NOVEL ASPECTS OF MOLECULAR BIOLOGY
AND PATHOLOGY OF LYOSOMAL STORAGE
DISORDERS. STUDIES WITH PARTIAL USE OF
Caenorhabditis elegans

Ph.D. thesis summary
Jakub Sikora

Supervisor: Prof. MUDr. Milan Elleder, DrSc.
Consultant: MUDr. Martin Hřebíček

Prague 2007
This thesis was elaborated within the postgraduate program for Biology and Pathology of the Cell at the Institute of Inherited Metabolic Disorders, 1st Faculty of Medicine, Charles University, Prague.

Chairman of the commision: Prof. MUDr. Karel Smetana, DrSc.
Institute of Hematology and Blood Transfusion, U nemocnice 1, Praha 2, 128 20

This work was funded by the grants 301/93/01854, 303/02/1324 provided by the Czech Scientific Agency (Grant Agency of the Czech Republic); and by the grant 8351-3 of the Ministry of Health of the Czech Republic. The institutional financial support was provided by the research projects 111100003 and 0021620806 of the Ministry of Education, Youth and Sports of the Czech Republic.

Note:
Authorship of all publications is in compliance with the criteria for authorship as accepted by the International Committee of Medical Journal Editors (ICMJE). Individual contributions can be provided on request if not listed directly in the publications.
In order to avoid any ambiguities, word mutation was omitted in the following text, if not stated otherwise, and was replaced by the term pathogenic variation in accordance with the valid recommendations.
Contents

1 Introduction ........................................................................................................................................... 5
  1.1 Lysosomes - 2007, overview ............................................................................................................. 5
  1.2 Lysosomal storage disorders - 2007, overview ................................................................................ 6

2 Summarized aims of the studies ........................................................................................................... 7

3 Saposin-like proteins in lysosomal biology and pathology – prosaposin, saposins and acid sphingomyelinase ................................................................. 8
  3.1 Prosaposin deficiency – enzyme activator precursor versus neurotrophic factor ............................ 8
    3.1.1 Prosaposin and saposins - gene, cellular trafficking, posttranslational modifications, functional versatility .......................................................... 8
    3.1.2 Prosaposin deficiency, overview ................................................................................................ 9
    3.1.3 Comments and discussion to the publication Sikora et al. (2007) .................................................. 10
    3.1.3.1 Specific features of neuronal affection in pSap deficiency ...................................................... 10
  3.2 Acid sphingomyelinase - Niemann-Pick disease type A/B – acid sphingomyelinase deficiency – model lysosomal enzymopathy – molecular studies .......................................................... 12
    3.2.1 Sphingomyelin and sphingomyelinas ......................................................................................... 12
    3.2.2 Molecular genetics of ASM (SMPD1 gene) ................................................................................ 12
    3.2.3 Cellular and organ pathology of ASM deficiency, consequences of the storage process .......... 13
    3.2.3.1 Neuronal storage compartment ............................................................................................ 13
    3.2.3.2 Visceral storage compartment .............................................................................................. 14
    3.2.4 Clinical classification of ASM deficiency, phenotypic heterogeneity versus historical perspective. .................................................................................................................................................. 14
    3.2.5 Molecular pathology of the SMPD1 gene, population genetics, genotype-phenotype correlations, implications of SMPD1 gene variations for ASM protein malfunction .......................................................................................................................... 15
    3.2.6 NPD A/B diagnostics – histopathology, enzyme biochemistry, molecular genetics, differential diagnosis, and therapy ................................................................................................................. 16
    3.2.7 Comments and discussion to the publications Sikora et al. (2003), Pavlíč-Perreira et al. (2005) and Dvořáková et al. (2006) ......................................................................................................................... 17
      3.2.7.1 Sikora et al. (2003) ................................................................................................................ 17
      3.2.7.2 Pavlíč-Perreira et al. (2005) .............................................................................................. 18
      3.2.7.2.1 ASM deficiency phenotypic variability and related implications ....................................... 18
      3.2.7.2.2 SMPD1 genotype in central Europe, implications for ASM enzymatic activity assays .... 18
      3.2.7.2.3 Proposals to the classification of ASM deficiency ........................................................ 19
      3.2.7.3 Dvořáková et al. (2006) .................................................................................................. 20
        3.2.7.3.1 Niemann-Pick disease type C (lysosomal storage disorder due to defective intracellular lipid trafficking) .......................................................... 21

4 Bioinformatic predictive study of acetyl-coenzyme A:α-glucosaminide N-acetyltransferase - Hřebiček et al. (2006) ......................................................................................................................... 22

5 Caenorhabditis elegans as a model organism for selected lysosomal storage diseases .......................................................... 25
  5.1 C. elegans, general characteristics ................................................................................................. 26
    5.1.1 Endosomal - lysosomal system in C. elegans .......................................................................... 26
  5.2 Comments and discussion to the publication - Hujová, Sikora, Dobrovolný et al. (2005) ............. 28
    5.2.1 gana-1 evolutionary context .................................................................................................. 28
    5.2.2 gana-1 enzymatic activity, RNAi induced compound enzymatic deficiency, cellular localization of gana-1 .......................................................................................................................... 29
1 Introduction

Christian de Duve and his colleagues proposed the existence of a novel type of membrane endowed intracellular granules rich in hydrolytic enzymes, and designated them as lysosomes, in the year 1955.

1.1 Lysosomes - 2007, overview

Lysosomes represent an integral part of membranous endosomal – lysosomal system that is defined by progressive intraluminal pH decrease realized by the action of vacuolar H+ ATPase. The tendency to subdivide the system, whose inherent function is continuous multidirectional flow of membranous constituents and luminal cargo molecules, sometimes results in unwanted simplifications. The intimate functional connection between late endosomes (mannose-6-phosphate receptor – M-6-PR positive) and lysosomes (M-6-PR negative) resulted in compound designation: late endosomal-lysosomal.

As stated above, endosomal - lysosomal system represents extremely dynamic compartment with widespread contacts in the cell. Spatial coordination of membrane and luminal contents shuttling is vast and in majority of aspects not fully understood. Many of these processes require complex molecular interactions of protein-protein or protein-lipid phases (see also sections 1.2 and 3). Late endosomal-lysosomal compartment exerts extensive contacts with other cellular compartments (Golgi apparatus, endoplasmic reticulum, mitochondria or cytoplasmic membrane). The system is further functionally and spatially stratified.

Current knowledge about lysosomal proteomics assigns nearly ten (or more) times more proteins to reside and function in this cellular compartment. Lysosomal proteins are either luminal or integral membrane proteins. As many substrate molecules are hydrophobic by nature and are distributed in layered membrane sheets, considerable proportion of luminal proteins functions in intimate membrane association and displays membrane perturbation potential (for further details see section 3). Common denominator of the two protein groups (lysosomal hydrolases and structural non-enzymatic proteins) is their acidic pH functional optimum.

Lysosomal membrane proteome and lipidic constitution are intensively studied at the present time. Growing number of lysosomal membrane proteins as well as diversity of their functions are striking. Latest advances including characterization of the gene coding for lysosomal membrane protein acetyl-coenzyme A:α-glucosaminide N-acetyltransferase (HGSNAT, section 4) or the issue of detergent-resistant microdomains (DRMs) and their participation in chaperon mediated autophagy (CMA) are discussed in sections 1.2 and 4. The composition, homeostasis, recycling and remodelling of the lysosomal membrane bilayer is a cornerstone of lysosomal function with implications for human pathology states.

The fundamental quality of lysosomal apparatus as ordered intracellular system is, besides others, maintained by energy dependent selective delivery of lysosomal proteins. There are three different protein targeting mechanisms that employ 3D secondary protein structure sequence tags: mannose-6-phosphate (M-6-P) recognition targeting, mannose receptor (MR) dependent pathway and sortillin protein receptor and carrier pathway employed by saposin-like lysosomal proteins (see 3.1.1 and 3.2.1.1). Two additional mechanisms of selective protein delivery to lysosomes employ conserved targeting sequences in the primary protein structure. Lysosomal membrane proteins are delivered on the basis of the presence of either tyrosin or dileucin motifs. Selective import of certain cytosolic


proteins into the lysosomal lumen is the basis of CMA mediated by LAMP 2a (lysosomal associated membrane protein) isoform and hsc-73 co-chaperon.\textsuperscript{18-20}

Substrate entry into the late endosomal-lysosomal compartment is realized by three distinct pathways: endocytosis, phagocytosis and autophagy.\textsuperscript{4} Proportional contribution of each of these delivery pathways is substrate, cell and tissue type specific, and is difficult to estimate, universally. Only autophagy will be briefly discussed. Autophagy is an evolutionary conserved mechanism used for degradation of inherent cellular contents. Current concept of autophagy recognizes three basic variants: macroautophagy (degradation of complete cytosolic regions with prior generation of autophagosomal membrane), microautophagy (degradation of cytosolic regions with direct engulfment by lysosomal membrane) and chaperon mediated autophagy (CMA, selective protein degradation pathway employing protein sequence tags).\textsuperscript{21,22} Latest publications delineated autophagosomal membrane constituting machinery of ubiquitin-like proteins including critical LC3 component\textsuperscript{23,24} in case of macroautophagy. For CMA, it was the characterization of the selectivity of lysosomal targeting of certain cytosolic proteins with KFERQ-like signal sequences, description of the critical protein recognition and trans-lysosomal membrane transfer and roles of DRMs in the whole process\textsuperscript{9,18,20}.

Autophagy represents one of the three basic mechanisms of protein elimination in eukaryotic cell (other two are cytosolic proteases and ubiquitin – proteasomal pathways). These three pathways function in partial overlap and substitutability.\textsuperscript{25-27} It was shown that autophagy (both macro- and CMA variants) represents a constitutively active and continuous process, which is indispensable for proper cellular protein clearance.\textsuperscript{28-32} The issues of cellular protein elimination or autophagy related programmed cell death (PCD II) should thus be considered integral part of lysosomal function.\textsuperscript{33}

\textbf{1.2 Lysosomal storage disorders - 2007, overview}

At the present time, lysosomal storage disorders (LSDs) represent a group of 48 entities\textsuperscript{34} that are defined on the molecular level and are inherited as recessive Mendelian monogenic traits (overall incidence of 1:6000 newborns).\textsuperscript{35} Common denominator in this group of molecular defects is the expansion of lysosomal compartment of affected cell types with foamy transformation of the cytoplasm.

Majority of LSDs (30 out of 48 entities) are single lysosomal enzymatic deficiencies (29 hydrolases and 1 acetyl-transferase). Summarizing list provides the following: ten mucopolysaccharidoses (MPS), seven glycoproteinoses, nine lipidoses, one glycogen storage disease, and three neuronal ceroid lipofuscinoses (NCLs) due to enzymatic deficiency. The gene discovery period of biochemically defined LSDs was closed by the characterization of the critical gene in MPS type IIIc (heparin acetyl-coenzyme A:α-glucosaminide N-acetyltransferase deficiency). MPS IIIc thus represents the only synthetic enzymatic deficiency (section 4) in the group of otherwise hydrolytic deficiencies.\textsuperscript{8}

Another eight LSDs result from enzymatic deficiency due to either malfunction of co- or post-translation protein processing (multi sulphatase deficiency, mucolipidosis II/III), lack of enzymatic protection by protective protein/Cathepsin A (PPCA, galactosialidosis), or from enzyme activator deficiency (GM\textsubscript{1} activator deficiency, saposin and prosaposin deficiencies, section 3.1).

