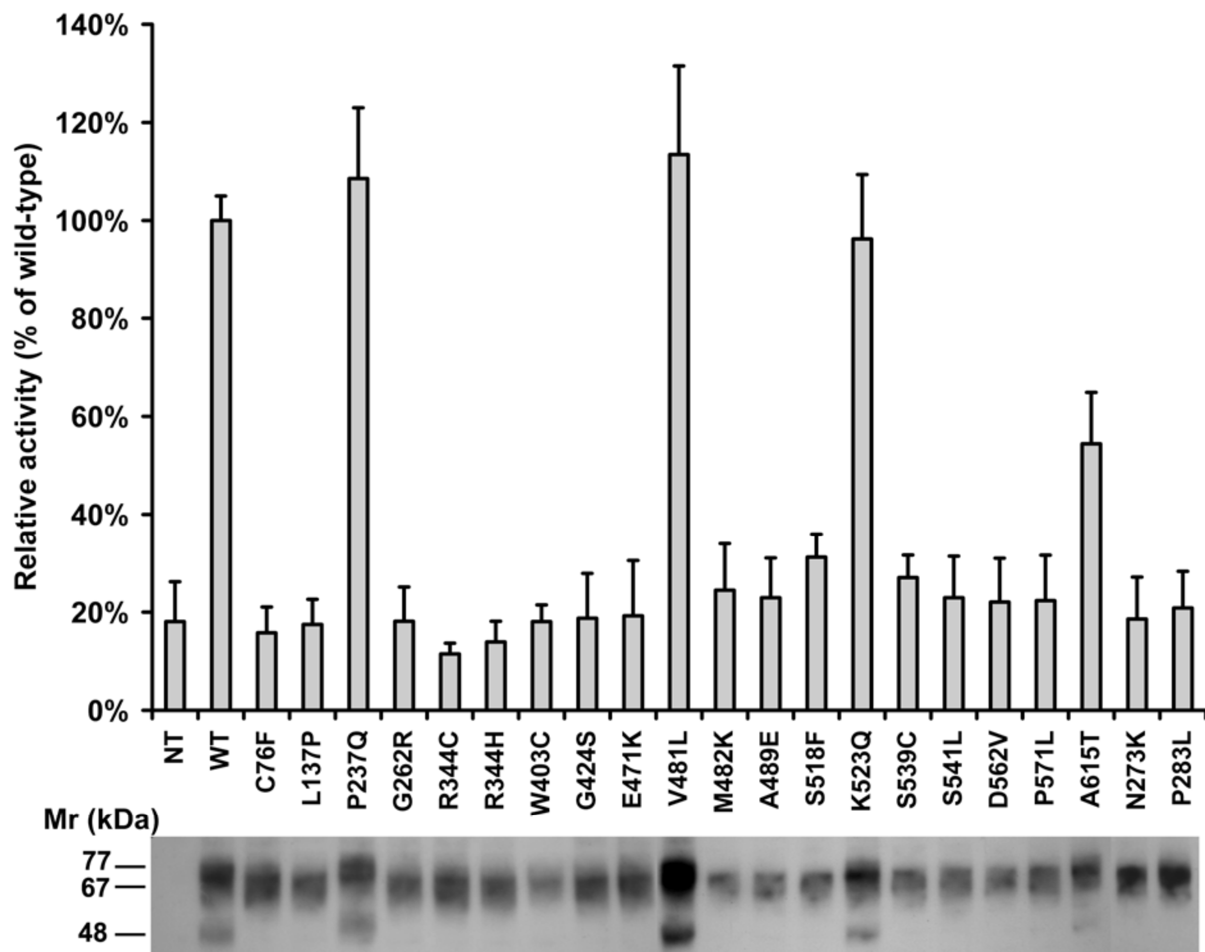



Figure 1

**Figure 2. Deglycosylation of HGSNAT by endoglycosidase H, PNGase F and tunicamycin.**

COS-7 cells expressing either the wild-type HGSNAT or the protein containing C76F, P237Q, G262R or V481L variants were harvested 42 h post-transfection and their homogenates were treated overnight with endoglycosidase H **(A)** or PNGase F **(B)**. **(C)** COS-7 cells expressing wild-type HGSNAT and C76F or P237Q variants were cultured for 48 h in the presence or absence of 1 μg/ml of tunicamycin added to the culture medium 5 h after the transfection. The treated and control homogenates were analyzed by Western blot using anti-CBP antibodies as described in Materials and Methods.

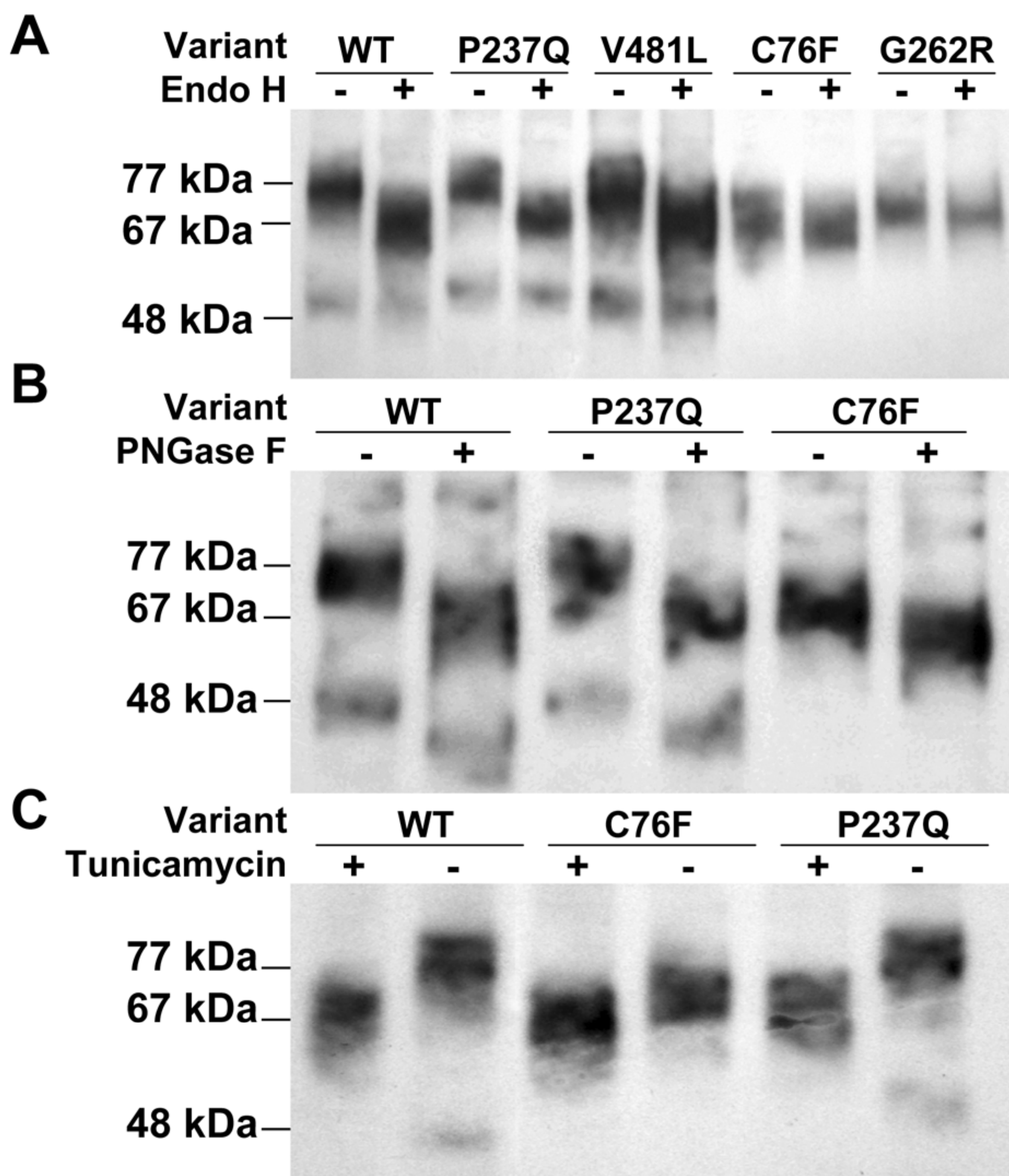



Figure 2

**Figure 3. Analysis of the wild-type recombinant HGSNAT by the anion-exchange FPLC.**

COS-7 cells were harvested 42 h after transfection with the wild-type HGSNAT plasmid, solubilized in a buffer containing 0.1% NP-40, applied to an ion-exhange Mono Q HR 5/5 column and eluted by 0-0.5 M NaCl gradient as described in Materials and Methods. Graph shows N-acetyltransferase activity (nmol/hr ml) in the collected fractions. Dashed line represents the NaCl gradient. An aliquot from each fraction was analyzed by Western blot using anti-CBP antibodies (inset).