Rest of lysosomal storage disorders (ten) result from molecular defects in lysosomal membrane proteins. Two of these are trans-membrane transporter defects (cystinosis, sialic acid storage disease). Remaining eight disorders are difficult to characterize on the basis of their common molecular pathogenesis. Some of these issues are discussed later (sections 3.1.2, 3.1.3 and 3.2.7.3.1). Cellular phenotype of these eight disorders originates from
defective membrane content handling (shuttling, targeting or redistribution) or from inability of lysosomal compartment to perform some of its principal goals, such as participation in autophagy.

The characteristics (morphological, biochemical) of cellular and tissue storage patterns differ considerably among individual disorders. The process of lysosomal storage due to accumulation of whatever material remains substantially enigmatic in the following areas: (i) functionality and turnover of the storage lysosomes; (ii) compensatory mechanisms in lysosomal biogenesis and lysosomal proteomics; (iii) recycling and redistribution of accumulated material; (iv) impact of storage process on partner cellular compartments and fundamental cellular regulations including protein clearance and cell death induction; (v) cell and tissue specific consequences of lysosomal storage; (vi) extralysosomal roles of both deficient proteins and accumulated moieties.

Impacts of lysosomal storage on the lysosomal biogenesis might be quite diverse, sometimes very pronounced, and were demonstrated on RNA and protein levels. The exact machinery involved and mechanisms of some of these phenomena are hypothetical, common transcription regulation of GC rich TAATA less promoters of lysosomal related genes included.

Redistribution of undegraded substrate molecules in the affected cells and their extralysosomal effects might be also very important in the pathogenesis of lysosomal storage. It is especially the issue of glycolipid recycling and genesis of toxic moieties that is currently in focus. Direct involvement in fundamental cellular regulatory pathways and non-lysosomal functions of proteins deficient in LSDs and their target molecules represent further contributory factors to LSD pathogenesis (see 3.1 and 3.2).

Lysosomal roles in programmed cell death (PCD) of either caspase mediated (type I) or autophagic (type II) are widespread. Direct participation of certain glycolipids or ceramide in apoptosis regulation or induction is a generally accepted fact (see section 3.2). Nevertheless, it seems more and more probable that combination of these phenomena (PCD type I and II) is the reality in LSDs. If one considers the extent of regulation of CMA and extensive consequences of its deficiency, protein clearance in LSDs is another issue contributing to cellular pathology.

As was demonstrated above, all of the historical LSDs were defined on the molecular level during the last 40 years. It is tempting to speculate whether other, unrecognized defects with “lysosomal” phenotype will be characterized in future.

2 Summarized aims of the studies

The detailed aims of the studies are listed and defined in the particular sections of the following text, this section provides introductory overview.

Saposin-like proteins in lysosomal biology and pathology: prosaposin, saposins and acid sphingomyelinase

- To evaluate neuropathologic changes in the available cases of pSap deficiency with special emphasis on the differences between neurolysosomal and non-neurolysosomal populations
- To evaluate roles of protein clearance mechanisms (ubiquitin-proteasome) in the pSap pathology and compare the findings with other selected LSDs
- To evaluate potential origins of massive postnatal loss of cortical neurons with respect to potential roles PCD type II in this process
To perform genotype-phenotype correlation studies in the selected populations of acid sphingomyelinase deficient patients (Niemann-Pick disease type A and B) with emphasis on the phenotypic diversity

Bioinformatic study of HGSNAT protein, critical protein in MPSIIIc pathogenesis
- To perform detailed predictive study of the evolutionary and functional traits of the newly defined HGSNAT protein involved in MPSIIIc pathogenesis. Special emphasis given to interspecies synteny conservation, inherent transmembrane protein domains and potential posttranslation modifications.

Caenorhabditis elegans as a model organism for selected lysosomal storage disorders
- To perform bioinformatic, biochemical and cell biological characterization of the single C.elegans ortholog of mammalian acid α-galactosidase and α-N-acetylgalactosaminidase with special emphasis on evolutionary clustering and diversification of catalytic functions.

3 Saposin-like proteins in lysosomal biology and pathology – prosaposin, saposins and acid sphingomyelinase

Saposin like proteins (SAPLIPs) are membrane interacting proteins of diverse biochemical and cellular functions. Their unifying and characterizing feature is the presence of saposin-like domain in their secondary protein structure. The saposin-like domain includes hydrophobic protein core with six invariable cysteine residues forming three disulphide bridges and a common position of four to five α-helices.

The phylogenetic origins of SAPLIPs are diverse and include proteins from eukaryotic and prokaryotic organisms and even from plants. Human SAPLIPs are, for example, prosaposin (pSap) and deriving Saps (see 3.1), acid sphingomyelinase (ASM, see 3.2.1.1), acyloxy acylase, surfactant protein B or granulysin, for overview see Bruhn (2005).

SAPLIP properties related to membrane (lipid) interactions can be clustered into three areas sharing common features: (i) membrane association with local disordering of the bilayer structure (e.g. ASM); (ii) fundamental membrane perturbation without direct permeabilization for the purposes of antigen presentation or enzymatic activity (e.g. Saps); and (iii) membrane permeabilization for defensive mechanisms (e.g. granulysin). It should be noted that SAPLIPs are directly involved in other cellular signaling pathways besides the direct membrane (lipid) interactions (see 3.1.2 and 3.1.3.1).

3.1 Prosaposin deficiency – enzyme activator precursor versus neurotrophic factor

3.1.1 Prosaposin and saposins - gene, cellular trafficking, posttranslational modifications, functional versatility

Human prosaposin gene (PSAP) is located on the chromosome 10 (10q22.1). PSAP mRNA encodes a polypeptide (~53 kDa, pSap) that contains signal peptide and four distinct but homologous sphingolipid (sphingolipidhydrolase) activating proteins (Sap A-D). Complex ontogenetic temporo-spatial expression patterns were determined for PSAP gene.
pSap (53 kDa) is posttranslationally processed by a succession of modifications. The 65 kDa molecule is associated with Golgi apparatus membranes and is destined to lysosomes for further proteolytic processing. The sorting step discriminating the secretory and lysosomal forms is only partially elucidated. It was demonstrated that Golgi/lysosomal sorting mechanism does not employ M-6-P receptor but is dependent on pSap’s interaction with sphingolipids and sortilin receptor protein. Similar mechanism of Golgi/lysosomal targeting was documented for acid sphingomyelinase (see 3.2.1.1) and GM2 activator protein.

Proteolytic processing of pSap to four Saps in the late endosome-lysosome occurs via cleavage by hydrolytic proteases. Secretory targeting of pSap is present in a number of cell types and pSap was detected in numerous body fluids. Functions of secreted pSap are diverse in different tissues. Some of the downstream cellular signaling exerted by pSap is mediated by its interaction with incompletely defined G0-protein associated receptor.

Human Saps (A-D) function as membrane glycosphingolipid mobilizing agents. Sap interaction with glycosphingolipid or phospholipid membrane sheets intensifies with lower pH and is directly related to overall negative surface charge of Sap molecules. Overall, Saps mediate intermembrane transport of gangliosides and glycosphingolipids, reorganization and destabilization of phospholipid membrane sheets, and fusion of acidic phospholipid vesicles.

Saps (A-D) activator potential is directly linked to their membrane perturbation potential. Saposins A-D function as activators of the following lysosomal hydrolases – galactocerebrosidase, arylsulphatase A, glucocerebrosidase and ceramidase, respectively. Sap B, in addition, functions as lysosomal acid α-galactosidase activator in degradation of globotriaosylceramide. Lactosyl ceramide hydrolysis by β-galactosylceramidase and β-galactosidase is associated with Sap A and B activation, respectively.

Prosaposin has diverse functions that are linked to its lipid interaction characteristics. Its glycosphingolipid and phospholipid binding affinity underlies pSap roles in transport of these moieties; it may also facilitate membrane fusions and remodeling. In addition, pSap was defined as important extracellular regulator with anti-apoptotic and cell cycle promoting functions. pSap was demonstrated to interact with G0 protein coupled receptor in different tissues and trigger MAP kinase and other signaling pathways directly associated with S phase promotion and apoptosis down regulation. Prosaposin’s neurotrophic and neuroprotective effects that are exerted by similar mechanisms of signalling are discussed in section 3.1.3.1. It should be stressed out that this functional duality of pSap molecule has direct consequences in the molecular pathology of pSap deficiency (see 3.1.3).

### 3.1.2 Prosaposin deficiency, overview

Prosaposin deficiency is a rapidly progressive fatal infantile neurovisceral lysosomal storage disorder. The pathogenic consequences must be attributed to the biochemical enzymatic deficiencies and to other inherent functions of pSap (see 3.1.3). pSap deficiency represents very complex biological and pathogenic problem with many different aspects to consider. Molecular basis of pSap deficiency lies in the complete absence of pSap and all four deriving Saps (A-D) in the cells and tissues due to specific pathogenic variations in PSAP gene that interfere either with functional mRNA transcription or with pSap protein translation initiation. Morphological (optical and electron microscopic) and biochemical (accumulated lipid spectra, cellular lipid turnover) alterations can be assigned to the enzymatic deficiency of 5 different lysosomal hydrolases associated with activator functions.
of Saps. Morphological and biochemical consequences of pSap deficiency can be attributed to a combination of changes found in five different and defined single lysosomal hydrolase deficiencies (metachromatic leukodystrophy, Farber, Fabry, Krabbe and Gaucher diseases). A substantial difference exists between the lysosomal storage in non-neuronal cells and neurons (see 3.1.4)\(^{49,50,58}\). This difference is related to predominant functionality of pSap protein in the particular cell type.

Morphological changes in non-neuronal cells (various types of epithelia or histiocytes) represent foamy cytoplasmic transformation by lysosomal storage. Electron microscopic findings include mixture of pleomorphic, sometimes oligolamellar lysosomal deposits, which might display some typical ultrastructural characteristics of the above mentioned isolated deficiencies. Lysosomal accumulation of multiple sphingolipids without contribution of either sphingomyelin or cholesterol is the predominant feature; lactosyl-ceramide is universally stored. Ceramide, glucosylceramide, galactosylceramide, globotriaosylceramide, sulphatide and GM1-3 gangliosides are accumulated in tissue dependent manner. The glycosphingolipid/sphingomyelin ratio or certain tissue specific storage patterns might serve as very efficient differential diagnostic tools\(^{49,50}\).

### 3.1.3 Comments and discussion to the publication Sikora et al. (2007)

This work demonstrates roles of pSap in human neurons and suggests extensive lysosomal sequestration of ubiquitin-tagged proteins most probably due to massive autophagy. In addition, it describes immense sensitivity of cortical neurons to pSap deficiency resulting in their death in the relatively short postnatal time period.

The work utilizes formaldehyde fixed paraffin embedded (FFPE) biological material from three previously described pSap-deficient infants\(^{49,50,56}\). This study represents detailed analysis of neural tissues in pSap deficiency employing optical and electron microscopic approaches.

#### 3.1.3.1 Specific features of neuronal affection in pSap deficiency

Neurons in all analyzed cases of pSap deficiency displayed cytoplasmic distension by fine eosinophilic granules, the changes attained granulovacuolar appearance in the oldest deceased patient (day 119). The granular-granulovacuolar appearance contrasted with foamy cytoplasmic transformation in the non-neuronal cells. There was no evidence of birefringent moieties, nor significant presence of glycolipids or sphingolipids. Glial elements (astrocytes and oligodendrocytes) displayed considerably milder morphological changes, microglial phagocytes are discussed later. Neuronal ultrastructural changes were featured by numerous vacuoles (0.5-1 \(\mu\)m in diameter, depending on the particular case) containing variably dense material with amorphous or coarse membranous substructure. Some of the vacuoles seemed to be endowed by double membrane. There was no substantial difference between central and peripheral neurons.

Observed changes were fundamentally different from the changes seen in non-neuronal cells in pSap deficiency or from changes observed in other lysosomal storage disorders characterized by membranous lipid deposits.

Immunophenotype of the neuronal granules was determined as late endosomal-lysosomal (LAMP 1, 2 and cathepsin D positive). Neurons in pSap deficiency displayed variable immunohistochemical positivity of subunit C of mitochondrial ATP synthase (SCMAS)\(^{59}\).
Novel and surprising finding in pSap deficiency was the uniform and strong granular immunopositivity of ubiquitin in the whole neuronal population including axonal spheroids. Such an extent of ubiquitination was not observed in non-neuronal cells affected by pSap deficiency. We evaluated not only the overall cellular ubiquitin (free and bound) but also ubiquitin protein conjugates (mono- and polyubiquitinated)\textsuperscript{60}, which were also considerably enhanced in presence. Aggresomal protein component - p62\textsuperscript{61} immunohistochemical detection was negative. We were not able to demonstrate comparably extensive and neurally selective ubiquitination in any other lysosomal storage disease besides some very specific situations (neurofibrillary tangles in Niemann-Pick disease type C or axonal spheroids).

We resorted to immunofluorescence multiple labeling in order to address the issues of ubiquitin positivity subcellular localization. The colocalization analysis was based on Pearson’s coefficient cross-correlation function (CCF) analysis of dual channel images\textsuperscript{62,63}. CCF analysis showed unambiguous colocalization (positive correlation) of ubiquitin signal with late endosomal-lysosomal markers. Unfortunately, we were not able to distinguish the prevalence of either of the two phenomena - ubiquitination of membranes or contents of the late endosomal-lysosomal compartment. The issue of direct demonstration of autophagy events by observing the organelar entrapment in lysosomal lumen, which would clearly complement electron microscopy, remained unanswered, most probably due to advanced stage of the intralysosomal epitope degradation.

Other important finding in pSap deficiency was the massive loss of cortical neuronal population in the early postnatal life as was demonstrated by comparison of the findings in the youngest deceased patient (day 27) and two other cases (days 89 and 119). This phenomenon does not seem to be related to improper functioning of subependymal germinal zone, neuronal migration or cortex remodeling, and it does not seem to be mediated via caspase dependent (apoptotic) cell death. As a consequence of such an extensive cellular loss, replacement of the neuronal cortical population by dense population of microglial phagocytes (CD 68 positive) and GFAP positive astrocytes occurred.

This ontogenetic stage dependent neuronal loss, most probably caused by the absence of pSap, occurs within a time period of 60-80 days in the early postnatal period. With respect to the extent and relative abruptness of the neuronal loss in pSap deficiency, and in order to express the limited potential of normally evolving neurons to survive, we designated it cortical neuronal survival crisis.

Neuronal affection in pSap deficiency displayed several fundamental characteristics that cannot be directly assigned to biochemical consequences of combined Saps activator deficiency. The true origin of these changes (massive ubiquitination, signs of autophagy and cortical neuronal survival crisis) is, in our opinion, reflection of the absence of pSap’s trophic and neuroprotective functions\textsuperscript{64-66}.

The true sequel of changes, as observed in pSap deficient neurons remains unanswered. The massive segregation of ubiquitin tagged proteins and observed ultrastructural changes can be compared to other processes previously described to result from macroautophagy\textsuperscript{67-69}. SCMAS detectability and ubiquitin-lysosomal colocalization further support autophagic origin of the changes\textsuperscript{70}. In addition, massive ubiquitination has been suggested to trigger autophagy under specific conditions\textsuperscript{28}. Despite the existence of respectable data about the link of pSap to apoptosis down regulation we were not able to assign the neuronal cortical crisis to caspase mediated cell death. We thus speculate that the impact of pSap deficiency and depletion of neuronal postnatal stimulation triggers massive protein ubiquitination and sequential enhanced autophagy.
3.2 Acid sphingomyelinase - Niemann-Pick disease type A/B – acid sphingomyelinase deficiency – model lysosomal enzymopathy – molecular studies

Niemann-Pick disease type A and B (NPD A/B, acid sphingomyelinase deficiency) represents a classical, prototypic lysosomal storage disorder caused by the deficiency of a single lysosomal hydrolase with SAPLIP characteristics - acid sphingomyelinase (ASM, MIM 257200 and 607616)71.

3.2.1 Sphingomyelin and sphingomyelinases

Sphingomyelin is a phospholipid composed of sphingenine, long chain fatty acid (together ceramide) and phosphorylcholine. Sphingomyelin comprises 5-20 % of the cellular phospholipids and is the essential constituent of the biological membranes, predominantly located in the outer bilayer leaflet. Sphingomyelinases catalyze the hydrolysis of sphingomyelin into ceramide and phosphorylcholine71. The classification proposed by Samet and Barenholz72 includes following mammalian enzymes: acid sphingomyelinase (ASM, E.C. 3.1.4.12) and its secretory Zn2+ activated variant (products of a single gene), neutral Mg2+ dependent sphingomyelinase, Mg2+ independent neutral sphingomyelinase and alkaline sphingomyelinase from the intestinal tract.

Sphingomyelinases act as soluble membrane associated proteins73. These enzymes carry out the catalysis on the surface of the lipid membrane bilayer and therefore, only the head group of sphingomyelin is predicted to be available for the substrate recognition (see 3.2.1.1).

Sphingomyelinases play crucial roles in the pathogenesis of some bacterial and viral infections (N. gonorrhoeae, P. aeruginosa) and have substantial implications to the apoptosis triggering by some natural or artificial cytotoxic agents or ionizing radiation74. Ceramide, a product of sphingomyelin hydrolysis, functions as a secondary messenger in a number of regulatory pathways including apoptosis, cell differentiation or proliferation.

3.2.1.1 Human ASM (E.C. 3.1.4.12)

ASM is a soluble metal ion (Zn2+) dependent, membrane associated SAPLIP glycoprotein with a pH optimum of ~ 4.5–5.5. Two forms of ASM have been described, one intracellular (lysosomal) and the other extracellular (secreted)75. The first (intracellular) form is targeted to lysosomes by means of sortilin receptor patway or mannose-6-phosphate receptor mediated transport and possessen acidic pH optimum of 4.5-5.075. The extracellular (secreted) form is released from the cell via default secretory pathway, its pH optimum is broader. Both forms of ASM are Zn2+ dependent, despite the differences in temporo-spatial coordination of the enzyme’s Zn2+ activation75.

The initial translated protein encoded by SMPD1 gene has 629 amino acid residues and consists of a signal peptide, predicted saposin domain, phosphoesterase domain containing predicted dimetal center, and carboxy terminal part.

Bioinformatic methods employing homology modeling based on structurally related mammalian purple acid phosphatase rendered a representative model of phosphoesterase domain of ASM. This model comprises residues 201-472 of ASM and provided valuable insights into the structure of this key ASM domain76 (see Figure 1). It suggests structural model of phosphorylcholine head group recognition site and establishes sound basis for the
catalytic mechanism of ASM mediated sphingomyelin hydrolysis, assigning the critical function of proton donation to His319 residue.

The substantial necessity of saposins (see 3.1.1) for sphingomyelin hydrolysis by ASM has not been unequivocally demonstrated. SAP-like domain might function in the mobilization of the substrate by disturbing the target membrane structure, as well as, by maintaining the folding stability of the whole enzyme\textsuperscript{40}.

### 3.2.2 Molecular genetics of ASM (SMPD1 gene)

ASM is coded by \textit{SMPD1} (sphingomyelin phosphodiesterase) gene\textsuperscript{77-79}. The overall size of the \textit{SMPD1} gene is \textasciitilde{}5kb including 5′ and 3′ untranslated regions, it has six exons (sized 77-773 bp) and five introns (sized 153-1059 bp). Only the full length (2347bp) mRNA encodes catalytically active enzyme. The open reading frame (ORF) of this full length mRNA is 1890 bp long and codes for 629 amino acids-long protein.

Only recently, results of the studies on Beckwith-Wiedemann syndrome (BWS) demonstrated genomic imprinting of the \textit{SMPD1} locus\textsuperscript{80}. These recent findings shed new light onto the previously described phenotypic variability of otherwise genotypically identical heteroallelic individuals with NPD A/B or possible phenotypic affection of NPD A/B carriers\textsuperscript{81}.

### 3.2.3 Cellular and organ pathology of ASM deficiency, consequences of the storage process

Basic morphological change characterizing the lysosomal storage process on the optical and electron microscopic levels is gradual and three dimensional distension of the cellular lysosomal compartment\textsuperscript{71,82,83}. The final stage of this process could be described as foamy transformation of the cell. Current state of knowledge considers this kind of morphological transformation as nonspecific.

Lysosomal storage caused by ASM deficiency (NPD A/B) is generalized, multicellular type affection. ASM deficiency is generalized sphingomyelinosis. The massive and generalized accumulation of sphingomyelin distinguishes NPD A/B from NPD type C (disorder with a completely different molecular basis, see 3.2.7.3.1), that is characterized by sphingomyelin accumulation in histiocytic elements; all other cell types in NPD type C accumulate a mixture of lipids, predominantly glycosphingolipids and cholesterol\textsuperscript{84}.

The consequences of the lysosomal storage process on the normal cellular functions are still controversial (see section 1.2). The fundamental questions regarding the dynamics and turnover of the storage compartment and accumulated substrate remain unanswered.

#### 3.2.3.1 Neuronal storage compartment

The extent of neuropathological affection can be quite variable and includes both central and peripheral nervous system. Affected neurons have distended foamy cytoplasm with detectable sphingomyelin birefringent crystals. The general tendency to atrophy and gliosis is a result of neuronal depopulation due to lysosomal storage and parallel proliferation of astroglia. Similar changes might be detected in the spinal cord, sympathetic ganglion cells or in adrenal medulla. There is sometimes detectable peripheral demyelinization with lysosomal storage in Schwann cells.
Retina is a complex histological structure with extensive neuronal population. It is easily accessible to noninvasive fundoscopic evaluation, and as such, retinal changes are considered important for the evaluation of neuronal lysosomal storage, not only in ASM deficiency. The cellular types affected by lysosomal storage in retina, however, have not yet been clearly defined. Grey-yellowish haze of retina appears when lipid deposits accumulate in the somata of ganglion cells. Fovea, under the conditions of lysosomal storage, contrasts in red with the surrounding yellowish retina.

3.2.3.2 Visceral storage compartment

Spleen is the most affected organ in ASM deficiency. Splenomegaly might be very variable, but 10 fold increases in overall size are not exceptional. Splenic architecture is considerably abnormal with the storage present especially in the histiocytes of both pulps. The intensity of the storage process in the hepatic lobule exhibits porto-central vectorial increase. In some very mild cases lysosomal distension in the hepatocytes can be hardly discernible on the optical microscopy level. Kupffer cells are always affected, even though to variable extent, suggesting their higher sphingomyelin turnover. Protracted storage process might also result in deposition of autofluorescent ceroid. Correlation between the extent of enzymatic ASM deficiency and histological tissue affection exists, but only to a limited extent.

Bone marrow and lymphatic nodes display variable admixture of foamy macrophages, sometimes with considerable amount of ceroid (sea-blue histiocytes) in protracted cases. Pulmonary parenchyma (alveolar septa, alveolar space) can be infiltrated by foamy cells of macrophagic origin, especially in visceral storage limited cases. Adrenal glands (especially cortex – epithelial cells and infiltrating macrophages) and thymus are enlarged. Other organs may exhibit mild storage changes but usually no substantial functional impairment.

3.2.4 Clinical classification of ASM deficiency, phenotypic heterogeneity versus historical perspective

The first step in pathogenesis-based classification of Niemann-Pick disease was taken, when type C of NPD was declared a separate pathogenic entity (see 3.2.7.3.1). Ever since, NPD due to ASM deficiency is described to have two distinct phenotypic variants – types A and B.