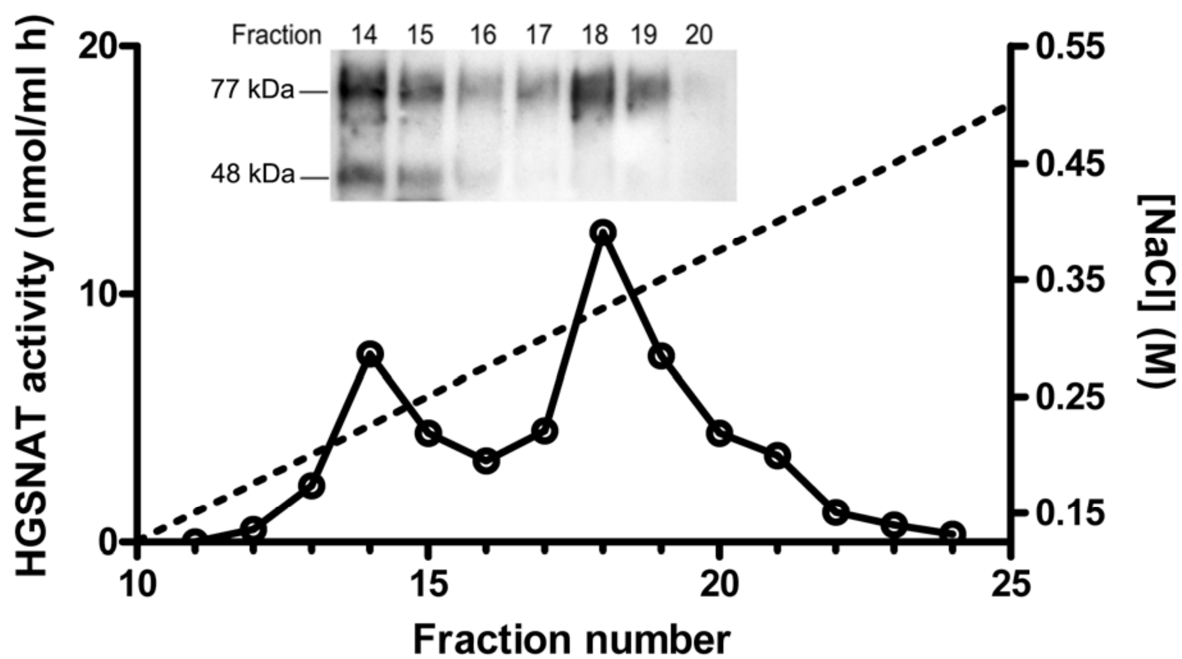



Figure 3

**Figure 4. Localization of HGSNAT mutants expressed in cultured human skin fibroblasts by immunofluorescence microscopy.**

The cells transfected with wild-type or mutant HGSNAT-TAP constructs as indicated were fixed and stained with either mouse monoclonal anti-LAMP-2 antibodies, Lysotracker Red DND-99 or mouse monoclonal anti-calnexin antibodies (red) and rabbit polyclonal anti-CBP antibodies (green) as indicated. Slides were studied on a Zeiss LSM510 inverted confocal microscope. Magnification 630x. Panels show representative images illustrating co-localization of anti-CBP antibodies (green) and lysosomal and ER markers (red) for the wild-type HGSNAT, active enzyme containing P237Q polymorphism and inactive L137P and P571L mutants. From 10 to 15 cells all showing similar localization patterns were studied for each variant. See Supplementary figure 2 for the data on other mutants, Appendix II.

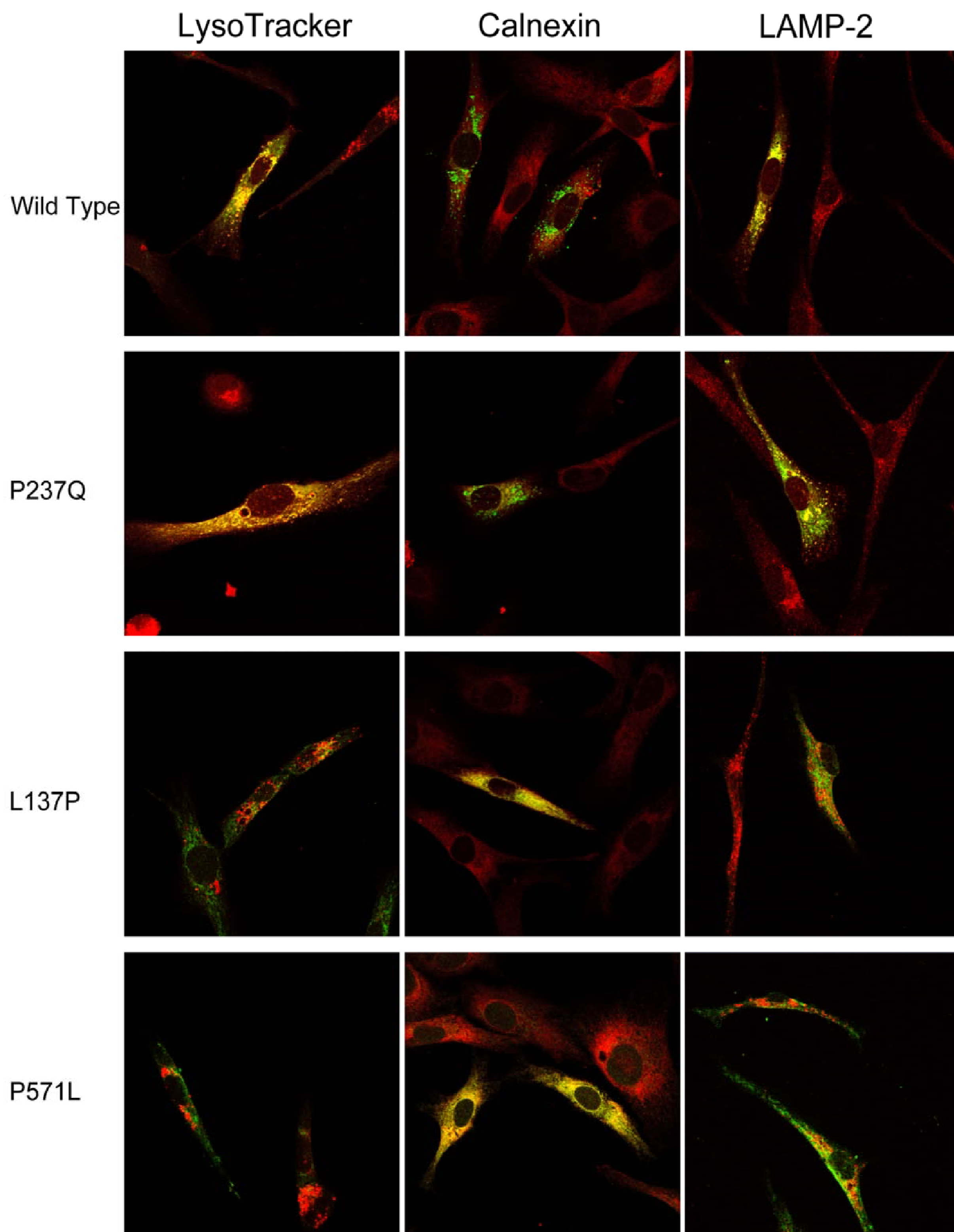



Figure 4

**Figure 5. Partial refolding of HGSNAT mutants by glucosamine.**
**A.** Fifty percent confluent immortalized skin fibroblasts from a MPS IIIC patient homozygous for N273K mutation [25] were cultured in the presence or absence of 14 mM D-(+)-glucosamine hydrochloride. Medium was replaced every day and at the indicated time intervals cells were harvested and assayed for the N-acetyltransferase activity. N-acetyltransferase activity is shown as a fraction of that measured in non-treated cells after 24 h of culturing. Data show mean values and standard error of 2 independent experiments. ** Significantly different ($p<0.01$) from non-treated cells according to repeated measurements ANOVA.
**B.** Same cells were cultured in the presence of increasing glucosamine concentrations (0-14 mM) for 5 days, harvested and assayed for N-acetyltransferase activity or β-hexosaminidase activity. Data show mean values and standard error of 2 independent experiments.
**C.** Primary skin fibroblasts of the MPS IIIC patients carrying the missense HGSNAT mutations: L137P/S518F **(P1)**, P283L/R344C **(P3)**, S518F/S518F **(P4)**, N273K/N273K **(P6)**, R344C/R344C **(P7),** S518F/S518F **(P8)** and E471K/D562V **(P9)** or a missense mutation in combination with a splice site (S541L/c.234+1G>A; **P2**) or nonsense (R344H/R384X, **P5**) mutation were cultured for 2 (P1, P2, P5, P6, P8, P9) or 3 (P3, P4, P7) days in the absence (open bars) or presence (filled bars) of 7 mM (P3-P9) or 14 mM (P1, P2) glucosamine, harvested and assayed for N-acetyltransferase activity. Data show mean values and standard error of 2 independent experiments. The residual N-acetyltransferase activity detectable in untreated cells most likely represents background chemical or enzymatic reactions occurring with the substrate in the presence of cell homogenates.
Significantly (*, $p<0.05$; **, $p<0.01$; ***, $p<0.001$) different from non-treated cells according to non-parametric t-test.

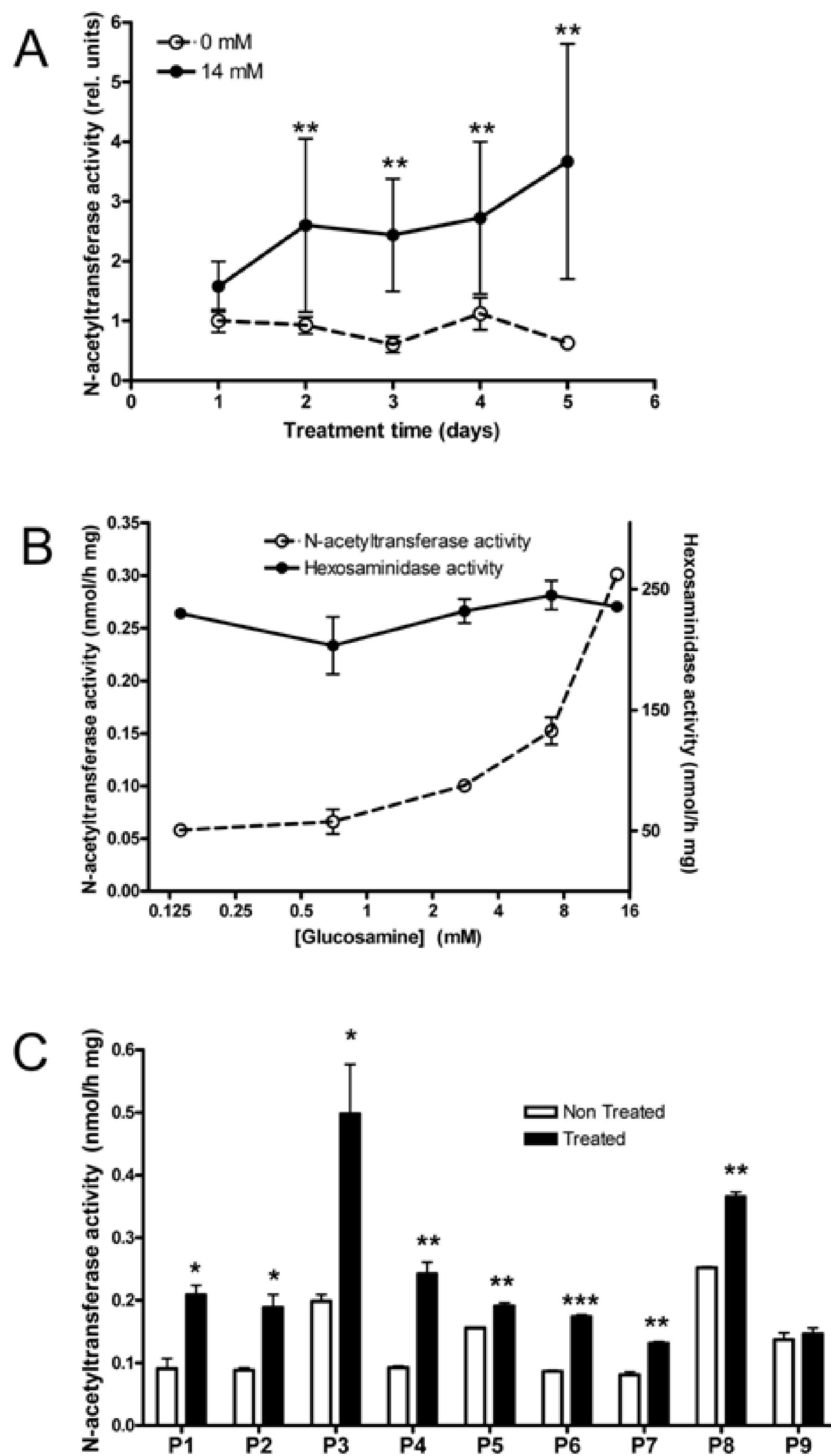



Figure 5

**Figure 6. Distribution of missense mutations in HGSNAT protein.**

Visual representation of HGSNAT membrane topology was created using the TMRPres2D software [26]. The deduced amino acid sequence of HGSNAT predicts 11 transmembrane domains and five potential N-glycosylation sites oriented towards the lysosomal lumen (shown in blue). Mutations that result in production of misfolded proteins are shown in red. Polymorphisms are shown in green. Figure was adapted from our previous work [13].

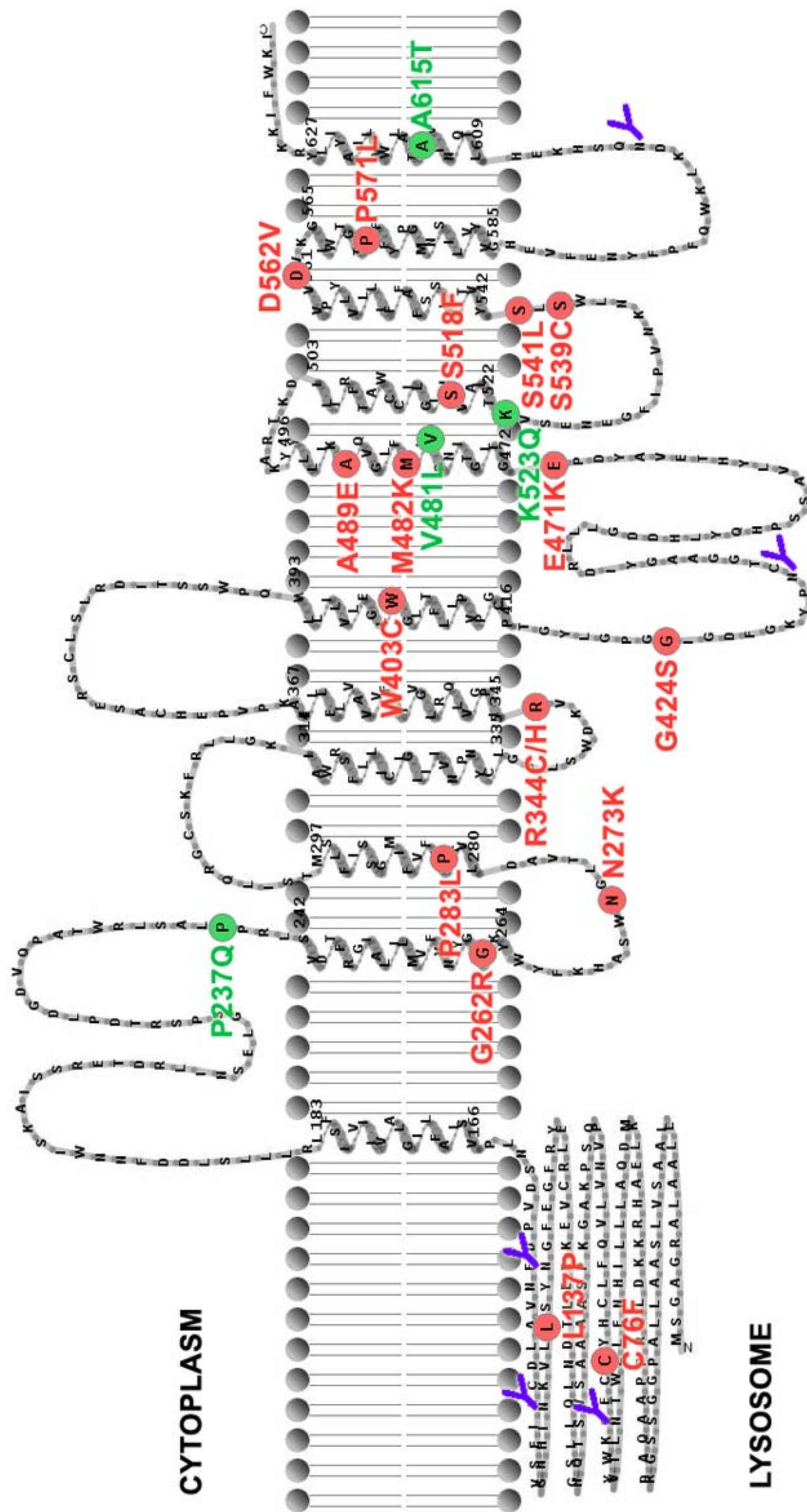



Figure 6

# CHAPTER 4

# GENERAL DISCUSSION
# AND CONCLUSIONS

# General Discussion

Our studies represent the first major characterization of the molecular defects of HGSNAT in MPS IIIC as well as the first comprehensive review of patient mutations. We have identified 10 novel mutations present in diverse patient populations and speculated on their contributions to pathogenesis. We further showed enzyme misfolding to be the underlying molecular defect in patients with missense mutations explaining their severe disease progression. Additionally we identified and utilized the pharmalogical chaperone glucosamine to rescue mutant enzyme activity in these patient fibroblasts.

To date 54 HGSNAT mutations have been reported in the literature including 13 splice-site mutations, 11 insertions and deletions, 8 nonsense, 18 missense and 4 polymorphisms. The mutations are found throughout the 18 exons and affect 8 of the 11 transmembrane domains of the protein as well as its cytosolic and luminal domain. Interestingly, a majority of the missense mutations are located within the predicted transmembrane domains or on adjacent residues within the lumen. These mutations are likely to affect the native folding of the enzyme by destabilizing the transmembrane helix as well as potentially interfering with sites important for binding of AcCoA, its transport through lysosomal membrane or areas important for the catalytic activity of the enzyme.

Despite the relative heterogeneity seen in patient mutations, there are several mutations with higher frequencies suggesting possible founder effects. Two of the more frequent mutations p.R344C and p.S518F are found primarily in Dutch patients. These mutations account for 22% and 29.3% of the probands of

Dutch origin respectively, representing potential founder mutations [81]. Founder mutations are generally frequent among Dutch patients due to a high genetic homogeneity among this population [83]. The p.S518F mutation was found exclusively in Dutch patients and one patient from Germany. Interestingly, 3 of 11 Italian patients were homozygous for the mutation c.852-1G>A, all of whom had origins in the south of Italy [80]. Another frequent mutation was c.525dupT reported in 5 of 6 alleles of Portuguese patients, suggesting a potential founder mutation in the Portuguese population. We further reported two different Portuguese patients each homozygous for the c.525dupT mutation supporting the idea of a founder mutation in the Portuguese patient population

Other reported mutations: splicing mutations, insertions, deletions and nonsense mutations will all predictably result in mRNA which is either improperly spliced, or contains a PTC resulting in rapid degradation of the transcript and therefore complete absence of the HGSNAT enzyme.