The neurovisceral storage associated form, type A, is characterized as rapid and fatal clinical phenotype with progressive neurodegeneration and death within the first three years of life. Type B is, on the other hand, characterized by visceral-only (non-neuronal) affection with mild, slowly progressive clinical course and survival till adulthood. Sphingomyelin lysosomal storage is, so far, the only accepted cause of clinical symptomatology. Neuronal lysosomal storage results in neurological affection and non-neuronal storage results in visceral symptomatology. It slowly became apparent that ASM deficiency displays wide phenotypic variability both in the neurological and visceral symptomatology, and typically affected patients (A or B phenotypes as defined) seem to represent a minor fraction of all ASM deficient patients.

Neurological affection ranges from overt perinatal crises, severe psychomotor retardation, severe muscle dysfunction or irritability to milder presentations such as peripheral neuropathy, mild mental retardation with learning disabilities, or only macular halo syndrome without additional abnormal neurological findings. Fundoscopic changes (cherry red macula...
or macular halo) might serve as direct and easily accessible indicators of neurolysosomal storage (see 3.2.3.1)\(^8\). Visceral affection was also repeatedly demonstrated as very variable. Liver failure due to massive hepatocellular storage in early years of life, or extensive pulmonary infiltration by storage laden macrophages contrast with cases featured by minute lysosomal storage in sessile tissue histiocytes and minimal organ dysfunction\(^8\).

It seems to be clear that clinical course (temporal and tissue related) and storage intensity are definitely not classical type (A or B) specific, and the accepted classification scheme oversimplifies the biological problem of ASM deficiency. Whatever modification of the current classification scheme is going to be achieved, it must reflect the immense variability of the clinical phenotype.

### 3.2.5 Molecular pathology of the \textit{SMPD1} gene, population genetics, genotype-phenotype correlations, implications of \textit{SMPD1} gene variations for ASM protein malfunction

NPD A/B is inherited as an autosomal recessive trait\(^7\). There are, at least, 99 different pathogenic sequence variations reported in the \textit{SMPD1} gene\(^9\). There are no reports demonstrating large scale rearrangements in the \textit{SMPD1} gene, but it should be repeated that the chromosomal region 11p15 can be a subject of imprinting\(^8\).

ASM deficiency, in all its phenotypic variants, should be regarded as a truly panethnic disorder with global distribution. Certain ethnic groups with tendency to socio-economic clustering and higher rates of consanguineous marriages show higher incidence of ASM deficiency (Ashkenazi Jews\(^9\), Israeli Arabs\(^9\) or Arabic population of North African region of Maghreb\(^9\)). The panethnic population frequency of NPD A/B including heterozygote carriers is not available despite the existence of demographically very heterogeneous series of genotyped patients\(^9\). Last fifteen years represent time of intensive genotype-phenotype correlation studies in the field of lysosomal storage disorders. General limitations of these studies (compound heterozygosity, heterogeneous phenotype and phenotypic overlaps, quality and relevance of evaluated phenotypic traits) can also be applied to NPD A/B.

Despite these drawbacks several pathogenic variations were clearly correlated with certain phenotypic variants of ASM deficiency (R496L, L302P and fsP330 are classical NPD type A\(^9\),99,100 pathogenic variations in Ashkenazi Jewish population, delR608 is panethnic classical type B associated variation even in compound heterozygotes\(^9,98,101\)). With respect to the above mentioned ongoing discussion on the matter of ASM deficiency phenotypic classification and relative paucity of classical NPD phenotypes (A and B) among affected patients, pathogenic variation Q292K\(^8,102,103\) associated with slowly progressive neurovisceral affection must be noted (see 3.2.7.2).

The impact of \textit{SMPD1} pathogenic variations on the ASM protein structure and function, evaluated by means of the residual activity, is variable. In certain aspects, residual ASM degradation capacity attained by loading the cells with labeled sphingomyelin substrate may be more informative than the expressed or natural residual activity measurements (see section 3.2.7.2).

ASM phosphoesterase domain model based on homology modeling provides very interesting platform to hypothesize about the impact of pathogenic variations on the ASM protein structure and function\(^7\). Based on this premise, pathogenic variations M382I and N383S interfere with substrate recognition, L302P might result in kinking of one of the \(\alpha\) helices and thus decreases protein stability, and W391G variation may destabilize the protein by disrupting important hydrophobic interactions. It should be noted that only one of \textit{SMPD1}
gene variations (D278A, see also 3.2.8.2) occurs in the predicted metal coordinating residues.

3.2.6 NPD A/B diagnostics – histopathology, enzyme biochemistry, molecular genetics, differential diagnosis, and therapy

The identification of the lysosomal storage of sphingomyelin can be usually performed by histochemical detection of sphingomyelin liquid crystals or additional ceroid admixture in the foamy cells. The measurement of activity of ASM is very efficient method to detect ASM deficiency; it also serves as a potent differential diagnostic tool to exclude other lysosomal storage disorders. There are several methods for ASM estimation, differing in the substrate used for the analysis (radiolabeled N-methyl-$^{14}$C sphingomyelin or chromogenic 2-N-(hexadecanoyl)-amino-4-nitrophenyl phosphorylcholine (HNP-PC) or 6-hexadecanoylamino-4-methylumbelliferyl-phosphorylcholine (HMU-PC)). It must be noted, that considerable care must be taken for the presence of contaminating neutral sphingomyelinase activity while performing and evaluating ASM activity assays.

The values of residual ASM activity in affected individuals, heterozygotes or healthy individuals are not strictly separated. Classical NPD type A patients do not express more than 2% of normal ASM activity. Classical NPD type B patients usually exhibit values of 5-10% of normal ASM activity. Inconclusive grey-zone values of ASM residual activity must be interpreted very cautiously and with regard to additional findings, including molecular genetic testing.

Additional method of evaluation of ASM degradative capacity is the use of LDL or liposome mediated sphingomyelin loading test in the cell culture with consecutive sphingomyelin degradation products determination. This type of assay setup might characterize the residual sphingomyelin degradative capacity of the lysosomal apparatus in a more “biological” way than the values of residual ASM activity obtained from cell homogenates.

The ultimate diagnostic tool of NPD A/B is the molecular genetic evaluation of the SMPD1 DNA sequence. Analysis of RNA may be useful in certain special cases (i.e. nonsense variations, intronic variations affecting consensus splicing factors binding sites etc.). Epigenetic modifications (i.e. imprinting of the SMPD1 locus) should not be ignored.

The differential diagnostics of NPD A/B require broad clinical, histopathological and laboratory considerations. The complexity of the problem is demonstrated in the publication – Dvořáková et al. (2006).

Goal of principal correction of the enzymatic deficiency in ASM deficiency has not been achieved yet. Therapeutic modalities being considered are: enzyme replacement therapy (ERT) or substrate reduction therapy, bone marrow transplantation or other means of stem cells therapy in conjunction with gene therapy.

Each of the above mentioned therapeutic approaches to ASM deficiency is facing multiple biological obstacles – generalized character of affection, relative pharmacokinetic inaccessibility of certain tissue compartments (especially central nervous system), phenotypic variability hampering efficient indication protocols, and economic demands for health care systems.

ASM deficiency is one of the candidates for enzyme replacement therapy (ERT). Recombinant acid sphingomyelinase has already entered advanced stages of clinical testing and should be available to NPD type B patients in a short time.
3.2.7 Comments and discussion to the publications Sikora et al. (2003), Pavlů-Perreira et al. (2005) and Dvořáková et al. (2006)

3.2.7.1 Sikora et al. (2003)

Publication by Sikora et al. (2003) describes a molecular study of \textit{SMPD1} pathogenic variations in a cohort of 7 NPD A/B patients provided by cooperating Dutch center involved in diagnostics of inherited metabolic disorders. The goal of the study was to evaluate genotype-phenotype correlations in this patient group.

We analyzed seven non-related NPD A/B patients (out of 27 diagnosed in the Netherlands in the years 1970-1999), whose diagnosis of ASM deficiency was based on previous ASM activity measurements using N-methyl-$^{14}$C sphingomyelin substrate. Three of the patients were of Dutch ethnic origin, the remaining four patients, even though Dutch residents were of Turkish origin. The clinical phenotype corresponded with the classical type A of NPD in three patients, one patient was described as slower but progressive type A patient (died at 5 years of age), and the remaining three patients were affected by typical type B phenotype of NPD.

The overall evaluation of the coding region of \textit{SMPD1} gene was performed by direct sequencing of the PCR products. The analysis of fourteen alleles disclosed eight different pathogenic variations, seven (G29fsX74, S248R, H319Y, P371S, F463S, P475L, and Y537H) of these variations were novel, at that time. Novel pathogenic variations were assessed in additional 100 wild type \textit{SMPD1} alleles to exclude the possibility that they represent common genetic polymorphisms.

The fundamental question regarding the pathogenic potential of the novel variations was based on the following premises: (i) no other nucleotide change, besides the common polymorphisms, was found in the \textit{SMPD1} coding sequence; (ii) all of the variations result in a non-conservative amino-acid exchange, except for P475L; (iii) one of the variations (G29fsX74) is a micro deletion with a reading frame shift resulting in a premature stop codon in the ASM ORF (open reading frame); (iv) all of the affected amino acid residues display broader evolutionary conservation; (v) found variations are not common polymorphisms; (vi) extensive searches in EST (expressed sequence tags) databases did not render a single positive hit when searched for these variant sequences.

Out of seven evaluated individuals five were homozygotes (all four Turkish patients and one Dutch patient) for respective pathogenic variations - H319Y, P371S, Y537H and delR608. H319Y and Y537H associated with NPD type A and P371S associated with NPD type B (see below). At the time of publication (year 2003), delR608 was considered to be prevalent pathogenic variation among the Arabic NPD type B population of North African region of Maghreb. The publication by Sikora et al. (2003) was one of the first suggesting that delR608 occurrence is panethnic, this suggestion was later confirmed by additional studies in the ethnically heterogeneous NPD type B patient cohorts\textsuperscript{91}. The genotype-phenotype correlation between delR608 and mild NPD type B phenotype without any neurological affection is an accepted fact of immense significance. With respect to the planned introduction of ERT into the therapy of ASM deficiency, genotype-based indication criterion (delR608 positivity) seems to be valid for inclusion of the affected pathogenic variation carrier into the therapeutic protocol.

The critical residue donating the proton for the sphingomyelin hydrolytic cleavage was suggested to be H319\textsuperscript{78}. One of the supportive arguments provided for this hypothesis is the
homzygous occurrence of H319Y in one of the above described NPD type A patients. On the contrary, P371S variation found in one of the NPD type B patients is predicted to be situated at the end of a β-strand between the two central β-sheets, but is not in the close vicinity of the catalytic center. Variant serine at this position might result in the slight change of the two subdomains spatial orientation and thereby affect the enzyme’s catalytic activity. Topology of the global phosphoesterase domain of ASM with the discussed residues is depicted on Figure 1.

The knowledge, if only of a bioinformatic, but highly representative protein structural prediction, remarkably extends the potential to reasonably hypothesize about the pathogenic variation impact on the protein structure and function. In the same logic, the information about pathogenic variation primarily generated for medical purposes might provide important clues for protein catalytic functions.

3.2.7.2 Pavlů-Perreira et al. (2005)

Publication by Pavlů-Perreira et al. (2005) represents an extensive multi parameter study of 25 NPD A/B patients from the region of former Czechoslovakia. This respectable cohort of patients was evaluated from different perspectives (clinical with special emphasis on the neurological and ophthalmologic affections, histopathological concentrating on the neuropathology and liver presentation of ASM deficiency, biochemical including sphingomyelin degradation capacity in cell-loading assays, and molecular genetic analysis of SMPD1 gene sequence). The complexity of the gathered data allowed extensive conclusions towards the genotype-phenotype correlations. The robustness of the studied patient cohort also provided sound basis to critically address the current classification of ASM deficiency.