As mentioned, a majority of MPS IIIC patients experience a severe clinical progression due to an almost complete absence of HGSNAT activity. Exceptionally, the 2 mutations p.G262R and p.S539C likely cause a late onset form of MPSIII as a combination of these mutations was only found in two sisters with the attenuated form of the disease.

In order to understand the reason behind the severe disease progression seen in patients with missense mutations; we have conducted the initial biochemical characterization of the HGSNAT enzyme, as well as mutational analysis of all known missense mutations. We studied the activity and biogenesis

of all 21 previously reported missense mutations by expressing recombinant forms in various cell types.

We determined that HGSNAT missense mutants with the exception of the 4 polymorphisms are inactive, misfolded and retained in the ER where they are abnormally glycosylated. As mentioned, many of the missense mutations affected residues within or adjacent to the transmembrane domains of the enzyme, therefore the introduction of amino acids with different charge, size or hydrophobicity can destabilize the helix and have drastic effects on protein folding.

Once we determined MPS IIIC to be a folding disorder we screened potential chaperone molecules for their ability to stabilize the active conformation of HGSNAT. We identified glucosamine as an active site-specific chaperone (ASSC) of HGSNAT and showed that it could rescue the active conformation in 8 of 9 patient cell lines tested. We therefore believe chaperone therapy to be a good treatment option for MPS IIIC as a majority of patients are affected with at least one missense mutation interfering with proper enzyme folding which could be partially rescued by the ASSC glucosamine.

Many questions about the biochemistry of HGSNAT remain unanswered. It is still not clear exactly how the enzyme exerts its N-acetyltransferase function, what residues are important for its catalytic activity and where the substrates glucosamine and AcCoA bind. We also do not know for sure in what oligomeric form the enzyme is present in vivo, or even how HGSNAT is targeted to the lysosome. Answers to these questions will provide a better understanding of the

disease mechanism and will allow for the development of better therapies to help patients suffering from this and similar disorders.

# Conclusions

1. We have identified 10 novel mutations in HGSNAT including 3 splicing mutations, 2 deletions, 1 insertion, 2 nonsense mutations and 2 missense mutations identified in 14 MPS IIIC patients. Molecular and clinical information on these as well as on all previously reported mutations was summarized in a comprehensive review.

2. We determined that 4 previously reported amino acid changes in HGSNAT were in fact polymorphisms which had no effect on enzyme activity, targeting or glycosylation and therefore were not clinically significant.

3. We expressed in COS 7 cells and human fibroblasts HGSNAT mutants containing all 17 reported amino acid changes and showed that they all lacked enzymatic activity, proper glycosylation and proteolytic cleavage occurring with the wild type enzyme in the lysosome.

4. We further showed that HGSNAT missense mutants in contrast to wild type enzyme and enzyme containing polymorphisms were not properly trafficked to the lysosome, but were retained in the ER.

5. We showed that a competitive inhibitor of HGSNAT, glucosamine can act as a molecular chaperone, rescuing HGSNAT activity in 8 of 9 patient fibroblast lines tested.

## Bibliography

[1] P. Saftig, J. Klumperman, Lysosome biogenesis and lysosomal membrane proteins: trafficking meets function, Nat Rev Mol Cell Biol 10 (2009) 623-635.

[2] J.S. Bonifacino, L.M. Traub, Signals for sorting of transmembrane proteins to endosomes and lysosomes, Annu Rev Biochem 72 (2003) 395-447.

[3] R.D. Bagshaw, D.J. Mahuran, J.W. Callahan, Lysosomal membrane proteomics and biogenesis of lysosomes, Mol Neurobiol 32 (2005) 27-41.

[4] N. Andrejewski, E.L. Punnonen, G. Guhde, Y. Tanaka, R. Lullmann-Rauch, D. Hartmann, K. von Figura, P. Saftig, Normal lysosomal morphology and function in LAMP-1-deficient mice, J Biol Chem 274 (1999) 12692-12701.

[5] K.K. Huynh, E.L. Eskelinen, C.C. Scott, A. Malevanets, P. Saftig, S. Grinstein, LAMP proteins are required for fusion of lysosomes with phagosomes, EMBO J 26 (2007) 313-324.

[6] A.H. Erickson, G.E. Conner, G. Blobel, Biosynthesis of a lysosomal enzyme. Partial structure of two transient and functionally distinct NH2-terminal sequences in cathepsin D, J Biol Chem 256 (1981) 11224-11231.

[7] T. Braulke, J.S. Bonifacino, Sorting of lysosomal proteins, Biochim Biophys Acta 1793 (2009) 605-614.

[8] P.Y. Tong, W. Gregory, S. Kornfeld, Ligand interactions of the cation-independent mannose 6-phosphate receptor. The stoichiometry of mannose 6-phosphate binding, J Biol Chem 264 (1989) 7962-7969.

[9] P.Y. Tong, S. Kornfeld, Ligand interactions of the cation-dependent mannose 6-phosphate receptor. Comparison with the cation-independent mannose 6-phosphate receptor, J Biol Chem 264 (1989) 7970-7975.

[10] L. Little, M. Alcouloumre, A.M. Drotar, S. Herman, R. Robertson, R.Y. Yeh, A.L. Miller, Properties of N-acetylglucosamine 1-

phosphotransferase from human lymphoblasts, Biochem J 248 (1987) 151-159.

[11] L.E. Little, O.T. Mueller, N.K. Honey, T.B. Shows, A.L. Miller, Heterogeneity of N-acetylglucosamine 1-phosphotransferase within mucolipidosis III, J Biol Chem 261 (1986) 733-738.

[12] M.L. Reitman, S. Kornfeld, Lysosomal enzyme targeting. N-Acetylglucosaminylphosphotransferase selectively phosphorylates native lysosomal enzymes, J Biol Chem 256 (1981) 11977-11980.

[13] M.L. Reitman, S. Kornfeld, UDP-N-acetylglucosamine:glycoprotein N-acetylglucosamine-1-phosphotransferase. Proposed enzyme for the phosphorylation of the high mannose oligosaccharide units of lysosomal enzymes, J Biol Chem 256 (1981) 4275-4281.

[14] A. Waheed, A. Hasilik, K. von Figura, UDP-N-acetylglucosamine:lysosomal enzyme precursor N-acetylglucosamine-1-phosphotransferase. Partial purification and characterization of the rat liver Golgi enzyme, J Biol Chem 257 (1982) 12322-12331.

[15] H.B. Lodish, A. Zipursky, SL. Masudaira, P. Baltimore, D. Darnell, J., Molecular Cell Biology, Textbook (1999).

[16] A. Tsuji, K. Omura, Y. Suzuki, I-cell disease: evidence for a mannose 6-phosphate independent pathway for translocation of lysosomal enzymes in lymphoblastoid cells, Clin Chim Acta 176 (1988) 115-121.

[17] A.J. Hassan, J. Zeng, X. Ni, C.R. Morales, The trafficking of prosaposin (SGP-1) and GM2AP to the lysosomes of TM4 Sertoli cells is mediated by sortilin and monomeric adaptor proteins, Mol Reprod Dev 68 (2004) 476-483.

[18] S. Lefrancois, J. Zeng, A.J. Hassan, M. Canuel, C.R. Morales, The lysosomal trafficking of sphingolipid activator proteins (SAPs) is mediated by sortilin, EMBO J 22 (2003) 6430-6437.

[19] S. Lefrancois, M. Canuel, J. Zeng, C.R. Morales, Inactivation of sortilin (a novel lysosomal sorting receptor) by dominant negative competition and RNA interference, Biol Proced Online 7 (2005) 17-25.

[20] J.M. Aerts, J. Heikoop, S. van Weely, W.E. Donker-Koopman, J.A. Barranger, J.M. Tager, A.W. Schram, Characterization of glucocerebrosidase in peripheral blood cells and cultured blastoid cells, Exp Cell Res 177 (1988) 391-398.

[21] J.G. Barriocanal, J.S. Bonifacino, L. Yuan, I.V. Sandoval, Biosynthesis, glycosylation, movement through the Golgi system, and transport to lysosomes by an N-linked carbohydrate-independent mechanism of three lysosomal integral membrane proteins, J Biol Chem 261 (1986) 16755-16763.

[22] C. Peters, K. von Figura, Biogenesis of lysosomal membranes, FEBS Lett 346 (1994) 108-114.

[23] F.G. Guarnieri, L.M. Arterburn, M.B. Penno, Y. Cha, J.T. August, The motif Tyr-X-X-hydrophobic residue mediates lysosomal membrane targeting of lysosome-associated membrane protein 1, J Biol Chem 268 (1993) 1941-1946.

[24] B.M. Pearse, C.J. Smith, D.J. Owen, Clathrin coat construction in endocytosis, Curr Opin Struct Biol 10 (2000) 220-228.

[25] D. Reczek, M. Schwake, J. Schroder, H. Hughes, J. Blanz, X. Jin, W. Brondyk, S. Van Patten, T. Edmunds, P. Saftig, LIMP-2 is a receptor for lysosomal mannose-6-phosphate-independent targeting of beta-glucocerebrosidase, Cell 131 (2007) 770-783.

[26] B. Chapuy, R. Tikkanen, C. Muhlhausen, D. Wenzel, K. von Figura, S. Honing, AP-1 and AP-3 mediate sorting of melanosomal and lysosomal membrane proteins into distinct post-Golgi trafficking pathways, Traffic 9 (2008) 1157-1172.

[27] H.G. Hers, The role of lysosomes in the pathogeny of storage diseases, Biochimie 54 (1972) 753-757.

[28] B. Winchester, A. Vellodi, E. Young, The molecular basis of lysosomal storage diseases and their treatment, Biochem Soc Trans 28 (2000) 150-154.

[29] P. Bifsha, K. Landry, L. Ashmarina, S. Durand, V. Seyrantepe, S. Trudel, C. Quiniou, S. Chemtob, Y. Xu, R.A. Gravel, R. Sladek, A.V. Pshezhetsky, Altered gene expression in cells from patients with lysosomal storage disorders suggests impairment of the ubiquitin pathway, Cell Death Differ 14 (2007) 511-523.

[30] C. Settembre, A. Fraldi, L. Jahreiss, C. Spampanato, C. Venturi, D. Medina, R. de Pablo, C. Tacchetti, D.C. Rubinsztein, A. Ballabio, A block of autophagy in lysosomal storage disorders, Hum Mol Genet 17 (2008) 119-129.

[31] O. Staretz-Chacham, T.C. Lang, M.E. LaMarca, D. Krasnewich, E. Sidransky, Lysosomal storage disorders in the newborn, Pediatrics 123 (2009) 1191-1207.

[32] P.J. Meikle, J.J. Hopwood, Lysosomal storage disorders: emerging therapeutic options require early diagnosis, Eur J Pediatr 162 Suppl 1 (2003) S34-37.

[33] P.J. Meikle, J.J. Hopwood, A.E. Clague, W.F. Carey, Prevalence of lysosomal storage disorders, JAMA 281 (1999) 249-254.

[34] G.A. Grabowski, R.J. Hopkin, Enzyme therapy for lysosomal storage disease: principles, practice, and prospects, Annu Rev Genomics Hum Genet 4 (2003) 403-436.

[35] M.B. Tropak, S.P. Reid, M. Guiral, S.G. Withers, D. Mahuran, Pharmacological enhancement of beta-hexosaminidase activity in fibroblasts from adult Tay-Sachs and Sandhoff Patients, J Biol Chem 279 (2004) 13478-13487.

[36] Y. Matoth, S. Chazan, A. Cnaan, I. Gelernter, C. Klibansky, Frequency of carriers of chronic (type I) Gaucher disease in Ashkenazi Jews, Am J Med Genet 27 (1987) 561-565.

[37] G.M. Petersen, J.I. Rotter, R.M. Cantor, L.L. Field, S. Greenwald, J.S. Lim, C. Roy, V. Schoenfeld, J.A. Lowden, M.M. Kaback, The Tay-Sachs disease gene in North American Jewish populations: geographic variations and origin, Am J Hum Genet 35 (1983) 1258-1269.

[38] J.C. Fratantoni, C.W. Hall, E.F. Neufeld, Hurler and Hunter syndromes: mutual correction of the defect in cultured fibroblasts, Science 162 (1968) 570-572.

[39] J.R. Hobbs, K. Hugh-Jones, A.J. Barrett, N. Byrom, D. Chambers, K. Henry, D.C. James, C.F. Lucas, T.R. Rogers, P.F. Benson, L.R. Tansley, A.D. Patrick, J. Mossman, E.P. Young, Reversal of clinical features of Hurler's disease and biochemical improvement after treatment by bone-marrow transplantation, Lancet 2 (1981) 709-712.

[40] C. Peters, C.G. Steward, P. National Marrow Donor, R. International Bone Marrow Transplant, E.B.M.T.G. Working Party on Inborn Errors, Hematopoietic cell transplantation for inherited metabolic diseases: an overview of outcomes and practice guidelines, Bone Marrow Transplant 31 (2003) 229-239.

[41] W. Krivit, J.H. Sung, E.G. Shapiro, L.A. Lockman, Microglia: the effector cell for reconstitution of the central nervous system following bone marrow transplantation for lysosomal and peroxisomal storage diseases, Cell Transplant 4 (1995) 385-392.

[42] R.E. Field, J.A. Buchanan, M.G. Copplemans, P.M. Aichroth, Bone-marrow transplantation in Hurler's syndrome. Effect on skeletal development, J Bone Joint Surg Br 76 (1994) 975-981.

[43] M. Cantz, H. Kresse, Sandhoff disease: defective glycosaminoglycan catabolism in cultured fibroblasts and its correction by beta-N-acetylhexosaminidase, Eur J Biochem 47 (1974) 581-590.

[44] R.J. Desnick, Enzyme replacement and enhancement therapies for lysosomal diseases, J Inherit Metab Dis 27 (2004) 385-410.

[45] J. Keutzer, J. Yee, Enzyme replacement therapy for lysosomal storage disorders, Hum Gene Ther 19 (2008) 857; author reply 858.

[46] R.O. Brady, Enzyme replacement for lysosomal diseases, Annu Rev Med 57 (2006) 283-296.

[47] S. Tomatsu, A.M. Montano, A. Ohashi, M.A. Gutierrez, H. Oikawa, T. Oguma, V.C. Dung, T. Nishioka, T. Orii, W.S. Sly, Enzyme replacement

therapy in a murine model of Morquio A syndrome, Hum Mol Genet 17 (2008) 815-824.

[48] L. Ellgaard, A. Helenius, Quality control in the endoplasmic reticulum, Nat Rev Mol Cell Biol 4 (2003) 181-191.

[49] L. Ellgaard, M. Molinari, A. Helenius, Setting the standards: quality control in the secretory pathway, Science 286 (1999) 1882-1888.

[50] A. Helenius, T. Marquardt, I. Braakman, The endoplasmic reticulum as a protein-folding compartment, Trends Cell Biol 2 (1992) 227-231.

[51] J. Matsuda, O. Suzuki, A. Oshima, Y. Yamamoto, A. Noguchi, K. Takimoto, M. Itoh, Y. Matsuzaki, Y. Yasuda, S. Ogawa, Y. Sakata, E. Nanba, K. Higaki, Y. Ogawa, L. Tominaga, K. Ohno, H. Iwasaki, H. Watanabe, R.O. Brady, Y. Suzuki, Chemical chaperone therapy for brain pathology in G(M1)-gangliosidosis, Proc Natl Acad Sci U S A 100 (2003) 15912-15917.

[52] M. Malinowska, F.L. Wilkinson, W. Bennett, K.J. Langford-Smith, H.A. O'Leary, J. Jakobkiewicz-Banecka, R. Wynn, J.E. Wraith, G. Wegrzyn, B.W. Bigger, Genistein reduces lysosomal storage in peripheral tissues of mucopolysaccharide IIIB mice, Mol Genet Metab 98 (2009) 235-242.

[53] M. Jeyakumar, R.A. Dwek, T.D. Butters, F.M. Platt, Storage solutions: treating lysosomal disorders of the brain, Nat Rev Neurosci 6 (2005) 713-725.

[54] A.L. Roberts, B.J. Thomas, A.S. Wilkinson, J.M. Fletcher, S. Byers, Inhibition of glycosaminoglycan synthesis using rhodamine B in a mouse model of mucopolysaccharidosis type IIIA, Pediatr Res 60 (2006) 309-314.

[55] M. Jeyakumar, D.A. Smith, I.M. Williams, M.C. Borja, D.C. Neville, T.D. Butters, R.A. Dwek, F.M. Platt, NSAIDs increase survival in the Sandhoff disease mouse: synergy with N-butyldeoxynojirimycin, Ann Neurol 56 (2004) 642-649.

[56] C. Bartsocas, H. Grobe, J.J. van de Kamp, K. von Figura, H. Kresse, U. Klein, M.A. Giesberts, Sanfilippo type C disease: clinical findings in four

patients with a new variant of mucopolysaccharidosis III, Eur J Pediatr 130 (1979) 251-258.

[57] E.D. Kakkis, J. Muenzer, G.E. Tiller, L. Waber, J. Belmont, M. Passage, B. Izykowski, J. Phillips, R. Doroshow, I. Walot, R. Hoft, E.F. Neufeld, Enzyme-replacement therapy in mucopolysaccharidosis I, N Engl J Med 344 (2001) 182-188.

[58] M.J. Valstar, G.J. Ruijter, O.P. van Diggelen, B.J. Poorthuis, F.A. Wijburg, Sanfilippo syndrome: A mini-review, J Inherit Metab Dis (2008).

[59] H.V.F. Kresse, K. Barttsocas C., Clinical and biochemical findings in a family with Sanfilippo disease, type C., Clin Genet 10 (1976).

[60] U. Klein, H. Kresse, K. von Figura, Sanfilippo syndrome type C: deficiency of acetyl-CoA:alpha-glucosaminide N-acetyltransferase in skin fibroblasts, Proc Natl Acad Sci U S A 75 (1978) 5185-5189.