3.2.7.2.1 ASM deficiency phenotypic variability and related implications

Cardinal observation in this work is related to the fact that classical NPD phenotypes comprised minor proportion of all evaluated patients (9 of 25 cases, 36%); rest of the patients (16 of 25 cases, 64%) displayed clinically variable phenotype that could not have been classified according to the currently accepted criteria. This phenotypic variability can be perceived as continuous spectra of severity of neurological or visceral symptomatology (or their overlaps).

It is especially the high proportion of patients with variable protracted but progressive neurological affection and associated visceral storage that makes this cohort unique. It must be noted that not only the neurological affection, but also the visceral affection displayed variant phenotypes among the evaluated patients.

According to the point taken in the paper, retinal lesions without any additional neurological affection may be considered as the most benign feature of neuronal lysosomal storage in ASM deficiency. Even though most benign feature, it is still a sign of neuronal lysosomal storage.

3.2.7.2.2 SMPD1 genotype in central Europe, implications for ASM enzymatic activity assays

In comparison to the publication by Sikora et al. (2003), genotype evaluation of the ASM deficient patients from former Czechoslovakia revealed one prevalent pathogenic variation - Q292K. This variation was demonstrated to clearly associate with slowly progressive neuronopathic phenotype of ASM deficiency when present on both SMPD1
alleles in one patient. Phenotypic presentation of Q292 K compound heterozygotes relies on the second contributing allelic variant.

Evaluation of the position of Q292 residue in the proposed structural model of phosphoesterase ASM domain demonstrates its location in one of the peripheral $\alpha$-helices (Figure 1). The impact of the substitution can be assumed in a very limited way. The role of the critical $\alpha$-helix is not easily predictable, and consequences of its kinking by introduction of Lys residue are unclear.

3.2.7.2.3 Proposals to the classification of ASM deficiency

Considering the phenotypic variability in the evaluated group of patients, the current classification of ASM deficiency is biologically restrictive. Dramatic change in the phenotypic perception of the ASM deficiency cannot be expected, especially when considerable number of authors in the field tends to adhere to historical connotations.

The proposal raised in our paper suggests inclusion of all variant phenotypes of ASM deficiency (neuronopathic and non-neuronopathic) into a category of “intermediate phenotype”. More importantly, the publication tried to express the necessity to perceive the existence of broad phenotypic variability of ASM deficiency by deeming that both the speed of the clinical course and storage intensity are not classic type specific.
Figure 1

**Structural model of ASM phosphoesterase domain**

A – Global representation of the predicted crystal structure of ASM phosphoesterase domain with twofold symmetry, each half containing βαβαβ structural motif (α-helices in red, β-sheets in green, loops in grey). Two Zn²⁺ ions are shown in violet; phosphorylcholine (substrate) is superposed into the protein model in orange. Dimetal coordinating residues (206, 278, 318, 425, and 457) are displayed in light pink. Predicted disulfide bond formed by Cys 385 and 431 (deep red) is in direct contact with the substrate recognition loop. Calculated molecular surface of the residues (382-390, light blue) involved in substrate recognition was rendered in transparent green color. Three discussed variant residues (Q292, H319, and P371) are presented in yellow. Residue Q292 lies outside of the predicted catalytic site of ASM phosphoesterase domain in one of the predicted short peripheral α-helices. Residue P371 is located at the C-terminal end of one of the central β-sheets; the potential impact to the protein structure is discussed in the text.

B – Detailed view of the catalytic site and substrate recognition loop. Coloring is identical to part A of this Figure. Arrow signifies the calculated H-bond between the proton donating H319 (yellow) and phosphorylcholine substrate (orange). Residue H319 is the predicted active residue in the catalytic site of ASM. The presented structure was downloaded from (PDB protein crystal structure database) and corresponds to the original 1X9O.pdb file. The images were rendered in DeepView/Swiss PDB Viewer software. Image pairs represent stereo views with rotation angle of ±2° from the central axis.

3.2.7.3 Dvořáková et al. (2006)

Differential diagnostic pitfalls of a slowly progressive lysosomal storage disorder with minute histological changes are demonstrated in the publication by Dvořáková et al. (2006). It also provides new insights into the phenotypic variability of Niemann-Pick disease type C. The discussion related to this article deals primarily with variant adult Niemann-Pick disease type C and its differentiation from ASM deficiency.
The report describes a 53-year-old female patient without significant medical history, who succumbed to cardio-respiratory failure after encountering an attack of pulmonary embolism. The autopsy disclosed previously asymptomatic hepatosplenomegaly and lymphadenopathy.

The microscopic affection was dominated by the presence of histiocytic foam cells, containing considerable amount of ceroid pigment. Maximum of these cells was found in the lymph nodes and in the liver. The lysosomal system seemed to be morphologically activated (distended) in hepatocytes and neurons, but lysosomal storage was on the verge of detection by means of optical microscopy. The initial diagnostic suggestion of Gaucher disease was rejected on the basis of histopathological changes, i.e. on the basis of absence of Gaucher cells (storage histiocytes with typical cytology). The described changes seemed to correlate with the findings in slowly progressive ASM deficiency (classical NPD type B, see section 3.2). Unfortunately, FFPE tissues available for analysis hampered in-situ detection of sphingomyelin liquid crystals.

ASM activity values assessed in obligate heterozygotes (mother and two children of the proband) were inconclusive. At this point, the only possibility to exclude or confirm ASM deficiency in the proband was to perform DNA analysis of the sequence of the \textit{SMPD1} gene. The DNA sequence analyses in obligatory heterozygotes rendered wild type \textit{SMPD1} sequence. At this stage, diagnostic considerations shifted into the category of wishful thinking and to previously rejected explanation of the findings by assigning them to extremely rare variant of adult onset visceral form of Niemann-Pick disease type C (NPC).

### 3.2.7.3.1 Niemann-Pick disease type C (lysosomal storage disorder due to defective intracellular lipid trafficking)

Lysosomal storage in NPC does not represent generalized sphingomyelinosis even though sphingomyelin is partly accumulated. The lysosomal storage in NPC has specific characteristics in different tissues and cell types, the most extensively affected cell types are monocytes/macrophages and neurons. Spleen and liver (mostly macrophages) accumulate predominantly unesterified cholesterol and sphingomyelin, and a comparably smaller proportion of bis(monoacylglycerol) phosphate, and glycolipids (glucosylceramide, lactosylceramide, and GM₃ gangliosides). In neurons, the ratio is reversed with predominance of glycolipids and minute accumulation of unesterified cholesterol and sphingomyelin.

The search for the fundamental molecular defect in NPC culminated in its causative association with one particular gene (\textit{NPC1})\textsuperscript{107}. Soon after the discovery of \textit{NPC1} gene, it became clear that there exists another complementation group in NPC phenotype, resulting in discovery of the second critical gene (\textit{NPC2})\textsuperscript{108}. The overall ratio of these two complementation groups is 95\% (\textit{NPC1}) and 5\% (\textit{NPC2}).

Findings that both protein products (NPC1 and NPC2) are directly involved in intracellular lipid (especially cholesterol) trafficking and homeostasis provided sound basis for the previous hypotheses proposed at time of NPC exclusion from ASM deficiency phenotype\textsuperscript{109,110}. NPC1 is a multi-transmembrane protein residing in the membranes of a specialized subset of late endosomal-lysosomal compartment (NPC1+/LAMP2+/M6PR-)\textsuperscript{111,112}. This specific membrane endowed subcompartment represents extremely dynamic intracellular structure that lively communicates with its cellular partners (especially Golgi apparatus and other constituents of endosomal-lysosomal compartment). NPC1 protein possesses a predicted sterol sensing domain on one of its extra membranous loops\textsuperscript{110}. Interpretation of the current understanding of the functions of NPC1 protein and NPC1 positive cellular compartment is the regulation of intracellular lipid sorting, trafficking, and
targeting. It is not only cholesterol, but also other lipid moieties (e.g. glycolipids) that are regulated by NPC1111.

NPC2 protein is, unlike NPC1 gene, soluble late endosomal-lysosomal luminal protein, functioning in cholesterol binding109. The fundamental principle of its function, similar to NPC1 protein, has not yet been fully elucidated. Despite that, it can be postulated that NPC1 and NPC2 proteins function in a non-redundant functional coordination, whether by direct interaction or successive action113 is not known.

Phenotypic presentation of Niemann-Pick disease type C is very variable. The phenotype almost always includes hepatosplenomegaly and neurological impairment84,114. Phenotypic presentations can be subdivided into perinatal (dominated by overt hepatopulmonary failure with cholestasis), infantile, classical juvenile and adult forms. It is especially the nearly omnipresent neurological affection that adopts extreme phenotypic variability from severe psychomotor retardation in infantile forms to presentation by psychiatric symptoms (e.g. bipolar psychosis) in adult NPC variants. Up to date, only two patients with visceral only NPC without neurological affection, presenting by isolated hepatosplenomegaly in their fourth and sixth decades, were reported115,116.

The diagnostics of NPC are, in principal, similar to ASM deficiency. Special emphasis should always be given to the non-generalized sphingomyelin storage and the histological distribution of accumulated cholesterol. The basis of NPC diagnostics is filipin staining test (increased in NPC) in cultured fibroblasts, and evaluation of the rate of LDL-derived cholesterol esterification (decreased in NPC)84. Delineation of the variant genotype on the DNA sequence level is the final step in the diagnostics.

In order to test the hypothesis that our patient represented third case of the extremely rare visceral only NPC affection in adult age, we resorted to DNA sequence evaluation in the obligate heterozygote relatives. Both NPC1 and NPC2 genes were evaluated. All three evaluated relatives were found to be heterozygous for either previously reported NPC1 pathogenic variation (S666N) or a novel pathogenic variation (N961S). Both variations were also found in the proband’s NPC1 DNA sequence isolated from FFPE tissues.

These results allowed us to conclude that this patient suffered from visceral only NPC without manifest neurological affection at the age of 53 years.

4 Bioinformatic predictive study of acetyl-coenzyme A:α-glucosaminide N-acetyltransferase - Hřebíček et al. (2006)

Publication by Hřebíček et al. (2006) describes a genetic linkage analysis performed in several families affected by mucopolysaccharidosis type IIIc (Sanfilippo type C syndrome, MPS IIIc, see sections 1.1 and 1.2). The main conclusion of the study is that MPS IIIc results from pathogenic variations in the gene designated TMEM 76, located in the pericentric region of chromosome 8. The critical TMEM 76 gene was designated as HGSNAT (heparin acetyl-coenzyme A:α-glucosaminide N-acetyltransferase), this designation is an official gene name7.

HGSNAT protein (663 amino acids, 73 kD), which displays considerable degree of instability under in-vitro conditions, was previously demonstrated to reside in the DRMs of lysosomal membranes10. HGSNAT primary function is to transfer acetyl moiety to heparan sulfate and render it available for consecutive intralysosomal degradation by α-N-acetyl glucosaminidase. Deficiency in this, otherwise synthetic, step (acytlylation) results in the accumulation of heparan sulfate in the lysosomal compartment. The molecular mechanisms of HGSNAT catalytic function are still enigmatic17,118.
The immense growth of the genomic information in the last decades due to availability of efficient and high throughput methods of nucleic acids analyses, resulted in the expansion of bioinformatic tools employing the “DNA sequence domain” to predict protein properties and their cellular functions\(^\text{119}\).