[61] U. Klein, J.J. van de Kamp, K. von Figura, R. Pohlmann, Sanfilippo syndrome type C: assay for acetyl-CoA: alpha-glucosaminide N-acetyltransferase in leukocytes for detection of homozygous and heterozygous individuals, Clin Genet 20 (1981) 55-59.

[62] R. Pinto, C. Caseiro, M. Lemos, L. Lopes, A. Fontes, H. Ribeiro, E. Pinto, E. Silva, S. Rocha, A. Marcao, I. Ribeiro, L. Lacerda, G. Ribeiro, O. Amaral, M.C. Sa Miranda, Prevalence of lysosomal storage diseases in Portugal, Eur J Hum Genet 12 (2004) 87-92.

[63] B.J. Poorthuis, R.A. Wevers, W.J. Kleijer, J.E. Groener, J.G. de Jong, S. van Weely, K.E. Niezen-Koning, O.P. van Diggelen, The frequency of lysosomal storage diseases in The Netherlands, Hum Genet 105 (1999) 151-156.

[64] J.T. Gallagher, Multiprotein signalling complexes: regional assembly on heparan sulphate, Biochem Soc Trans 34 (2006) 438-441.

[65] G. Yogalingam, J.J. Hopwood, Molecular genetics of mucopolysaccharidosis type IIIA and IIIB: Diagnostic, clinical, and biological implications, Hum Mutat 18 (2001) 264-281.

[66] L.H. Rome, D.F. Hill, K.J. Bame, L.R. Crain, Utilization of exogenously added acetyl coenzyme A by intact isolated lysosomes, J Biol Chem 258 (1983) 3006-3011.

[67] K.J. Bame, L.H. Rome, Acetyl coenzyme A: alpha-glucosaminide N-acetyltransferase. Evidence for a transmembrane acetylation mechanism, J Biol Chem 260 (1985) 11293-11299.

[68] K.J. Bame, L.H. Rome, Genetic evidence for transmembrane acetylation by lysosomes, Science 233 (1986) 1087-1089.

[69] K.J. Bame, L.H. Rome, Acetyl-coenzyme A:alpha-glucosaminide N-acetyltransferase. Evidence for an active site histidine residue, J Biol Chem 261 (1986) 10127-10132.

[70] J. Ausseil, J.C. Loredo-Osti, A. Verner, C. Darmond-Zwaig, I. Maire, B. Poorthuis, O.P. van Diggelen, T.J. Hudson, T.M. Fujiwara, K. Morgan, A.V. Pshezhetsky, Localisation of a gene for mucopolysaccharidosis IIIC to the pericentromeric region of chromosome 8, J Med Genet 41 (2004) 941-945.

[71] V. Seyrantepe, F. Tihy, A.V. Pshezhetsky, The microcell-mediated transfer of human chromosome 8 restores the deficient N-acetylytransferase activity in skin fibroblasts of Mucopolysaccharidosis type IIIC patients, Hum Genet 120 (2006) 293-296.

[72] M. Hrebicek, L. Mrazova, V. Seyrantepe, S. Durand, N.M. Roslin, L. Noskova, H. Hartmannova, R. Ivanek, A. Cizkova, H. Poupetova, J. Sikora, J. Urinovska, V. Stranecky, J. Zeman, P. Lepage, D. Roquis, A. Verner, J. Ausseil, C.E. Beesley, I. Maire, B.J. Poorthuis, J. van de Kamp, O.P. van Diggelen, R.A. Wevers, T.J. Hudson, T.M. Fujiwara, J. Majewski, K. Morgan, S. Kmoch, A.V. Pshezhetsky, Mutations in TMEM76* cause mucopolysaccharidosis IIIC (Sanfilippo C syndrome), Am J Hum Genet 79 (2006) 807-819.

[73] X. Fan, H. Zhang, S. Zhang, R.D. Bagshaw, M.B. Tropak, J.W. Callahan, D.J. Mahuran, Identification of the gene encoding the enzyme deficient in mucopolysaccharidosis IIIC (Sanfilippo disease type C), Am J Hum Genet 79 (2006) 738-744.

[74] C. Freeman, P.R. Clements, J.J. Hopwood, Acetyl CoA:alpha-glucosaminide N-acetyl transferase: partial purification from human liver, Biochem Int 6 (1983) 663-671.

[75] J.J. Hopwood, H. Elliott, Selective depolymerisation of heparin to produce radio-labelled substrates for sulfamidase, 2-acetamido-2-deoxy-alpha-D-glucosidase, acetyl-CoA:2-amino-2-deoxy-alpha-D-glucoside N-acetyltransferase, and 2-acetamido-2-deoxy-D-glucose 6-sulfate sulfatase, Carbohydr Res 91 (1981) 165-190.

[76] U. Klein, K. von Figura, A 3H-labelled trisaccharide from heparin as substrate for acetyl-CoA: 2-amino-2-deoxy-alpha-D-glucoside N-acetyltransferase, Carbohydr Res 78 (1980) 249-256.

[77] J.J. Hopwood, H. Elliott, The diagnosis of the Sanfilippo C syndrome, using monosaccharide and oligosaccharide substrates to assay acetyl-CoA: 2-amino-2-deoxy-alpha-glucoside N-acetyltransferase activity, Clin Chim Acta 112 (1981) 67-75.

[78] P.J. Meikle, A.M. Whittle, J.J. Hopwood, Human acetyl-coenzyme A:alpha-glucosaminide N-acetyltransferase. Kinetic characterization and mechanistic interpretation, Biochem J 308 ( Pt 1) (1995) 327-333.

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

[80] A.O. Fedele, M. Filocamo, M. Di Rocco, G. Sersale, T. Lubke, P. di Natale, M.P. Cosma, A. Ballabio, Mutational analysis of the HGSNAT gene in Italian patients with mucopolysaccharidosis IIIC (Sanfilippo C syndrome). Mutation in brief #959. Online, Hum Mutat 28 (2007) 523.

[81] G.J. Ruijter, M.J. Valstar, J.M. van de Kamp, R.M. van der Helm, S. Durand, O.P. van Diggelen, R.A. Wevers, B.J. Poorthuis, A.V. Pshezhetsky, F.A. Wijburg, Clinical and genetic spectrum of Sanfilippo type C (MPS IIIC) disease in The Netherlands, Mol Genet Metab 93 (2008) 104-111.

[82] M.F. Coutinho, L. Lacerda, M.J. Prata, H. Ribeiro, L. Lopes, C. Ferreira, S. Alves, Molecular characterization of Portuguese patients with

mucopolysaccharidosis IIIC: two novel mutations in the HGSNAT gene, Clin Genet 74 (2008) 194-195.

[83] M.P. Zeegers, F. van Poppel, R. Vlietinck, L. Spruijt, H. Ostrer, Founder mutations among the Dutch, Eur J Hum Genet 12 (2004) 591-600.