The only confirmed protein (both by genomic and proteomic means) present in the available public databases was murine ortholog of human HGSNAT. The only other proteins homologous to HGSNAT belong to COG4299 bacterial protein family (uncharacterized), a cohort of protein predictions from a broad range of bacteria. Functions of these proteins are only inferred, but their roles in transmembrane acetylation and glycosaminoglycans metabolism cannot be excluded.

Syntenic aligning is an approach that is capable to document evolutionary conservation and is based on the use of available genomic DNA data that are fully experimental and thus not based on prediction. These methods allow alignments of long (in order of tens of kb to Mbs) DNA sequences to evaluate the extent of similarity among them. The two most widely used algorithms are VISTA and PIPMaker\(^\text{120,121}\). The results of multiple sequences VISTA alignment of chromosomal regions syntenic to human chromosome 8, carrying \textit{HGSNAT} gene, are shown on Figure 2. The major advantage of this kind of sequence comparison lies in its capability to efficiently predict conserved non-coding sequence stretches (functionally important, CNS) in the syntenic DNA sequences. Figure 2 depicts considerable evolutionary conservation of the coding (exonic) sequence in the \textit{HGSNAT} locus, besides that, it shows CNS regions that might be involved in either expression, RNA splicing or other regulations.

In order to evaluate previously proposed transmembrane (TM) nature of HGSNAT protein, its primary sequence was searched for membrane spanning domains using TMMOD algorithm\(^\text{122}\). The resultant prediction suggests 11 TM domains. The critical issue of TM orientation (cytosolic vs. non-cytosolic) was addressed, and assigned the C-terminal end of the protein to face the cytosol. For the orientation of the TM domains refer to Figure 3. In addition, an attempt was made to predict posttranslational modifications (PTM) using an integrated www interface\(^\text{123}\). HGSNAT primary sequence was evaluated for the presence of potential signal peptide (SignalIP, see Figure 3) and N-linked glycosylation sites (NetNGlyc). As demonstrated on Figure 3, there are at least four potential N-glycosylation sites present in the protein sequence, all four sites are predicted to face lysosomal lumen, which is a finding in accordance with the concepts of lysosomal membrane protein properties.

The presented results demonstrate advantageous use of bioinformatic approaches to predict protein structure, PTMs and function, even though the only available data is the primary protein sequence and the number of homologous genes is small, and there are no structural data available for homology modeling.
Figure 2
Syntenic alignment of appropriate HGSNAT chromosomal regions

Alignment of multiple syntenic sequences of the critical pericentric region of chromosome 8, annotation based on Ensemble database. The alignment was performed using mVISTA algorithm provided by Berkeley genome pipeline. The aligned sequences are listed below the image and follow the order given (top line represents human HGSNAT gene with annotated 5'-3' orientation, Pt - *Pan troglodytes*, Macm - *Macaca mulatta*, Cf - *Canis familiaris*, Bt - *Bos taurus*, Rn - *Ratus norvegicus*, Mm - *Mus musculus*, Gg - *Gallus gallus*, Xt - *Xenopus tropicalis*, Md - *Monodelphis domestica*). The length of the aligned sequences ranged from 27 to 97 kb.

The conserved exonic sequences are depicted in light blue, coding non-conserved sequences (CNS) are shown in pink. Exon sequences are generally conserved from humans to frog (*Xenopus tropicalis*). Regions of CNS are located in the promoter regions, of note are CNS sequences in the vicinity of alternatively spliced exons 9 and 10. For the details about the alignment construction see text.
Figure 3

ClustalW\textsuperscript{125} alignment of human and murine HGSNAT (TMEM76) protein sequences.
Conserved amino acid residues are shown in yellow, non-conserved residues are in light blue. Predicted signal sequence in the human protein is in red (signal protein cleavage site at the residue 47). Predicted N-glycosylation sites are in violet. Transmembrane domains are underlined and in bold letters. The residues in \textit{italics} represent predicted cytosol facing sequence stretches.

5 \textit{Caenorhabditis elegans} as a model organism for selected lysosomal storage diseases

There are numerous animal models available for the lysosomal storage disorders. Some of them are naturally occurring variants, but a considerable number of them were prepared \textit{in-vitro} by different genetic manipulations. In general, mammalian species are advantageous for the study of human pathology states, but there are several established non-mammalian model organisms available as well. Nematode \textit{Caenorhabditis elegans} (\textit{C. elegans}) is one of these. Unfortunately, the criterion of biological-evolutionary similarity has its limits with respect to laboratory utilization. \textit{C. elegans} is a respected model organism for concerted genetic, ultrastructural and behavioral studies\textsuperscript{126}. \textit{C. elegans} was the first of complex multicellular eukaryotic organisms for which the complete genomic sequence became known. About thirty-six percent of the genes are homologous to human genes, including those implicated in human pathology\textsuperscript{127,128}. RNA-mediated interference (RNAi) technique allows inhibition of \textit{C. elegans} genes expression at the level of the whole organism with the ease unavailable in any other eukaryotic organism\textsuperscript{129,130}. 

\begin{tabular}{|c|c|c|c|}
\hline
\textbf{H. sapiens} & (1) & MTG---ARASAAEQRRAGRSGQARAAERAAGMSGAGRALAALLLAASVLSAALLAPGGSSGRDAQAAPPR & 70 \\
\textbf{M. musculus} & (1) & MTGGSSSRRRRAEERSSAAGTERNSRREAVGGMGAGPALAALLLAGSVLSATLLAPG--------RRAEP & 71 \\
\hline
\textbf{H. sapiens} & (68) & DLDKKRHAELKMDQALLLIHNELLWTNLTVYWKSECCYHCLFQVLVNVPQSPKAGKPSAAAASVSTQHGS & 140 \\
\textbf{M. musculus} & (63) & DLDEKRNVELKMDQALLLIHNELLGTSLTVYWKSDDCYQCTFQPLANVSHGGKPAKPSVAPVSVSTQHGS & 141 \\
\hline
\textbf{H. sapiens} & (138) & VSIAFLIGLAVIIV & 210 \\
\textbf{M. musculus} & (133) & VSIAFLVGLALIVA & 211 \\
\hline
\textbf{H. sapiens} & (208) & ISFL & 280 \\
\textbf{M. musculus} & (203) & VSLL & 281 \\
\hline
\textbf{H. sapiens} & (278) & ALILMVFVNYGGGKY & 350 \\
\textbf{M. musculus} & (271) & ALVLMVFVNGGGKY & 351 \\
\hline
\textbf{H. sapiens} & (348) & ILQVNSTSEERAACRLEYKFGEFGNYSLLVQHASSGANKIACDIIVNENPVDSNLP & 420 \\
\textbf{M. musculus} & (341) & ILQVIVNFNYCL & 421 \\
\hline
\textbf{H. sapiens} & (418) & HFWQ & 490 \\
\textbf{M. musculus} & (411) & HFWQ & 491 \\
\hline
\textbf{H. sapiens} & (488) & LLICIGIIIVNPNYCL & 560 \\
\textbf{M. musculus} & (481) & LLICIGIIIVNPNYCL & 492 \\
\hline
\textbf{H. sapiens} & (558) & WPQ & 630 \\
\textbf{M. musculus} & (551) & WPQ & 493 \\
\hline
\textbf{H. sapiens} & (628) & KDNQSHKEH & 666 \\
\textbf{M. musculus} & (621) & KDNQSHKEH & 631 \\
\hline
\end{tabular}
5.1 C. elegans, general characteristics

*C. elegans* is a small (~1 mm long), translucent, soil nematode feeding primarily on bacteria and reproducing with a life cycle of about 3 days. Fundamentals of *C. elegans* genetics were established by Sydney Brenner in the 1970s and are valid up to now\(^{131}\). Very efficient transgenic techniques have been developed for *C. elegans*. It is easy to follow the expression of transgenes (GFP or X-Gal linked), single cell ablations can be performed very efficiently\(^{132}\) and the laboratory maintenance of the nematode cultures is also easy.

The sequence of the complete *C. elegans* genome, including 15kb of mitochondrial DNA, was determined and assembled by the concerted efforts of the *C. elegans* sequencing consortium\(^{128}\). The 97 Mb genome is very compact with an average density of one gene per 5kb. *C. elegans* genome includes (January 2007) 7821 confirmed genes, 10741 partially confirmed genes, and 4650 predicted ORFs. Complete annotated genomic sequence of *C. elegans* is freely available via www database\(^{133}\) (Wormbase).

There are more protein similarities between *C. elegans* and *H. sapiens* than with any other multicellular organism (with the exception of other nematodes, typically *Caenorhabditis briggsae*), with the average predicted protein similarity of 36%. Many conserved genes, exhibit immense conservation, in some proteins as high as 97 %\(^{128,134}\). The identification and comparison of orthologs of human proteins in the nematode shows interesting clues as to the respective functions of proteins in the two species.

In *C. elegans*, the introduction of double-stranded RNA (dsRNA) results in the specific and potent inactivation of genes containing homologous sequences\(^{129,130}\). Gene-specific loss-of-function results from the degradation of endogenous mRNA by a set of consecutive molecular steps\(^{135}\). Studies involving RNAi have shown that it is possible to initiate RNAi by injection of dsRNA, however, as the RNA can be absorbed through the nematode’s gut, the effect was also observed in animals soaked in dsRNA solution and also in worms fed with Escherichia coli expressing double-stranded RNA\(^{136}\). Not surprisingly, RNAi was quickly embraced as a tool for determining the functions of specific genes not only in *C. elegans*.

5.1.1 Endosomal - lysosomal system in *C. elegans*

The protein constituents of the cellular system involved in endocytosis, sorting, transport, substrate degradation and redistribution are well conserved in *C. elegans* in comparison to man. Table 1 lists selected proteins participating in the function of endosomal-lysosomal system and their potential orthologs found by BLASTP (basic local alignment search tool) algorithm in the *C. elegans* genome, the evaluated parameter being the expectancy value (*e*-value).