# APPENDIX I

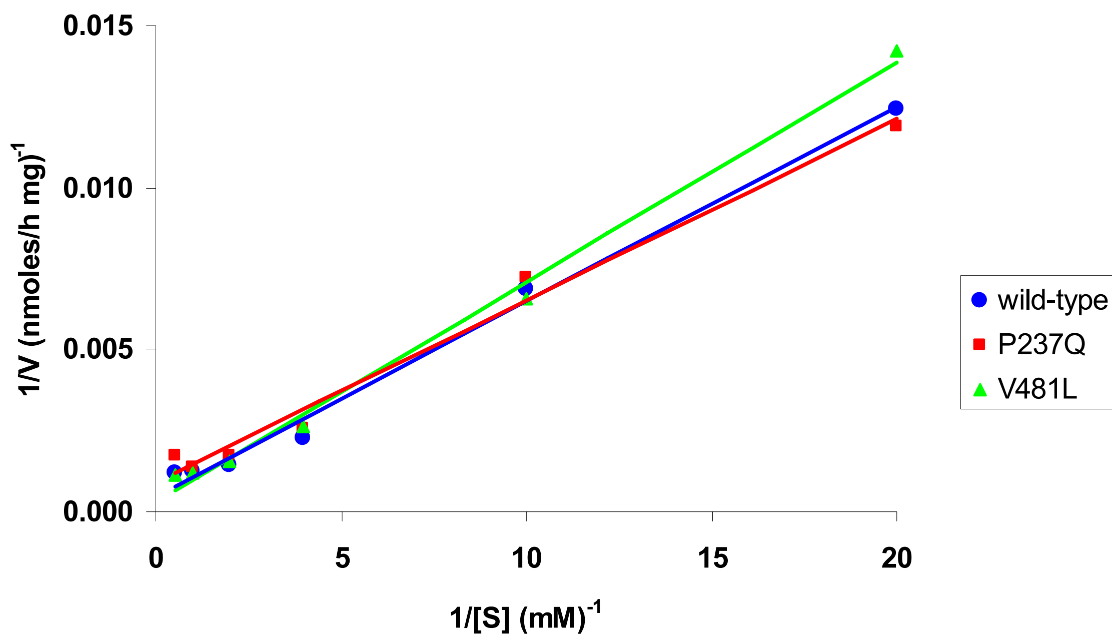
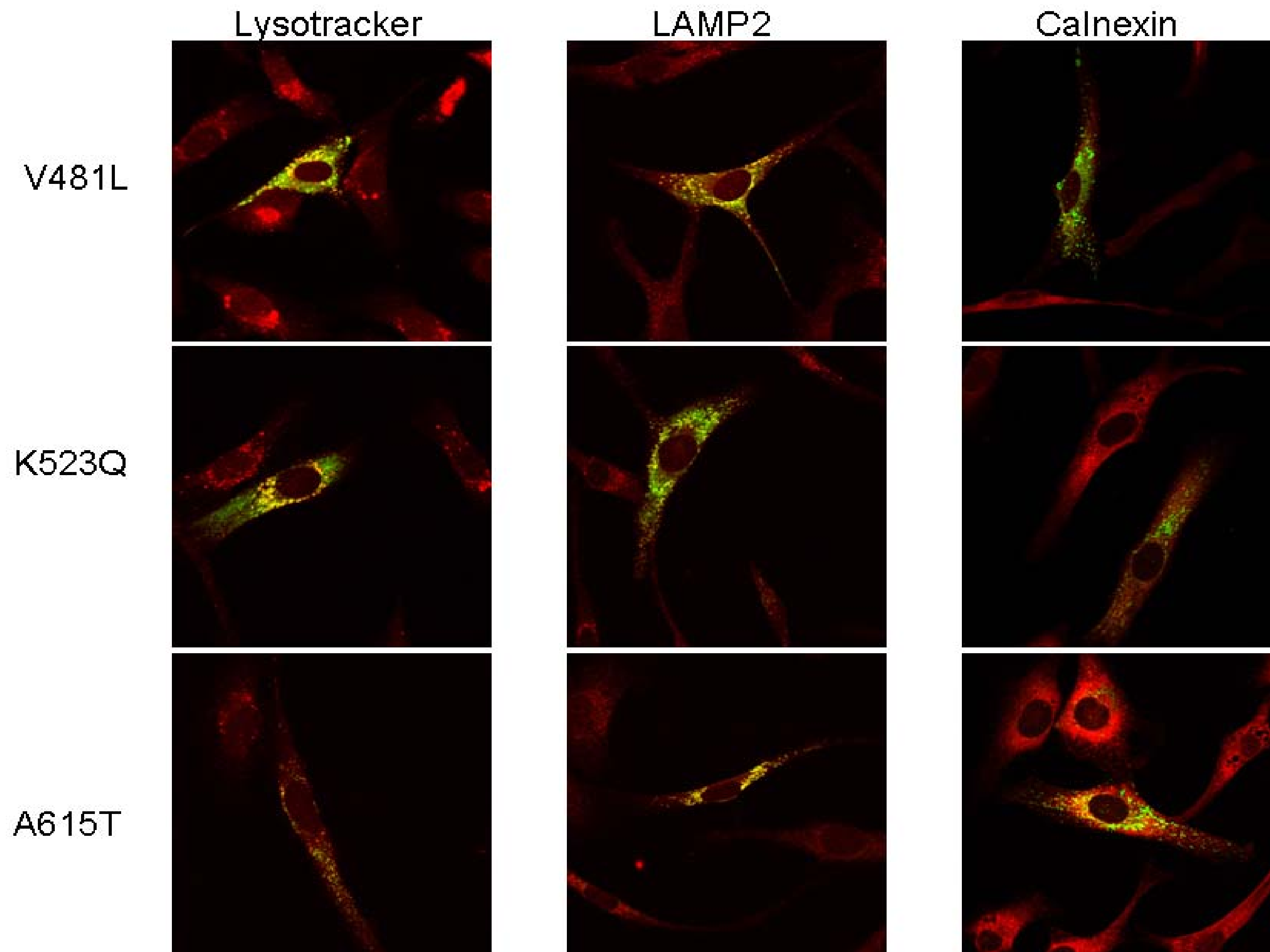
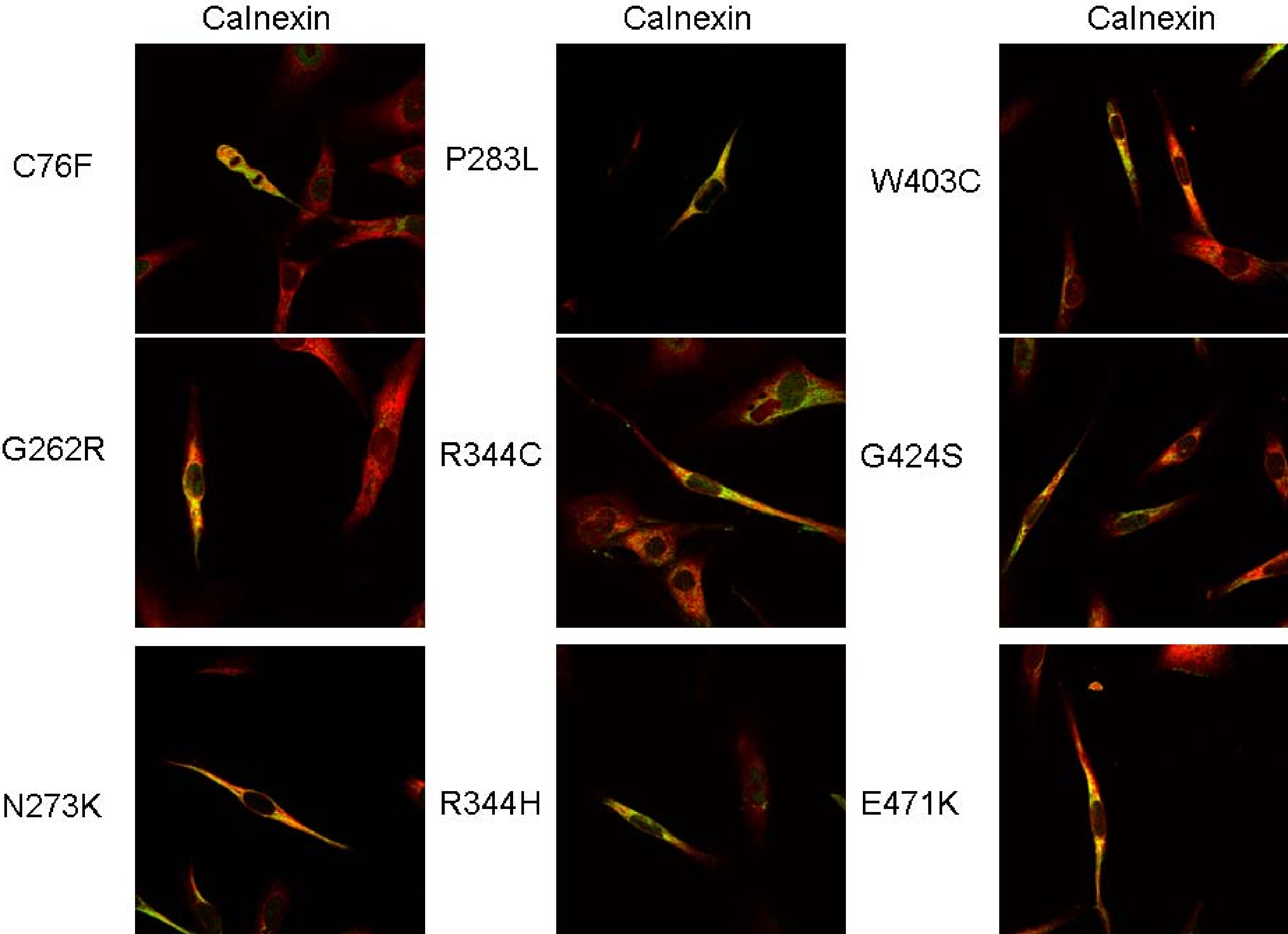