*C. elegans* has already been utilized in numerous works dealing with issues regarding basic endosomal-lysosomal system functioning\(^{137-141}\). *C. elegans* has also been repetitively used as a good model organism for different human pathology states, metabolic disorders and lysosomal storage disorders included\(^{127,142}\). *C. elegans* was used in modeling of Niemann-Pick disease type C\(^{143}\) or neuronal ceroid lipofuscinosis type 3\(^{144}\). *C. elegans* was demonstrated to have two types of acid sphingomyelinase\(^{145}\), and to possess defined cathepsin D\(^{146}\) or LAMP/CD68\(^{147}\) like proteins. Pioneering character can be assigned to the works on *cup-5*\(^{148-152}\), *C. elegans* ortholog of mucolipin-1, a protein that is defective in one of the LSDs (mucolipidosis type IV).
<table>
<thead>
<tr>
<th>LSD</th>
<th>Variant human protein</th>
<th>e-value (expectancy)</th>
<th>C. elegans ORF</th>
<th>C. elegans gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fabry disease</td>
<td>alpha galactosidase</td>
<td>2.3e -87</td>
<td>R07B7.11</td>
<td>gana-1</td>
</tr>
<tr>
<td>Farber disease</td>
<td>acid ceramidase</td>
<td>2.1e -84</td>
<td>HK11D2.2</td>
<td></td>
</tr>
<tr>
<td>Gaucher disease</td>
<td>beta glucocerebrosidase</td>
<td>1.2e -106</td>
<td>Y4C8B.6</td>
<td></td>
</tr>
<tr>
<td>GM1 gangliosiosis</td>
<td>beta galactosidase</td>
<td>5.6e -116</td>
<td>T19B10.3</td>
<td></td>
</tr>
<tr>
<td>Tay-Sachs disease</td>
<td>alpha hexosaminidase</td>
<td>2.6e -95</td>
<td>T14F9.3</td>
<td></td>
</tr>
<tr>
<td>Sandhoff disease</td>
<td>beta hexosaminidase</td>
<td>9.1e -84</td>
<td>T14F9.3</td>
<td></td>
</tr>
<tr>
<td>Krabbe disease</td>
<td>galactocerebrosidase</td>
<td>2.1e -47</td>
<td>C29E4.10</td>
<td></td>
</tr>
<tr>
<td>Metachromatic</td>
<td>arylsulfatase A</td>
<td>1.5e -43</td>
<td>D1014.1</td>
<td></td>
</tr>
<tr>
<td>leukodystrophy</td>
<td>mucoolin</td>
<td>1.2e -79</td>
<td>R13A5.1</td>
<td></td>
</tr>
<tr>
<td>Niemann-Pick A,B</td>
<td>acid sphingomyelinase</td>
<td>2.0e -86</td>
<td>B0252.2</td>
<td></td>
</tr>
<tr>
<td>Schindler disease</td>
<td>N-acetylgalactosaminidase</td>
<td>3.7e -87</td>
<td>R07B7.11</td>
<td></td>
</tr>
<tr>
<td>Aspartylglycosaminuria</td>
<td>aspartylglycosaminidase</td>
<td>1.7e -68</td>
<td>R04B3.2</td>
<td></td>
</tr>
<tr>
<td>Fabry disease</td>
<td>alpha galactosidase</td>
<td>2.9e -96</td>
<td>W03G11.3</td>
<td></td>
</tr>
<tr>
<td>Galactosialdosis</td>
<td>cathepsin A</td>
<td>8.1e -92</td>
<td>F41C3.5</td>
<td></td>
</tr>
<tr>
<td>alpha-mannosidosis</td>
<td>alpha mannosidase</td>
<td>1.1e -157</td>
<td>F35G10.1</td>
<td></td>
</tr>
<tr>
<td>beta-mannosidosis</td>
<td>alpha mannosidase</td>
<td>5.8e -150</td>
<td>C33G3.4</td>
<td></td>
</tr>
<tr>
<td>MPSII</td>
<td>iduronate 2 sulfatase</td>
<td>2.7e -07</td>
<td>D1014.1</td>
<td></td>
</tr>
<tr>
<td>MPSV1</td>
<td>arylsulfatase B</td>
<td>4.6e -18</td>
<td>D1014.1</td>
<td></td>
</tr>
<tr>
<td>MPSVIA</td>
<td>galactosamin sulfate sulfatase</td>
<td>5.9e -26</td>
<td>D1014.1</td>
<td></td>
</tr>
<tr>
<td>MPSIIIA</td>
<td>sulfatase</td>
<td>2.5e -07</td>
<td>D1014.1</td>
<td></td>
</tr>
<tr>
<td>MPSIIIB</td>
<td>alpha N-acetyl glucosaminidase</td>
<td>4.5e -98</td>
<td>K99E4.4</td>
<td></td>
</tr>
<tr>
<td>MPSIIID</td>
<td>N-acetylgalactosaminidase</td>
<td>2.3e -91</td>
<td>K99E4.4</td>
<td></td>
</tr>
<tr>
<td>MPSVII</td>
<td>beta glucuronidase</td>
<td>6.1e -114</td>
<td>Y15E48.11</td>
<td></td>
</tr>
<tr>
<td>Pompéa disease</td>
<td>alpha glucosidase</td>
<td>2.7e -134</td>
<td>D299B3</td>
<td></td>
</tr>
<tr>
<td>Wolman disease</td>
<td>acid lipase</td>
<td>3.2e -81</td>
<td>Y57B72B.B</td>
<td></td>
</tr>
<tr>
<td>cystinosis</td>
<td>solute carrier family 3 member 1</td>
<td>2.0e -36</td>
<td>F2D10.9</td>
<td></td>
</tr>
<tr>
<td>Lysosomal storage</td>
<td>solute carrier family 17 member 5</td>
<td>2.1e -100</td>
<td>C38C10.2</td>
<td></td>
</tr>
<tr>
<td>disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCL1</td>
<td>palmitoyl-thioesterase 1</td>
<td>1.2E -81</td>
<td>F44C4.5</td>
<td>ctn-3.1, ctn-3.2, ctn-3.3</td>
</tr>
<tr>
<td>NCL3</td>
<td>CLN3 protein</td>
<td>7.6e -80</td>
<td>F07B10.1</td>
<td>ctn-3.3</td>
</tr>
<tr>
<td>congenital NCL</td>
<td>cathepsin D</td>
<td>1.5e+018</td>
<td>R12H7.2</td>
<td>cad-1</td>
</tr>
<tr>
<td>Niemann-Pick C1</td>
<td>NPC1 protein</td>
<td>8.7e -07</td>
<td>F92B6.6</td>
<td>imp-1</td>
</tr>
<tr>
<td>prosaposin</td>
<td>prosaposin</td>
<td>3.2e -09</td>
<td>C28C12.7</td>
<td></td>
</tr>
<tr>
<td>Structural protein</td>
<td>subunit</td>
<td>e-value (expectancy)</td>
<td>C. elegans ORF</td>
<td></td>
</tr>
<tr>
<td>rab4 A</td>
<td></td>
<td>7.2e -59</td>
<td>K99A9.2</td>
<td></td>
</tr>
<tr>
<td>rab4B</td>
<td></td>
<td>1.8e -60</td>
<td>K99A9.2</td>
<td></td>
</tr>
<tr>
<td>rab5A</td>
<td></td>
<td>3.1e -83</td>
<td>F26H9.5</td>
<td></td>
</tr>
<tr>
<td>rab7</td>
<td></td>
<td>3.3e -78</td>
<td>W03C9.3</td>
<td></td>
</tr>
<tr>
<td>rab9</td>
<td></td>
<td>3.1e -51</td>
<td>W03C9.3</td>
<td></td>
</tr>
<tr>
<td>rab5B</td>
<td></td>
<td>2.4e -83</td>
<td>F26H9.5</td>
<td></td>
</tr>
<tr>
<td>LAMP1</td>
<td></td>
<td>7.8e -08</td>
<td>C05D9.2</td>
<td></td>
</tr>
<tr>
<td>LAMP2</td>
<td></td>
<td>7.2e -06</td>
<td>C05D9.2</td>
<td></td>
</tr>
<tr>
<td>coatomer proteins</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td></td>
<td>0</td>
<td>Y71F9AL.17</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>0</td>
<td>F38E11.5</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td></td>
<td>3.8e -46</td>
<td>F45G2.4</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td>4.6e -128</td>
<td>C13B9.3</td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td></td>
<td>5.2e -49</td>
<td>F59E10.3</td>
<td></td>
</tr>
<tr>
<td>G2</td>
<td></td>
<td>8.6e -262</td>
<td>T14G10.5</td>
<td></td>
</tr>
<tr>
<td>B2</td>
<td></td>
<td>0</td>
<td>F38E11.5</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td></td>
<td>8.9e -260</td>
<td>T14G10.5</td>
<td></td>
</tr>
<tr>
<td>clathrin protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLTA</td>
<td></td>
<td>1.9e -09</td>
<td>T05H11.3</td>
<td></td>
</tr>
<tr>
<td>LTB</td>
<td></td>
<td>1.2e -06</td>
<td>T05H11.3</td>
<td></td>
</tr>
<tr>
<td>CLH1</td>
<td></td>
<td>0</td>
<td>T26G5.1</td>
<td></td>
</tr>
<tr>
<td>CLH2</td>
<td></td>
<td>0</td>
<td>T26G5.1</td>
<td></td>
</tr>
</tbody>
</table>
Table 1

List of selected human proteins involved in pathogenesis of lysosomal storage diseases or functioning in endosomal-lysosomal system.

*C. elegans* orthologs found by BLASTP algorithm in the nematode’s ORF database. The value of expectancy (*e*-value) in principle represents the probability that BLASTP result represents a random positive hit directly related to the size of the database. In other words, smaller the *e*-value (closer to zero) higher the probability that the two sequences are non-randomly related. Expectancy values are given in the order of 10 to the minus value given in the table.

Note near identity of certain structural proteins and near random relation for others (prosaposin, sulfamidase or iduronate-2-sulfatase). The example of prosaposin demonstrates the potential drawback of similarity thresholding algorithms, *C. elegans* most probably possesses pSap ortholog, but it must be disclosed on the level of secondary protein structure.

5.2 Comments and discussion to the publication - Hujová, Sikora, Dobrovolný et al. (2005)

The publication by Hujová, Sikora, Dobrovolný et al. (2005, equal contributors) describes the use of *C. elegans* in delineation of biochemical and molecular aspects of evolution of human lysosomal enzymes degrading moieties with terminal α-galactose and α-N-acetylgalactosamine. The respective enzymes are lysosomal α-galactosidase (α-GAL) and α-N-acetylactosaminidase (α-NAGA). Both α-GAL and α-NAGA belong to the melibiase protein family (glycoside hydrolase family 27, clan D), and they represent true paralogs. The overall similarity between α-GAL and α-NAGA is extensive. The evolutionary concept of protein duplication and additional functional divergence has been previously proposed, but never unambiguously demonstrated.

Enzymatic properties of both proteins, including substrate specificities, further support the theory of common evolutionary origin. α-GAL is active against substrates containing terminal α-galactose, while α-NAGA is participating in intralysosomal degradation of glycoconjugates containing terminal α-N-acetylgalactosaminoyl(α-galNAc) but partially also against moieties with terminal α-galactose.

Both enzymes are soluble lysosomal hydrolases with acid pH optima targeted to lysosomes by mannose-6-phosphate receptor mediated pathway. Hereditary deficiency of either α-GAL or α-NAGA activity results in Fabry or Schindler (or Kanzaki) diseases, respectively. Both these disorders represent distinct lysosomal storage entities with different phenotypic presentation. Biological material accumulating in α-GAL deficiency is almost exclusively GB3 ceramide in comparison to α-NAGA deficiency which results in the predominant storage of α-galNAc bearing glycoproteins.

Simple preliminary BLAST search for *C. elegans* orthologs of human α-GAL and α-NAGA, revealed only a single ORF with comparable *e*-values (Table 1). This ORF (R0B7.11), later designated as *gana-1* (α-GAL and α-NAGA ortholog) is the only ortholog of these two genes in the nematode’s genome.

The goal of the study was to evaluate and characterize *gana-1* evolutionary relations to α-GAL and α-NAGA as well as its functional properties in order to demonstrate that this single *C. elegans* ortholog adopts functions of both human (vertebrate) proteins.

5.2.1 *gana-1* evolutionary context

Prior to all experimental work *gana-1* gene structure was evaluated, both on the level of genomic DNA and cDNA (RNA), the *in-silico* prediction was confirmed by experimental data.
Multiple protein sequence alignment with other members of melibiase protein family disclosed extensive similarity among the aligned sequences and clearly delineated plant sequences from their animal counterparts. We used multiple bootstrapped tree analysis to construct the final evolutionary trees with following results: (i) gana-1 protein was always positioned between animal and plant sequences; (ii) it was located very close (closest) to primary tree node (animal vs. plant), suggesting its lowest evolutionary position among the evaluated animal sequences; (iii) when it was included into one of the sequence clusters (GALS vs. NAGAs), it was always among α-NAGAs, a fact that corresponded well with its further defined characteristics.

Following bioinformatic step was the construction of a structural model of gana-1 protein based on the on the crystal structures of rice and human α-GAL using homology modeling approaches. Our predicted structure suggests two-domain organization with homodimeric tertiary organization. The active site and the N-acetyl recognition loop of gana-1, as it was modeled, adopted a structure more similar to α-NAGA, suggesting its capability to utilize substrates of both types.

5.2.2 gana-1 enzymatic activity, RNAi induced compound enzymatic deficiency, cellular localization of gana-1

α-GAL and α-NAGA activity measurements in the mixed culture C. elegans homogenates showed presence of both activities, α-NAGA activity predominated over α-GAL. The addition of N-acetyl-D-galactosamine (α-NAGA inhibitor) resulted in the considerable activity drop of both (α-GAL and α-NAGA) enzymatic activities, which is a phenomenon undetectable in vertebrate (human) proteins. N-acetyl-D-galactosamine selectively inhibits α-NAGA but does not influence α-GAL activity. In comparison, the addition of D-galactose (α-GAL inhibitor) also resulted in the drop of both activities. Parallel and proportional drops of both enzymatic activities were observed in repetitive experiments after RNAi treatment. The induced activity decreases were in the order of tens of per cents in comparison to controls.

Despite RNAi mediated introduction of considerable enzymatic deficiency in the nematodes, we were not able to document specific morphological changes (lysosomal storage) associated with either of the phenotypes. This could be explained by relatively high residual enzymatic activity after RNAi. This situation is reminiscent of LSD heterozygotes status, when despite having enzymatic activities as low as 10% of controls these individuals do not suffer from lysosomal storage.

In order to elucidate the cellular functions of gana-1 we performed C terminal EGFP (enhanced green fluorescent protein) tagging of gana-1. gana-1::GFP transgene was introduced into the standard Bristol N2 C. elegans strain and the expression pattern was studied in vivo. We were able to document GFP presence in the vacuolar compartment of subset of six C. elegans cells with high endocytic turnover called coelomocytes.

The results and conclusions of this publication clearly demonstrated usefulness of C. elegans for the study of human pathology states. Suggested evolutionary scheme of the two paralogous genes/proteins may represent important premise for future design of either enzymatic assays setups or drugs. The integrative conjunction of bioinformatic methods and experiments in easily genetically modifiable animal model generated results that would not be obtainable in other (mammalian) species in the same time. The fact that we were not able to replicate the morphological storage phenotype does not lower the results; it just demonstrates the limits of RNAi. As can be demonstrated on some other studies using C. elegans as a
model for human diseases, lack of comparable phenotypic presentation does not necessarily mean lack of important and applicable biological message.

6 Concluding notes

The last 40 years were dedicated to the studies on substantial monogenic bases of lysosomal storage disorders. Present time in the LSD research represents a transitory step between the hypothesis-driven and curiosity-driven experimental approaches, which are gaining more and more attention. Today’s LSD research could thus be characterized by intensive demand for integration of the available data with outcomes towards pathogenic mechanisms, novel diagnostic approaches or therapeutic modalities.

Purpose of this section of the thesis is to provide final summarizing overview of the performed and published findings. Reader is refered to the previous detailed discussions on the presented topics. The publications represent novel findings in the area of saposin-like proteins related lysosomal storage disorders. The reported “non-enzymatic” functions of pSap are integral part of pSap deficiency pathogenesis and are in complete accordance with the recently opened topics of protein clearance and autophagy related cell impairment in LSDs. In a closely related manner, HGSNAT protein (deficient in MPS IIIc), which resides in lysosomal DRMs, is most probably directly or indirectly involved in chaperon mediated autophagy and implications of its function might be relevant to a broader spectrum of LSDs.

Presented genotype-phenotype correlations in ASM (another Sap-like protein) deficiency rendered important data towards ASM catalytic functions and re-opened debate on phenotypic diversity and Niemann-Pick disease type A and B classification. These data should not be perceived as mere pathogenic variation reports as they contain broad pathogenic considerations and implications.

Last but not least, *C. elegans* was demonstrated as relevant model organism for lysosomal enzymopathies. The data demonstrating evolution of two related lysosomal hydrolases could have never been achieved with such an ease in mammalian model. These findings also point to sometimes overlooked neccessity to study the evolutionary relations in order to answer questions of human pathology.
# 7 Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D</td>
<td>three dimensional</td>
</tr>
<tr>
<td>α-GAL</td>
<td>alpha-galactosidase</td>
</tr>
<tr>
<td>α-NAGA</td>
<td>alpha-N-acetylgalactosaminidase</td>
</tr>
<tr>
<td>ASM</td>
<td>acid sphingomyelinase</td>
</tr>
<tr>
<td>BLASTN/P</td>
<td>basic local alignment search tool nucleotide/protein</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>BWS</td>
<td>Beckwith-Wiedemann syndrome</td>
</tr>
<tr>
<td>C.elegans</td>
<td>Caenorhabditis elegans</td>
</tr>
<tr>
<td>CCF</td>
<td>cross-correlation function</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary deoxyribonucleic acid</td>
</tr>
<tr>
<td>CMA</td>
<td>chaperon mediated autophagy</td>
</tr>
<tr>
<td>CNS</td>
<td>conserved non-coding sequence</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DRM</td>
<td>detergent resistant microdomain</td>
</tr>
<tr>
<td>EST</td>
<td>expressed sequence tag</td>
</tr>
<tr>
<td>e-value</td>
<td>expectancy value</td>
</tr>
<tr>
<td>FFPE</td>
<td>formaldehyde fixed paraffin embedded</td>
</tr>
<tr>
<td>gana-1</td>
<td>α-GAL and α-NAGA ortholog</td>
</tr>
<tr>
<td>GFAP</td>
<td>glial fibrillary acidic protein</td>
</tr>
<tr>
<td>HGSNAT</td>
<td>heparin acetyl-coenzyme A:α-glucosaminide N-acetyltransferase</td>
</tr>
<tr>
<td>HMU-PC</td>
<td>6-hexadecanoylamino-4-methylumbelliferyl-phosphorylcholine</td>
</tr>
<tr>
<td>HNP-PC</td>
<td>2-N-(hexadecanoyl)-amino-4-nitrophenyl phosphorylcholine</td>
</tr>
<tr>
<td>kb</td>
<td>kilo base</td>
</tr>
<tr>
<td>kDa</td>
<td>kiloDalton</td>
</tr>
<tr>
<td>LAMP</td>
<td>lysosome associated membrane protein</td>
</tr>
<tr>
<td>LC3</td>
<td>microtubule-associated protein 1 light chain 3</td>
</tr>
<tr>
<td>LSD</td>
<td>lysosomal storage disorder</td>
</tr>
<tr>
<td>M-6-P</td>
<td>mannose-6-phosphate</td>
</tr>
<tr>
<td>M-6-PR</td>
<td>mannose-6-phosphate receptor</td>
</tr>
<tr>
<td>Mb</td>
<td>mega base</td>
</tr>
<tr>
<td>MIM</td>
<td>mendelian inheritance in man</td>
</tr>
<tr>
<td>MPS</td>
<td>mucopolysaccharidosis</td>
</tr>
<tr>
<td>MR</td>
<td>mannose receptor</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger ribonucleic acid</td>
</tr>
<tr>
<td>NCL</td>
<td>neuronal ceroid lipofuscinosis</td>
</tr>
<tr>
<td>NPC</td>
<td>Niemann-Pick disease type C</td>
</tr>
<tr>
<td>NPD</td>
<td>Niemann-Pick disease</td>
</tr>
<tr>
<td>NPD A/B</td>
<td>Niemann-Pick disease type A/B</td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
</tr>
<tr>
<td>OST</td>
<td>ORF sequence tag</td>
</tr>
<tr>
<td>PCD</td>
<td>programmed cell death</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PPCA</td>
<td>protective protein / Cathepsin A</td>
</tr>
<tr>
<td>pSap</td>
<td>prosaposin</td>
</tr>
<tr>
<td>PTM</td>
<td>post-translational modification</td>
</tr>
<tr>
<td>RNAi</td>
<td>RNA mediated interference</td>
</tr>
<tr>
<td>Sap</td>
<td>saposin</td>
</tr>
<tr>
<td>SAPLIP</td>
<td>saposin-like protein</td>
</tr>
<tr>
<td>SCMAS</td>
<td>subunit c of mitochondrial ATP synthase</td>
</tr>
<tr>
<td>SMDP</td>
<td>sphingomyelin phosphodiesterase</td>
</tr>
<tr>
<td>TM</td>
<td>transmembrane</td>
</tr>
<tr>
<td>TMMOD</td>
<td>improved transmembrane hidden Markov model</td>
</tr>
</tbody>
</table>
8 References

1. Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication - http://www.icmje.org/


99. The Human Gene Mutation Database [http://www.hgmd.cf.ac.uk](http://www.hgmd.cf.ac.uk)


106. PDB - Protein Data Bank - [http://www.rcsb.org/pdb](http://www.rcsb.org/pdb)


133. Wormbase - http://www.wormbase.org/
155. Wormbase BLAST or BLAT Search - http://www.wormbase.org/db/searches/blat
9 List of author's publications and presentations

9.1 Publications related to the thesis (sorted chronologically, IF 2006):

IF - 3.192

cited by:
Ricci V, et al. Hum Mutat. 2004
Pittis MG, et al. Hum Mutat. 2004

* contributed equally
IF - 2.652

cited by:

IF - 1.722

cited by:

IF - 1.722

cited by:

IF - 12.649

IF - 2.527
9.2 Other publications (sorted chronologically, IF 2006)


cited by:


9.3 Published abstracts

Sikora J, Perereira H, Elleder M

Sikora J, Perereira H, Elleder M, Wevers RA


9.4 Other selected presentations

Pavlů-Pereira H, Sikora J and Elleder M
Mutational analysis of Czech and Slovak patients with acid sphingomyelinase deficiency.
12th ESGLD workshop, September 1999, Vidago, Portugal

Elleder M, Pavlů-Perreira H, Sikora J, Asfaw B, Poupětová H, Berná L
Sphingomyelinase deficiency- genotype and multilevel phenotype study of 25 Czech and Slovak patients
13th ESGLD workshop, September 2001, Woudschoten, Netherlands

Caenorhabditis elegans as a model for lysosomal glycosidase deficiencies
18. pracovní dny Dědičné Metabolické Poruchy, May 2003, Slušovice, Czech Republic
First price for the scientists under 30 years of age

Caenorhabditis elegans as a model for acid α−glucosidase deficiency
14th ESGLD Workshop, September 2003, Poděbrady, Czech Republic

Caenorhabditis elegans as a model for Fabry and Schindler diseases
14th ESGLD Workshop, September 2003, Poděbrady, Czech Republic

Caenorhabditis elegans has only one alpha galactosidase and alpha N-acetylgalactosaminidase ortholog
5th students’ scientific conference 1.LF UK, April 2004, Prague, Czech Republic
1st price (PhD students’ section)

Elleder M, Sikora J, Harzer K
Neurolysosomes in human prosaposin deficiency
15th ESGLD Workshop, September 2005, Oslo, Norway

Sikora J, Elleder M
Vacuolar autophagic disorder with predominant involvement of liver, heart, and smooth muscle. Accidental autopsy finding in an elderly patient.
15th ESGLD Workshop, September 2005, Oslo, Norway

Úskalí sekvenční analýzy DNA z formolem fixoványh tkání: praktické aplikace na příkladu familiální formy Creutzfeldt-Jakobovy choroby a netypických forem Niemann-Pickovy choroby typu C.
Molekulární patologie začátků III. tisíciletí: využití bioppticích vzorků pro molekulární analýzu, April 2005, Olomouc, Czech Republic

Sikora J, Elleder M
A case of excessive autophagy with multiorgan involvement and low clinical penetrance. A novel lysosomal disorder?
33rd Congress of the Czech Society of Pathologists and 2nd Satellite Symposium and Workshop on Molecular Pathology, May 2006, Olomouc, Czech Republic