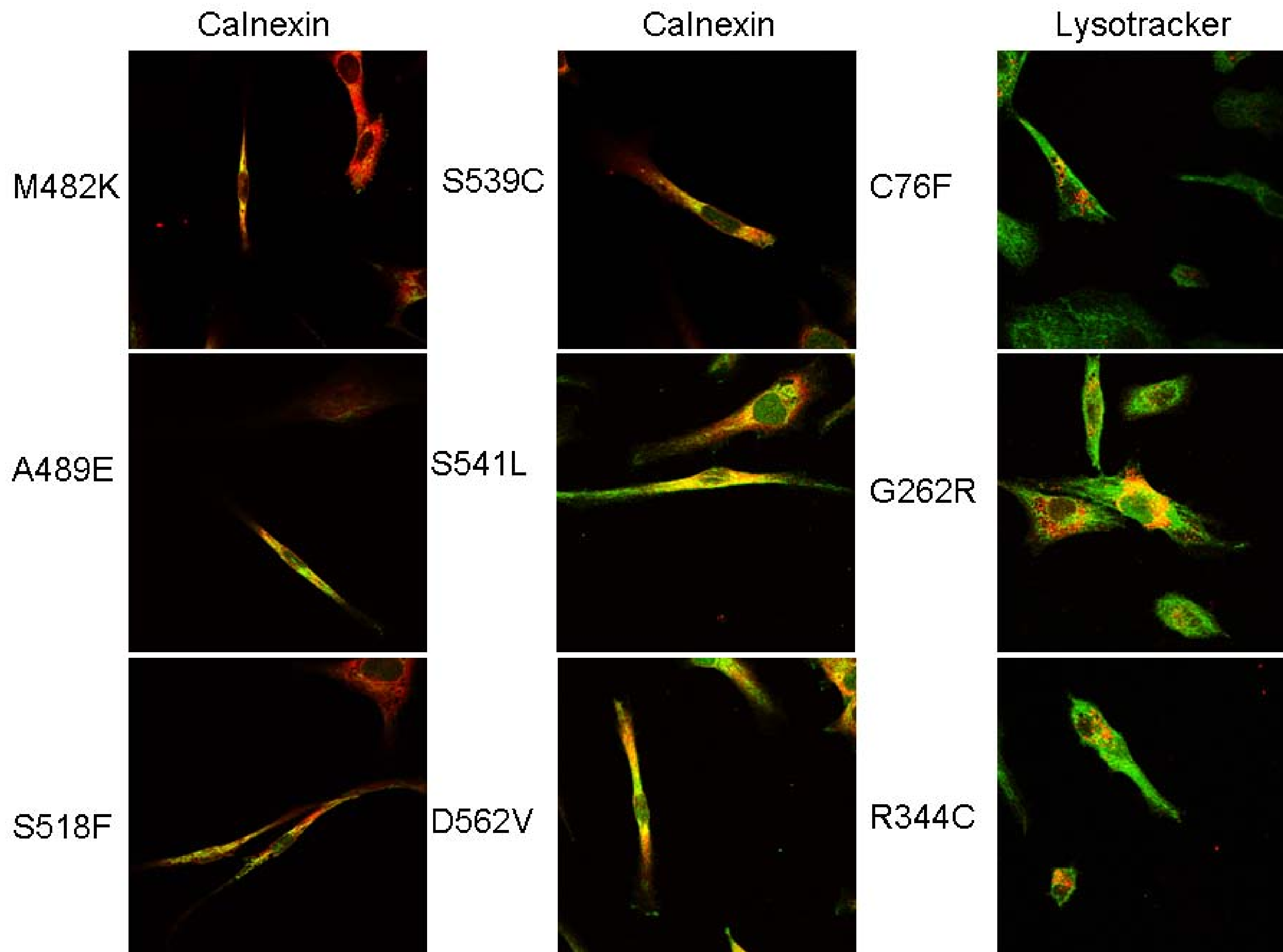

**Figure S1**. Lineweaver-Burk plot of substrate dependance for partially purified HGSNAT wild-type and P237Q and V481L mutants. COS-7 cells were harvested 42 hrs after transfection with the wild-type or mutant HGSNAT plasmids and suspended in lysis buffer (40 mM Tris-HCl, 300 mM KCl, pH 7.5, 0.1% NP-40, 1 mM PMSF and Sigma P8340 protease inhibitor cocktail at 10 µl per 1 ml of cell suspension). The homogenate was sonicated, gently shaked at 4°C for 2 h and centrifuged at 13,000 rpm for 30 min. The supernatant was first passed through an avidin-agarose column (Sigma A9207) then affinity purification of TAP-tagged HGSNAT was performed using streptavidin resin (Stratagene) according to the manufacturer's protocol. N-acetyltransferase activity was assayed as described in Material and Methods using 0.05 to 2.0 mM 4MU-βGlcN and 18 h incubation time. KM and VMAX values for all 3 enzymes were similar within the statistical error.

# APPENDIX II


Lysotracker
LAMP2
Calnexin
V481L
K523Q
A615T



Calnexin
Calnexin
Calnexin
C76F
P283L
W403C
G262R
R344C
G424S
N273K
R344H
E471K



Calnexin
Calnexin
Lysotracker
M482K
S539C
C76F
A489E
S541L
G262R
S518F
D562V
R344C

**Figure S2**. Localization of HGSNAT mutants expressed in cultured human skin fibroblasts by immunofluorescence microscopy. The cells transfected with wild-type or mutant HGSNAT-TAP constructs as indicated were fixed and stained with either mouse monoclonal anti-LAMP-2 antibodies, Lysotracker Red DND-99 or mouse monoclonal anti-calnexin antibodies (red) and rabbit polyclonal anti-CBP antibodies (green) as indicated. Slides were studied on a Zeiss LSM510 inverted confocal microscope. Magnification 630x. Panels show representative images showing co-localization of anti-CBP antibodies (green) and lysosomal and ER markers (red) for the active enzyme containing polymorphisms and all inactive mutants.

# APPENDIX III

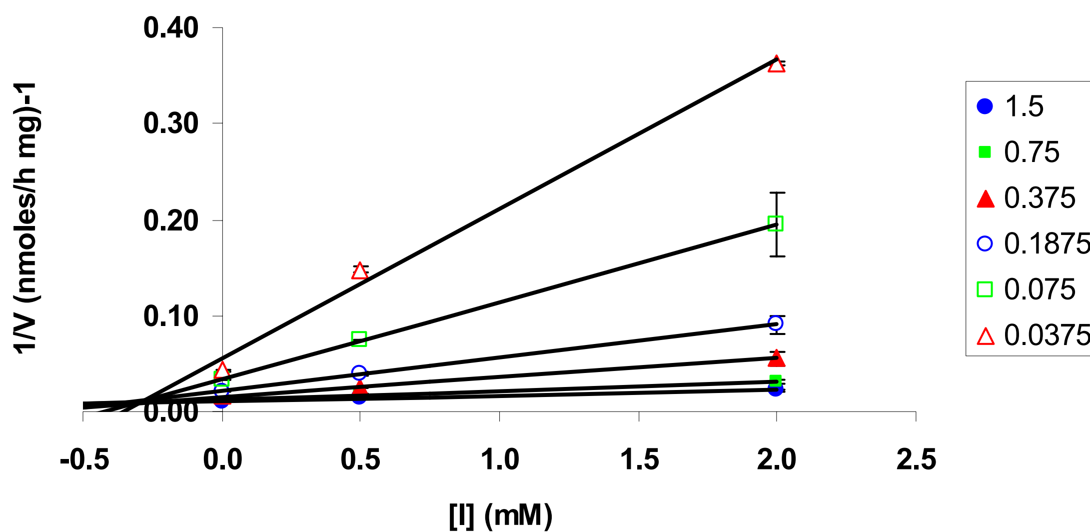



**Figure S3**. Dixon plot showing the inhibition of HGSNAT by glucosamine. COS-7 cells were harvested 42 hrs after transfection with the wild-type HGSNAT plasmid and N-acetyltransferase activity was measured in the homogenates for 3 h at 37°C in the presence of 2 mM AcCoA, 0.0375 to 1.5 mM 4MU-βGlcN and 0 to 2 mM D-(+)-glucosamine hydrochloride.

# APPENDIX IV

| Mutant | Sense Primer | Antisense Primer |
|---|---|---|
| C76F | ctactggaaatctgaatgcttttatcactgcttgtttcagg | cctgaaacaagcagtgataaaagcattcagatttccagtag |
| L137P | tggagaatttggaaactattctcccttggtaaagaacatccataat | attatggatgttctttaccaagggagaatagtttccaaattctcca |
| P237Q | atctgccctgcagccccgcctcc | ggaggcggggctgcagggcagat |
| G262R | tggtctttgtcaattatggaggaagaaaatattggtacttcaaacat | atgtttgaagtaccaatattttcttcctccataattgacaaagacca |
| N273K | tacttcaaacatgcaagttggaaggggctgacagtg | cactgtcagccccttccaacttgcatgtttgaagta |
| P283L | tggctgacctcgtgttcctgtggtttgtatttattatg | cataataaatacaaaccacaggaacacgaggtcagcca |
| R344C | tcttgggacaaggtgtgcattcctggtgtgc | gcacaccaggaatgcacaccttgtcccaaga |
| R344H | cttgggacaaggtgcacattcctggtgtgct | agcacaccaggaatgtgcaccttgtcccaag |
| W403C | gctggaaggcctgtgtctgggcttgacattc | gaatgtcaagcccagacacaggccttccagc |
| G424S | gttatcttggtcctgggagcattggagattttggc | gccaaaatctccaatgctcccaggaccaagataac |
| E471K | ggcctatgaccccaagggcatcctggg | ccggatactggggttcccgtaggaccc |
| V481L | gcaccatcaactccatcttgatggccttttttagga | tcctaaaaaaggccatcaagatggagttgatggtgc |
| M482K | gcaccatcaactccatcgtgaaggcctttttagg | cctaaaaaggccttcacgatggagttgatggtgc |
| A489E | atggcctttttaggagttcaggaaggaaaaatactattgtattaca | tgtaatacaatagtatttttccttcctgaactcctaaaaaggccat |
| S518F | ttgtattcttgggctcatttttgttgctctgacgaagg | ccttcgtcagagcaacaaaaatgagcccaagaatacaa |
| K523Q | catttctgttgctctgacgcaggtttctgaaaatgaagg | ccttcattttcagaaacctgcgtcagagcaacagaaatg |
| S539C | ccagtaaacaaaaatctctggtgcctttcgtatgtcactac | gtagtgacatacgaaaggcaccagagatttttgtttactgg |
| S541L | aaaatctctggtcccttttgtatgtcactacgctcag | ctgagcgtagtgacatacaaaagggaccagagatttt |
| D562V | tgtacccagttgtggttgtgaaggggctgtg | cacagccccttcacaaccacaactgggtaca |
| P571L | gctgtggacaggaaccctattcttttatccaggaa | ttcctggataaaagaatagggttcctgtccacagc |
| A615T | ctcagaacatcgtcaccactgccctctgg | ccagagggcagtggtgacgatgttctgag |

**Table S1.** Primers for site-directed mutagenesis of HGSNAT-TAP plasmid.